Lectures on Differential Geometry
Math 240BC

John Douglas Moore
Department of Mathematics
University of California
Santa Barbara, CA, USA 93106
e-mail: moore@math.ucsb.edu

June 5, 2009

Preface

This is a set of lecture notes for the course Math 240BC given during the Winter and Spring of 2009. The notes evolved as the course progressed and are still somewhat rough, but we hope they are helpful. Starred sections represent digressions are less central to the core subject matter of the course and can be omitted on a first reading.

Our goal was to present the key ideas of Riemannian geometry up to the generalized Gauss-Bonnet Theorem. The first chapter provides the foundational results for Riemannian geometry. The second chapter provides an introduction to de Rham cohomology, which provides prehaps the simplest introduction to the notion of homology and cohomology that is so pervasive in modern geometry and topology. In the third chapter we provide some of the basic theorem relating the curvature to the topology of a Riemannian manifold—the idea here is to develop some intuition for curvature. Finally in the fourth chapter we describe Cartan's method of moving frames and focus on its application to one of the key theorems in Riemannian geometry, the generalized Gauss-Bonnet Theorem.

The last chapter is more advanced in nature and not usually treated in the first-year differential geometry course. It provides an introduction to the theory of characteristic classes, explaining how these could be generated by looking for extensions of the generalized Gauss-Bonnet Theorem, and describes applications of characteristic classes to the Atiyah-Singer Index Theorem and to the existence of exotic differentiable structures on seven-spheres.

Contents

1 Riemannian geometry **2**
1.1 Review of tangent and cotangent spaces . . . 2
1.2 Riemannian metrics . . . 5
1.3 Geodesics . . . 9
1.3.1 Smooth paths . . . 9
1.3.2 Piecewise smooth paths . . . 13
1.4 Hamilton's principle . . . 14
1.5 The Levi-Civita connection . . . 20
1.6 First variation of J . . . 24
1.7 Lorentz manifolds . . . 27
1.8 The Riemann-Christoffel curvature tensor . . . 30
1.9 Curvature symmetries; sectional curvature . . . 36
1.10 Gaussian curvature of surfaces . . . 39
1.11 Review of Lie groups . . . 42
1.12 Lie groups with biinvariant metrics . . . 45
1.13 Grassmann manifolds . . . 49
1.14 The exponential map . . . 53
1.15 The Gauss Lemma . . . 56
1.16 Curvature in normal coordinates . . . 57
1.17 Riemannian manifolds as metric spaces . . . 60
1.18 Completeness . . . 62
1.19 Smooth closed geodesics . . . 64

2 Differential forms **68**
2.1 Tensor algebra . . . 68
2.2 The exterior derivative . . . 71
2.3 Integration of differential forms . . . 75
2.4 Theorem of Stokes . . . 80
2.5 de Rham Cohomology . . . 83
2.6 Poincaré Lemma . . . 85
2.7 Mayer-Vietoris Sequence . . . 90
2.8 Singular homology* . . . 94
2.8.1 Definition of singular homology* . . . 94
2.8.2 Singular cohomology* . . . 98

2.8.3 Proof of the de Rham Theorem* 102
2.9 The Hodge star 106
2.10 The Hodge Laplacian 112
2.11 The Hodge Theorem 116
2.12 d and δ in terms of moving frames 119
2.13 The rough Laplacian 122
2.14 The Weitzenböck formula 123
2.15 Ricci curvature and Hodge theory 125
2.16 The curvature operator and Hodge theory 126
2.17 Proof of Gallot-Meyer Theorem* 128

3 Curvature and topology **131**
3.1 The Hadamard-Cartan Theorem 131
3.2 Parallel transport along curves 133
3.3 Geodesics and curvature 134
3.4 Proof of the Hadamard-Cartan Theorem 138
3.5 The fundamental group 140
3.5.1 Definition of the fundamental group 140
3.5.2 Homotopy lifting 142
3.5.3 Universal covers 145
3.6 Uniqueness of simply connected space forms 147
3.7 Non simply connected space forms 149
3.8 Second variation of action 151
3.9 Myers' Theorem 153
3.10 Synge's Theorem 155

4 Cartan's method of moving frames **159**
4.1 An easy method for calculating curvature 159
4.2 The curvature of a surface 163
4.3 The Gauss-Bonnet formula for surfaces 167
4.4 Application to hyperbolic geometry 171
4.5 Vector bundles 177
4.6 Connections on vector bundles 179
4.7 Metric connections 183
4.8 Curvature of connections 185
4.9 The pullback construction 187
4.10 Classification of connections in complex line bundles 189
4.11 Classification of $U(1)$-bundles* 195
4.12 The Pfaffian 196
4.13 The generalized Gauss-Bonnet Theorem 197
4.14 Proof of the generalized Gauss-Bonnet Theorem 200

5 Characteristic classes **208**
5.1 The Chern character 208
5.2 Chern classes 213
5.3 Examples of Chern classes 215
5.4 Invariant polynomials 217
5.5 The universal bundle* 220
5.6 The Clifford algebra 225
5.7 The spin group 232
5.8 Spin structures and spin connections 235
5.9 The Dirac operator 240
5.10 The Atiyah-Singer Index Theorem 243
5.10.1 Index of the Dirac operator 243
5.10.2 Spin^c structures* 246
5.10.3 Dirac operators on general manifolds* 248
5.10.4 Topological invariants of four-manifolds* 252
5.11 Exotic spheres* 253

Bibliography **257**

Chapter 1

Riemannian geometry

1.1 Review of tangent and cotangent spaces

We will assume some familiarity with the theory of smooth manifolds, as presented, for example, in the first four chapters of [5].

Suppose that M is a smooth manifold and $p \in M$, and that $\mathcal{F}(p)$ denotes the space of pairs (U, f) where U is an open subset of M containing p and $f : U \to \mathbb{R}$ is a smooth function. If $\phi = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a smooth coordinate system on M with $p \in U$, and $(U, f) \in \mathcal{F}(p)$, we define

$$\left.\frac{\partial}{\partial x^i}\right|_p (f) = D_i(f \circ \phi^{-1})(\phi(p)) \in \mathbb{R},$$

where D_i denotes differentiation with respect to the i-th component. We thereby obtain an $\mathbb{R}$-linear map

$$\left.\frac{\partial}{\partial x^i}\right|_p : \mathcal{F}(p) \longrightarrow \mathbb{R},$$

called a *directional derivative operator*, which satisfies the Leibniz rule,

$$\left.\frac{\partial}{\partial x^i}\right|_p (fg) = \left(\left.\frac{\partial}{\partial x^i}\right|_p (f) \right) g(p) + f(p) \left(\left.\frac{\partial}{\partial x^i}\right|_p (g) \right),$$

and in addition depends only on the "germ" of f at p,

$$f \equiv g \quad \text{on some neighborhood of } p \quad \Rightarrow \left.\frac{\partial}{\partial x_i}\right|_p (f) = \left.\frac{\partial}{\partial x_i}\right|_p (g).$$

The set of all linear combinations

$$\sum_{i=1}^{n} a^i \left.\frac{\partial}{\partial x^i}\right|_p$$

of these basis vectors comprises the *tangent space* to M at p and is denoted by T_pM. Thus for any given smooth coordinate system $(x^1, \ldots, x^n)$ on M, we have a corresponding basis

$$\left(\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right)$$

for the tangent space T_pM.

The notation we have adopted makes it easy to see how the components (a^i) of a tangent vector transform under change of coordinates. If $\psi = (y^1, \ldots, y^n)$ is a second smooth coordinate system on M, the new basis vectors are related to the old by the chain rule,

$$\left.\frac{\partial}{\partial y^i}\right|_p = \sum_{j=1}^n \frac{\partial x^j}{\partial y^i}(p) \left.\frac{\partial}{\partial x^j}\right|_p, \quad \text{where} \quad \frac{\partial x^j}{\partial y^i}(p) = D_i(x^j \circ \psi^{-1})(\psi(p)).$$

The disjoint union of all of the tangent spaces forms the *tangent bundle*

$$TM = \bigcup \{T_pM : p \in M\},$$

which has a projection $\pi : TM \to M$ defined by $\pi(T_pM) = p$. If $\phi = (x^1, \ldots, x^n)$ is a coordinate system on $U \subset M$, we can define a corresponding coordinate system

$$\tilde{\phi} = (x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n) \quad \text{on} \quad \pi^{-1}(U) \subset TM$$

by letting

$$x^i \left(\sum_{j=1}^n a^j \left.\frac{\partial}{\partial x^j}\right|_p \right) = x^i(p), \qquad \dot{x}^i \left(\sum_{j=1}^n a^j \left.\frac{\partial}{\partial x^j}\right|_p \right) = a^i. \tag{1.1}$$

For the various choices of charts (U, ϕ), the corresponding charts $(\pi^{-1}(U), \tilde{\phi})$ form an atlas making TM into a smooth manifold of dimension $2n$, as you saw in Math 240A.

The *cotangent space* to M at p is simply the dual space T_p^*M to T_pM. Thus an element of T_p^*M is simply a linear map

$$\alpha : T_pM \longrightarrow \mathbb{R}.$$

Corresponding to the basis

$$\left(\left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right)$$

of T_pM is the dual basis

$$\left(dx^1|_p, \ldots, dx^n|_p \right), \quad \text{defined by} \quad dx^i|_p \left(\left.\frac{\partial}{\partial x^j}\right|_p \right) = \delta^i_j = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

The elements of T_p^*M, called cotangent vectors, are just the linear combinations of these basis vectors

$$\sum_{i=1}^n a_i dx^i|_p$$

Once again, under change of coordinates the basis elements transform by the chain rule,

$$dy^i|_p = \sum_{j=1}^n \frac{\partial y^i}{\partial x^j}(p) dx^j|_p.$$

An important example of cotangent vector is the differential of a function at a point. If $p \in U$ and $f : U \to \mathbb{R}$ is a smooth function, then the *differential* of f at p is the element $df|_p \in T_p^*M$ defined by $df|_p(v) = v(f)$. If $(x^1, \ldots, x^n)$ is a smooth coordinate system defined on U, then

$$df|_p = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(p) dx^i|_p.$$

Just as we did for tangent spaces, we can take the disjoint union of all of the cotangent spaces forms the *cotangent bundle*

$$T^*M = \bigcup \{T_p^*M : p \in M\},$$

which has a projection $\pi : TM \to M$ defined by $\pi(T_pM) = p$. If $\phi = (x^1, \ldots, x^n)$ is a coordinate system on $U \subset M$, we can define a corresponding coordinate system

$$\tilde{\phi} = (x^1, \ldots, x^n, p_1, \ldots, p_n) \quad \text{on} \quad \pi^{-1}(U) \subset TM$$

by letting

$$x^i \left(\sum_{j=1}^n a^j \left. \frac{\partial}{\partial x^j} \right|_p \right) = x^i(p), \qquad p_i \left(\sum_{j=1}^n a_j dx^j|_p \right) = a_i.$$

For the various choices of charts (U, ϕ), the corresponding charts $(\pi^{-1}(U), \tilde{\phi})$ form an atlas making T^*M into a smooth manifold of dimension $2n$.

We can generalize this construction and consider *tensor products* of tangent and cotangent spaces. For example, the tensor product of the cotangent space with itself, denoted by $\otimes^2 T_p^*M$, is the linear space of bilinear maps

$$g : T_pM \times T_pM \longrightarrow \mathbb{R}.$$

If $\phi = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a smooth coordinate system on M with $p \in U$, we can define

$$dx^i|_p \otimes dx^j|_p : T_pM \times T_pM \longrightarrow \mathbb{R}$$

$$\text{by} \quad dx^i|_p \otimes dx^j|_p \left(\left. \frac{\partial}{\partial x^k} \right|_p, \left. \frac{\partial}{\partial x^l} \right|_p \right) = \delta^i_k \delta^j_l.$$

Then

$$\left\{dx^i|_p \otimes dx^j|_p : 1 \le i \le n, 1 \le j \le n\right\}$$

is a basis for $\otimes^2 T_p^* M$, and a typical element of $\otimes^2 T_p^* M$ can be written as

$$\sum_{i,j=1}^{n} g_{ij}(p) dx^i|_p \otimes dx^j|_p,$$

where the $g_{ij}(p)$'s are elements of $\mathbb{R}$.

1.2 Riemannian metrics

Definition. Let M be a smooth manifold. A *Riemannian metric* on M is a function which assigns to each $p \in M$ a (positive-definite) inner product $\langle \cdot, \cdot \rangle_p$ on $T_p M$ which "varies smoothly" with $p \in M$. A *Riemannian manifold* is a pair $(M, \langle \cdot, \cdot \rangle)$ consisting of a smooth manifold M together with a Riemannian metric $\langle \cdot, \cdot \rangle$ on M.

Of course, we have to explain what we mean by "vary smoothly." This is most easily done in terms of local coordinates. If $\phi = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a smooth coordinate system on M, then for each choice of $p \in U$, we can write

$$\langle \cdot, \cdot \rangle_p = \sum_{i,j=1}^{n} g_{ij}(p) dx^i|_p \otimes dx^j|_p.$$

We thus obtain functions $g_{ij} : U \to \mathbb{R}$. To say that $\langle \cdot, \cdot \rangle_p$ varies smoothly with p simply means that the functions g_{ij} are smooth. We call the functions g_{ij} the *components* of the metric.

Note that the functions g_{ij} satisfy the symmetry condition $g_{ij} = g_{ji}$ and the condition that the matrix (g_{ij}) be positive definite. We will sometimes write

$$\langle \cdot, \cdot \rangle|_U = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j.$$

If $\psi = (y^1, \ldots, y^n)$ is a second smooth coordinate system on $V \subseteq M$, with

$$\langle \cdot, \cdot \rangle|_V = \sum_{i,j=1}^{n} h_{ij} dy^i \otimes dy^j,$$

it follows from the chain rule that, on $U \cap V$,

$$g_{ij} = \sum_{k.l=1}^{n} h_{kl} \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j}.$$

We will sometimes adopt the Einstein summation convention and leave out the summation sign:

$$g_{ij} = h_{kl} \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j}.$$

We remark in passing that this is how a "covariant tensor field of rank two" transforms under change of coordinates.

Using a Riemannian metric, one can "lower the index" of a tangent vector at p, producing a corresponding cotangent vector and vice versa. Indeed, if $v \in T_pM$, we can construct a corresponding cotangent vector α_v by the formla

$$\alpha_v(w) = \langle v, w \rangle_p.$$

In terms of components,

$$\text{if} \quad v = \sum_{i=1}^{n} a^i \left. \frac{\partial}{\partial x^i} \right|_p, \quad \text{then} \quad \alpha_v = \sum_{i,j=1}^{n} g_{ij}(p) a^j dx^i|_p.$$

Similarly, given a cotangent vector $\alpha \in T_p^*M$ we "raise the index" to obtain a corresponding tangent vector $v_\alpha \in T_pM$. In terms of components,

$$\text{if} \quad \alpha = \sum_{i=1}^{n} a_i dx^i|_p, \quad \text{then} \quad v_\alpha = \sum_{i,j=1}^{n} g^{ij}(p) a_j \left. \frac{\partial}{\partial x^i} \right|_p,$$

where $(g^{ij}(p))$ is the matrix inverse to $(g_{ij}(p))$. Thus a Riemannian metric transforms the differential $df|_p$ of a function to a tangent vector

$$\operatorname{grad}(f)(p) = \sum_{i,j=1}^{n} g^{ij}(p) \frac{\partial f}{\partial x^j}(p) \left. \frac{\partial}{\partial x^i} \right|_p,$$

called the *gradient* of f at p. Needless to say, in elementary several variable calculus this raising and lowering of indices is done all the time using the usual Euclidean dot product as Riemannian metric.

Example 1. Indeed, the simplest example of a Riemannian manifold is n-dimensional *Euclidean space* $\mathbb{E}^n$, which is simply $\mathbb{R}^n$ together with its standard rectangular cartesian coordinate system $(x^1, \ldots, x^n)$, and the Euclidean metric

$$\langle \cdot, \cdot \rangle_E = dx^1 \otimes dx^1 + \cdots + dx^n \otimes dx^n.$$

In this case, the components of the metric are simply

$$g_{ij} = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

We will often think of the Euclidean metric as being defined by the dot product,

$$\left\langle \sum_{i=1}^{n} a^i \left. \frac{\partial}{\partial x^i} \right|_p, \sum_{j=1}^{n} b^j \left. \frac{\partial}{\partial x^j} \right|_p \right\rangle = \left(\sum_{i=1}^{n} a^i \left. \frac{\partial}{\partial x^i} \right|_p \right) \cdot \left(\sum_{j=1}^{n} b^j \left. \frac{\partial}{\partial x^j} \right|_p \right) = \sum_{i=1}^{n} a^i b^i.$$

Example 2. Suppose that M is an n-dimensional smooth manifold and that $F : M \to \mathbb{R}^N$ is a smooth imbedding. We can give $\mathbb{R}^N$ the Euclidean metric defined in the preceding example. For each choice of $p \in M$, we can then define an inner product $\langle \cdot, \cdot \rangle_p$ on T_pM by

$$\langle v, w \rangle_p = F_{*p}(v) \cdot F_{*p}(w), \quad \text{for } v, w \in T_pM.$$

Here F_{*p} is the differential of F at p defined in terms of a smooth coordinate system $\phi = (x^1, \ldots, x^n)$ by the explicit formula

$$F_{*p}\left(\left. \frac{\partial}{\partial x^i} \right|_p \right) = D_i(F \circ \phi^{-1})(\phi(p)) \in \mathbb{R}^N.$$

Clearly, $\langle v, w \rangle_p$ is symmetric, and it is positive definite because F is an immersion. Moreover,

$$\begin{aligned} g_{ij}(p) = \left\langle \left. \frac{\partial}{\partial x^i} \right|_p, \left. \frac{\partial}{\partial x^j} \right|_p \right\rangle &= F_{*p}\left(\left. \frac{\partial}{\partial x^i} \right|_p \right) \cdot F_{*p}\left(\left. \frac{\partial}{\partial x^j} \right|_p \right) \\ &= D_i(F \circ \phi^{-1})(\phi(p)) \cdot D_j(F \circ \phi^{-1})(\phi(p)), \end{aligned}$$

so $g_{ij}(p)$ depends smoothly on p. Thus the imbedding F induces a Riemannian metric $\langle \cdot, \cdot \rangle$ on M which we call the *induced metric*, and we write

$$\langle \cdot, \cdot \rangle = F^* \langle \cdot, \cdot \rangle_E.$$

It is an interesting fact that this construction includes *all* Riemannian manifolds.

Definition. Let $(M, \langle \cdot, \cdot \rangle)$ be a Riemannian manifold, and suppose that $\mathbb{E}^N$ denotes $\mathbb{R}^N$ with the Euclidean metric. An imbedding $F : M \to \mathbb{E}^N$ is said to be *isometric* if $\langle \cdot, \cdot \rangle = F^* \langle \cdot, \cdot \rangle_E$.

Nash's Imbedding Theorem *If $(M, \langle \cdot, \cdot \rangle)$ is any smooth Riemannian manifold, there exists an isometric imbedding $F : M \to \mathbb{E}^N$ into some Euclidean space.*

This was regarded as a landmark theorem when it first appeared [28]. The proof is difficult, involves subtle techniques from the theory of nonlinear partial differential equations, and is beyond the scope of this course.

A special case of Example 2 consists of two-dimensional smooth manifolds which are imbedded in $\mathbb{E}^3$. These are usually called *smooth surfaces* in $\mathbb{E}^3$ and are studied extensively in undergraduate courses in "curves and surfaces." This subject was extensively developed during the nineteenth century and was summarized in 1887-96 in a monumental four-volume work, *Leçons sur la théorie générale des surfaces et les applications géométriques du calcul infinitésimal*, by Jean Gaston Darboux. Indeed, the theory of smooth surfaces in $\mathbb{E}^3$ still provides much geometric intuition regarding Riemannian geometry of higher dimensions.

What kind of geometry does a Riemannian metric provide a smooth manifold M? Well, to begin with, we can use a Riemannian metric to define the lengths of tangent vectors. If $v \in T_pM$, we define the length of v by the formula

$$\|v\| = \sqrt{\langle v, v\rangle_p}.$$

Second, we can use the Riemannian metric to define angles between vectors: The angle θ between two nonzero vectors $v, w \in T_pM$ is the unique $\theta \in [0, \pi]$ such that

$$\langle v, w\rangle_p = \|v\|\|w\| \cos\theta.$$

Third, one can use the Riemannian metric to define lengths of curves. Suppose that $\gamma : [a, b] \to M$ is a smooth curve with velocity vector

$$\gamma'(t) = \sum_{i=1}^{n} \frac{dx^i}{dt} \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)} \in T_{\gamma(t)}M, \quad \text{for } t \in [a, b].$$

Then the *length* of γ is given by the integral

$$L(\gamma) = \int_a^b \sqrt{\langle \gamma'(t), \gamma'(t)\rangle_{\gamma(t)}}dt.$$

We can also write this in local coordinates as

$$L(\gamma) = \int_a^b \sqrt{\sum_{i,j=1}^{n} g_{ij}(\gamma(t))\frac{dx^i}{dt}\frac{dx^j}{dt}}dt.$$

Note that if $F : M \to \mathbb{E}^N$ is an isometric imbedding, then $L(\gamma) = L(F \circ \gamma)$. Thus the lengths of a curve on a smooth surface in $\mathbb{E}^3$ is just the length of the corresponding curve in $\mathbb{E}^3$. Since any Riemannian manifold can be isometrically imbedded in some $\mathbb{E}^N$, one might be tempted to try to study the Riemannian geometry of M via the Euclidean geometry of the ambient Euclidean space. However, this is not necessarily an efficient approach, since sometimes the isometric imbedding is quite difficult to construct.

Example 3. Suppose that $\mathbb{H}^2 = \{(x, y) \in \mathbb{R}^2 : y > 0\}$, with Riemannian metric

$$\langle \cdot, \cdot\rangle = \frac{1}{y^2}(dx \otimes dx + dy \otimes dy).$$

A celebrated theorem of David Hilbert states that $(\mathbb{H}^2, \langle \cdot, \cdot\rangle)$ has no isometric imbedding in $\mathbb{E}^3$ and although isometric imbeddings in Euclidean spaces of higher dimension can be constructed, none of them is easy to describe. The Riemannian manifold $(\mathbb{H}^2, \langle \cdot, \cdot\rangle)$ is called the *Poincaré upper half-plane*, and figures prominently in the theory of Riemann surfaces.

1.3 Geodesics

Our first goal is to generalize concepts from Euclidean geometry to Riemannian geometry. One of principal concepts in Euclidean geometry is the notion of straight line. What is the analog of this concept in Riemannian geometry? One candidate would be the curve between two points in a Riemannian manifold which has shortest length (if such a curve exists).

1.3.1 Smooth paths

Suppose that p and q are points in the Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$. If a and b are real numbers with $a < b$, we let

$$\Omega_{[a,b]}(M; p, q) = \{ \text{ smooth paths } \gamma : [a, b] \to M : \gamma(a) = p, \gamma(b) = q\}.$$

We can define two functions $L, J : \Omega_{[a,b]}(M; p, q) \to \mathbb{R}$ by

$$L(\gamma) = \int_a^b \sqrt{\langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)}} dt, \qquad J(\gamma) = \frac{1}{2} \int_a^b \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)} dt.$$

Although our goal is to understand the *length* L, it is often convenient to study this by means of the closely related *action* J. Notice that L is invariant under reparametrization of γ, so once we find a single curve which minimizes L we have an infinite-dimensional family. This, together with the fact that the formula for L contains a troublesome radical in the integrand, make J far easier to work with than L.

It is convenient to regard J as a smooth function on the "infinite-dimensional" manifold $\Omega_{[a,b]}(M; p, q)$. At first, we use the notion of infinite-dimensional manifold somewhat informally, but later we will return to make the notion precise.

Proposition 1. $[L(\gamma)]^2 \leq 2(b - a)J(\gamma)$. *Moreover, equality holds if and only if* $\langle \gamma'(t)\gamma'(t) \rangle$ *is constant if and only if* γ *has constant speed.*

Proof: We use the Cauchy-Schwarz inequality:

$$\begin{aligned} L(\gamma)]^2 &= \left[\int_a^b \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt \right]^2 \\ &\leq \left[\int_a^b dt \right] \left[\int_a^b \langle \gamma'(t), \gamma'(t) \rangle dt \right] = 2(b - a)J(\gamma). \end{aligned} \tag{1.2}$$

Equality holds if and only if the functions $\langle \gamma'(t), \gamma'(t) \rangle$ and 1 are linearly dependent, that is, if and only if γ has constant speed.

Proposition 2. *Suppose that* M *has dimension at least two. An element* $\gamma \in \Omega_{[a,b]}(M; p, q)$ *minimizes* J *if and only if it minimizes* L *and has constant speed.*

Sketch of proof: One direction is easy. Suppose that γ has constant speed and minimizes L. Then, if $\lambda \in \Omega_{[a,b]}(M;p,q)$,

$$2(b-a)J(\gamma) = [L(\gamma)]^2 \leq [L(\lambda)]^2 \leq 2(b-a)J(\lambda),$$

and hence $J(\gamma) \leq J(\lambda)$.

We will only sketch the proof of the other direction for now; later a complete proof will be available. Suppose that γ minimizes J, but does not minimize L, so there is $\lambda \in \Omega$ such that $L(\lambda) < L(\gamma)$. Approximate λ by an immersion λ_1 such that $L(\lambda_1) < L(\gamma)$; this is possible by a special case of an approximation theorem due to Whitney (see [15], page 27). Since the derivative λ_1' is never zero, the function $s(t)$ defined by

$$s(t) = \int_a^t |\lambda_1'(t)| dt$$

is invertible and λ_1 can be reparametrized by arc length. It follows that we can find an element of $\lambda_2 : [a,b] \to M$ which is a reparametrization of λ_1 of constant speed. But then

$$2(b-a)J(\lambda_2) = [L(\lambda_2)]^2 = [L(\lambda_1)]^2 < [L(\gamma)]^2 \leq 2(b-a)J(\gamma),$$

a contradiction since γ was supposed to minimize J. Hence γ must in fact minimize L. By a similar argument, one shows that if γ minimizes J, it must have constant speed.

The preceding propositions motivate use of the function $J : \Omega_{[a,b]}(M;p,q) \to \mathbb{R}$ instead of L. We want to develop enough of the calculus on the "infinite-dimensional manifold" $\Omega_{[a,b]}(M;p,q)$ to enable us to find the critical points of J. To start with, we need the notion of a smooth curve

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega_{[a,b]}(M;p,q) \quad \text{such that} \quad \bar{\alpha}(0) = \gamma,$$

where γ is a given element of Ω.

We would like to define smooth charts on the path space $\Omega_{[a,b]}(M;p,q)$, but for now a simpler approach will suffice. We will say that a *variation* of γ is a map

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega_{[a,b]}(M;p,q)$$

such that $\bar{\alpha}(0) = \gamma$ and if

$$\alpha : (-\epsilon, \epsilon) \times [a,b] \to M \quad \text{is defined by} \quad \alpha(s,t) = \bar{\alpha}(s)(t),$$

then α is smooth.

Definition. An element $\gamma \in \Omega_{[a,b]}(M;p,q)$ is a *critical point* for J if

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = 0, \quad \text{for every variation } \bar{\alpha} \text{ of } \gamma. \tag{1.3}$$

Definition. An element $\gamma \in \Omega_{[a,b]}(M;p,q)$ is called a *geodesic* if it is a critical point for J.

Thus the geodesics are the candidates for curves of shortest length from p to q, that is candidates for the notion of straight line in Riemannian geometry.

We would like to be able to determine the geodesics in Riemannian manifolds. It is easiest to do this for the case of a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ that has been provided with an isometric imbedding in $\mathbb{E}^N$. Thus we imagine that $M \subseteq \mathbb{E}^N$ and thus each tangent space T_pM can be regarded as a linear subspace of $\mathbb{R}^N$. Moreover,

$$\langle v, w\rangle_p = v \cdot w, \quad \text{for} \quad v, w \in T_pM \subseteq \mathbb{E}^N,$$

where the dot on the right is the dot product in $\mathbb{E}^N$. If

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega_{[a,b]}(M;p,q)$$

is a variation of an element $\gamma \in \Omega_{[a,b]}(M;p,q)$, with corresponding map

$$\alpha : (-\epsilon, \epsilon) \times [a,b] \to M \subseteq \mathbb{E}^N,$$

then

$$\begin{aligned} \frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} &= \frac{d}{ds}\left[\frac{1}{2}\int_a^b \frac{\partial\alpha}{\partial t}(s,t)\cdot\frac{\partial\alpha}{\partial t}(s,t)dt\right]\Bigg|_{s=0} \\ &= \int_a^b \frac{\partial^2\alpha}{\partial s\partial t}(s,t)\cdot\frac{\partial\alpha}{\partial t}(s,t)dt\Bigg|_{s=0} = \int_a^b \frac{\partial^2\alpha}{\partial s\partial t}(0,t)\cdot\frac{\partial\alpha}{\partial t}(0,t)dt, \end{aligned}$$

where α is regarded as an $\mathbb{E}^N$-valued function. If we integrate by parts, and use the fact that

$$\frac{\partial\alpha}{\partial s}(0,b) = 0 = \frac{\partial\alpha}{\partial s}(0,a),$$

we find that

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} = -\int_a^b \frac{\partial\alpha}{\partial s}(0,t)\cdot\frac{\partial^2\alpha}{\partial t^2}(0,t)dt = -\int_a^b V(t)\cdot\gamma''(t)dt, \tag{1.4}$$

where $V(t) = (\partial\alpha/\partial s)(0,t)$ is called the *variation field* of the variation field $\bar{\alpha}$.

Note that $V(t)$ can be an arbitrary smooth $\mathbb{E}^N$-valued function such that

$$V(a) = 0, \quad V(b) = 0, \quad V(t) \in T_{\gamma(t)}M, \quad \text{for all } t \in [a,b],$$

that is, V can be an arbitrary element of the "tangent space"

$$\begin{aligned} T_\gamma\Omega = \{ &\text{ smooth maps } V : [a,b] \to \mathbb{E}^N \\ &\text{such that } V(a) = 0 = V(b), V(t) \in T_{\gamma(t)}M \text{ for } t \in [a,b] \ \}. \end{aligned}$$

We can define a linear map $dJ(\gamma) : T_\gamma\Omega \to \mathbb{R}$ by

$$dJ(\gamma)(V) = -\int_a^b \langle V(t), \gamma''(t)\rangle \, dt = \frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0},$$

for any variation $\bar{\alpha}$ with variation field V. We think of $dJ(\gamma)$ as the *differential* of J at γ.

If $dJ(\gamma)(V) = 0$ for all $V \in T_\gamma\Omega$, then $\gamma''(t)$ must be perpendicular to $T_{\gamma(t)}M$ for all $t \in [a, b]$. In other words, $\gamma : [a, b] \to M$ is a geodesic if and only if

$$\left(\gamma''(t)\right)^\top = 0, \quad \text{for all } t \in [a, b], \tag{1.5}$$

where $\left(\gamma''(t)\right)^\top$ denotes the orthogonal projection of $\gamma''(t)$ into $T_{\gamma(t)}M$. To see this rigorously, we choose a smooth function $\eta : [a, b] \to \mathbb{R}$ such that

$$\eta(a) = 0 = \eta(b), \quad \eta > 0 \quad \text{on} \quad (a, b),$$

and set

$$V(t) = \eta(t)\left(\gamma''(t)\right)^\top,$$

which is clearly an element of $T_\gamma\Omega$. Then $dJ(\gamma)(V) = 0$ implies that

$$\int_a^b \eta(t)\left(\gamma''(t)\right)^\top \cdot \gamma''(t) dt = \int_a^b \eta(t)\left\|\left(\gamma''(t)\right)^\top\right\|^2 dt = 0.$$

Since the integrand is nonnegative it must vanish identically, and (1.5) must indeed hold.

We have thus obtained a simple equation (1.5) which characterizes geodesics in a submanifold M of $\mathbb{E}^N$. The geodesic equation is a generalization of the simplest second-order linear ordinary differential equation, the equation of a particle moving with zero acceleration in Euclidean space, which asks for a vector-valued function

$$\gamma : (a, b) \longrightarrow \mathbb{E}^N \qquad \text{such that} \qquad \gamma''(t) = 0.$$

Its solutions are the constant speed straight lines. The simplest way to make this differential equation nonlinear is to consider an imbedded submanifold M of $\mathbb{E}^N$ with the induced Riemannian metric, and ask for a function

$$\gamma : (a, b) \longrightarrow M \subset \mathbb{E}^N \qquad \text{such that} \qquad (\gamma''(t))^\top = 0.$$

In simple terms, we are asking for the curves which are as straight as possible subject to the constraint that they lie within M.

Example. Suppose that

$$M = \mathbb{S}^n = \{(x^1, \ldots, x^{n+1}) \in \mathbb{E}^{n+1} : (x^1)^2 + \cdots + (x^{n+1})^2 = 1\}.$$

Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be two unit-length vectors in $\mathbb{S}^n$ which are perpendicular to each other and define the unit-speed great circle $\gamma : [a, b] \to \mathbb{S}^n$ by

$$\gamma(t) = \cos t\mathbf{e}_1 + \sin t\mathbf{e}_2.$$

Then a direct calculation shows that

$$\gamma''(t) = -\cos t\mathbf{e}_1 - \sin t\mathbf{e}_2 = -\gamma(t).$$

Hence $(\gamma''(t))^\top = 0$ and γ is a geodesic on $\mathbb{S}^n$. We will see later that all geodesics on $\mathbb{S}^n$ are obtained in this manner.

1.3.2 Piecewise smooth paths

Instead of smooth paths, we could have followed Milnor [25], §11, and considered the space of piecewise smooth maps,

$$\hat{\Omega}_{[a,b]}(M;p,q) = \{ \text{ piecewise smooth paths } \gamma : [a,b] \to M : \gamma(a) = p, \gamma(b) = q\}.$$

By piecewise smooth, we mean γ is continuous and there exist $t_0 < t_1 < \cdots < t_N$ with $t_0 = a$ and $t_N = b$ such that $\gamma|[t_{i-1}, t_i]$ is smooth for $1 \leq i \leq N$. In this a *variation* of γ is a map

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \hat{\Omega}_{[a,b]}(M;p,q)$$

such that $\bar{\alpha}(0) = \gamma$ and if

$$\alpha : (-\epsilon, \epsilon) \times [a,b] \to M \quad \text{is defined by} \quad \alpha(s,t) = \bar{\alpha}(s)(t),$$

then there exist $t_0 < t_1 < \cdots < t_N$ with $t_0 = a$ and $t_N = b$ such that

$$\alpha|(-\epsilon, \epsilon) \times [t_{i-1}, t_i]$$

is smooth for $1 \leq i \leq N$. As before, we find that

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = \int_a^b \frac{\partial^2 \alpha}{\partial s \partial t}(0,t) \cdot \frac{\partial \alpha}{\partial t}(0,t)dt,$$

but now the integration by parts is more complicated because $\gamma'(t)$ is not continuous at $t_1, \ldots, t_{N-1}$. If we let

$$\gamma'(t_i-) = \lim_{t \to t_i-} \gamma'(t), \quad \gamma'(t_i+) = \lim_{t \to t_i+} \gamma'(t),$$

a short calculation shows that (1.4) becomes

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = -\int_a^b V(t) \cdot \gamma''(t)dt - \sum_{i=1}^{N-1} V(t_i) \cdot (\gamma'(t_i+) - \gamma'(t_i-)),$$

whenever $\bar{\alpha}$ is any variation of γ with variation field V. If

$$dJ(\gamma)(V) = \frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = 0$$

for all variation fields V in the tangent space

$$T_\gamma \hat{\Omega} = \{ \text{ piecewise smooth maps } V : [a,b] \to \mathbb{E}^N \\ \text{such that } V(a) = 0 = V(b), V(t) \in T_{\gamma(t)} M \text{ for } t \in [a,b] \ \},$$

it follows that $\gamma'(t_i+) = \gamma'(t_i-)$ for every i and $(\gamma''(t))^\top = 0$. Thus critical points on the more general space of piecewise smooth paths are also smooth geodesics.

Exercise I. Suppose that M^2 is the right circular cylinder defined by the equation $x^2 + y^2 = 1$ in $\mathbb{E}^3$. Show that for each choice of real numbers a and b the curve

$$\gamma : \mathbb{R} \to M^2 \subseteq \mathbb{E}^3 \quad \text{defined by} \quad \gamma_{a,b}(t) = \begin{pmatrix} \cos(at) \\ \sin(at) \\ bt \end{pmatrix}$$

is a geodesic.

1.4 Hamilton's principle

Of course, we would like a formula for geodesics that does not depend upon the existence of an isometric imbedding. To derive such a formula, it is convenient to regard the action J in a more general context, namely that of classical mechanics.

Definition. A *simple mechanical system* is a triple $(M, \langle \cdot, \cdot \rangle, \phi)$, where $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold and $\phi : M \to \mathbb{R}$ is a smooth function.

We call M the *configuration space* of the simple mechanical system. If $\gamma : [a,b] \to M$ represents the motion of the system,

$$\frac{1}{2}\langle \gamma'(t), \gamma'(t) \rangle = (\text{kinetic energy at time } t), \\ \phi(\gamma(t)) = (\text{potential energy at time } t).$$

Example 1. If a planet of mass m is moving around a star of mass M with $M >> m$, the star assumed to be stationary, we might take

$$M = \mathbb{R}^3 - \{(0,0,0)\}, \\ \langle \cdot, \cdot \rangle = m(dx \otimes dx + dy \otimes dy + dz \otimes dz), \\ \phi(x,y,z) = \frac{-GMm}{\sqrt{x^2 + y^2 + z^2}}.$$

Here G is the gravitational constant. Sir Isaac Newton derived Kepler's three laws from this simple mechanical system.

Example 2. To construct an interesting example in which the configuration space M is not Euclidean space, we take $M = SO(3)$, the group of real orthogonal 3×3 matrices of determinant one, regarded as the space of "configurations" of a rigid body B in $\mathbb{R}^3$ which has its center of mass located at the origin. We want to describe the motion of the rigid body as a path $\gamma : [a, b] \to M$. If p is a point in the rigid body with coordinates (x^1, x^2, x^3) at time $t = 0$, we suppose that the coordinates of this point at time t will be

$$\gamma(t) \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}, \qquad \text{where} \qquad \gamma(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & a_{13}(t) \\ a_{21}(t) & a_{22}(t) & a_{23}(t) \\ a_{31}(t) & a_{32}(t) & a_{33}(t) \end{pmatrix} \in SO(3),$$

and $\gamma(0) = I$, the identity matrix. Then the velocity $v(t)$ of the particle p at time t will be

$$v(t) = \gamma'(t) \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^3 a'_{1i}(t)x^i \\ \sum_{i=1}^3 a'_{2i}(t)x^i \\ \sum_{i=1}^3 a'_{3i}(t)x^i \end{pmatrix},$$

and hence

$$v(t) \cdot v(t) = \sum_{i,j,k=1}^{3} a'_{ki}(t)a'_{kj}(t)x^i x^j.$$

Suppose now that $\rho(x^1, x^2, x^3)$ is the density of matter at (x^1, x^2, x^3) within the rigid body. Then the total kinetic energy within the rigid body at time t will be given by the expression

$$K = \frac{1}{2} \sum_{i,j,k} \left(\int_B \rho(x^1, x^2, x^3) x^i x^j dx^1 dx^2 dx^3 \right) a'_{ki}(t)a'_{kj}(t).$$

We can rewrite this as

$$K = \frac{1}{2} \sum_{i,j,k} c_{ij} a'_{ki}(t)a'_{kj}(t), \qquad \text{where} \qquad c_{ij} = \int_B \rho(x^1, x^2, x^3) x^i x^j dx^1 dx^2 dx^3,$$

and define a Riemannian metric $\langle \cdot, \cdot \rangle$ on $M = SO(3)$ by

$$\langle \gamma'(t), \gamma'(t) \rangle = \sum_{i,j,k} c_{ij} a'_{ki}(t)a'_{kj}(t).$$

Then once again $(1/2)\langle \gamma'(t), \gamma'(t) \rangle$ represents the kinetic energy, this time of the rigid body B when its motion is represented by the curve $\gamma : (a, b) \to M$. We remark that the constants

$$I_{ij} = \text{Trace}(c_{ij})\delta_{ij} - c_{ij}$$

are called the *moments of inertia* of the rigid body.

A smooth function $\phi : SO(3) \to \mathbb{R}$ can represent the potential energy for the rigid body. In classical mechanics books, the motion of a top is described by means of a simple mechanical system which has configuration space $SO(3)$ with a suitable left-invariant metric and potential ϕ. Applied to the rotating earth, the same equations explain the precession of the equinoxes, according to which the axis of the earth traverses a circle in the celestial sphere once every 26,000 years.

In Lagrangian mechanics, the equations of motion for a simple mechanical system are derived from a variational principle. The key step is to define the *Lagrangian* to be the kinetic energy minus the potential energy. More precisely, for a simple mechanical system $(M, \langle \cdot, \cdot \rangle, \phi)$, we define the Lagrangian $\mathcal{L} : TM \to \mathbb{R}$ by

$$\mathcal{L}(v) = \frac{1}{2}\langle v, v\rangle - \phi(\pi(v)),$$

where $\pi : TM \to M$ is the usual projection. We can then define the *action* $J : \Omega_{[a,b]}(M; p, q) \to \mathbb{R}$ by

$$J(\gamma) = \int_a^b \mathcal{L}(\gamma'(t))dt.$$

As before, we say that $\gamma \in \Omega$ is a critical point for J if (1.3) holds. We can than formulate Lagrangian classical mechanics as follows:

Hamilton's principle. *If γ represents the motion of a simple mechanical system, then γ is a critical point for J.*

Thus the problem of finding curves from p to q of shortest length is put into a somewhat broader context.

It can be shown that if $\gamma \in \Omega_{[a,b]}(M; p, q)$ is a critical point for J and $[c, d] \subseteq [a, b]$, then the restriction of γ to $[c, d]$ is also a critical point for J, this time on the space $\Omega_{[c,d]}(M; r, s)$, where $r = \gamma(c)$ and $s = \gamma(d)$. Thus we can assume that $\gamma([a, b]) \subseteq U$, where $(U, x^1, \ldots, x^n)$ is a coordinate system on M, and we can express $\mathcal{L}$ in terms of the coordinates $(x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n)$ on $\pi^{-1}(U)$ described by (1.1). If

$$\gamma(t) = (x^1(t), \ldots, x^n(t)), \quad \text{and} \quad \gamma'(t) = (x^1(t), \ldots, x^n(t), \dot{x}^1(t), \ldots, \dot{x}^n(t)),$$

then

$$\mathcal{L}(\gamma'(t)) = \mathcal{L}(x^1(t), \ldots, x^n(t), \dot{x}^1(t), \ldots, \dot{x}^n(t)).$$

Theorem 1. *A point $\gamma \in \Omega_{[a,b]}(M; p, q)$ is a critical point for the action J $\Leftrightarrow$ its coordinate functions satisfy the Euler-Lagrange equations*

$$\frac{\partial \mathcal{L}}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) = 0. \tag{1.6}$$

Proof: We prove only the implication $\Rightarrow$, and leave the other half (which is quite a bit easier) as an exercise. We make the assumption that $\gamma([a, b]) \subseteq U$, where U is the domain of local coordinates as described above.

For $1 \leq i \leq n$, let $\psi^i : [a,b] \to \mathbb{R}$ be a smooth function such that $\psi^i(a) = 0 = \psi^i(b)$, and define a variation

$$\alpha : (-\epsilon, \epsilon) \times [a,b] \to U \quad \text{by} \quad \alpha(s,t) = (x^1(t) + s\psi^1(t), \dots, x^n(t) + s\psi^n(t)).$$

Let $\dot{\psi}^i(t) = (d/dt)(\psi^i)(t)$. Then

$$J(\bar{\alpha}(s)) = \int_a^b \mathcal{L}(\cdots, x^i(t) + s\psi^i(t), \dots, \dot{x}^i(t) + s\dot{\psi}^i(t), \dots) dt,$$

so it follows from the chain rule that

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} = \int_a^b \sum_{i=1}^n \left[\frac{\partial \mathcal{L}}{\partial x^i}(x^i(t), \dot{x}^i(t))\psi^i(t) + \frac{\partial \mathcal{L}}{\partial \dot{x}^i}(x^i(t), \dot{x}^i(t))\dot{\psi}^i(t)\right] dt.$$

Since $\psi^i(a) = 0 = \psi^i(b)$,

$$0 = \int_a^b \sum_{i=1}^n \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\psi^i\right) dt = \int_a^b \sum_{i=1}^n \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right)\psi^i dt + \int_a^b \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial \dot{x}^i}\dot{\psi}^i dt,$$

and hence

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} = \int_a^b \sum_{i=1}^n \left[\frac{\partial \mathcal{L}}{\partial x^i}\psi^i - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right)\psi^i\right] dt.$$

Thus if γ is a critical point for J, we must have

$$0 = \int_a^b \sum_{i=1}^n \left[\frac{\partial \mathcal{L}}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right)\right]\psi^i dt,$$

for every choice of smooth functions $\psi(t)$. In particular, if $\eta : [a,b] \to \mathbb{R}$ is a smooth function such that

$$\eta(a) = 0 = \eta(b), \quad \eta > 0 \quad \text{on} \quad (a,b),$$

and we set

$$\psi^i(t) = \eta(t)\left[\frac{\partial \mathcal{L}}{\partial x^i}(x^i(t), \dot{x}^i(t)) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}(x^i(t), \dot{x}^i(t))\right)\right],$$

then

$$\int_a^b \eta(t)\left[\frac{\partial \mathcal{L}}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right)\right]^2 dt = 0.$$

Since the integrand is nonnegative, it must vanish identically and hence (1.6) must hold.

For a simple mechanical system, the Euler-Lagrange equations yield a derivation of Newton's second law of motion. Indeed, if

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^n g_{ij} dx^i dx^j,$$

then in the standard coordinates $(x^1, \dots, x^n, \dot{x}^1, \dots, \dot{x}^n)$,

$$\mathcal{L}(\gamma'(t)) = \frac{1}{2} \sum_{i,j=1}^{n} g_{ij}(x^1, \dots, x^n) \dot{x}^i \dot{x}^j - \phi(x^1, \dots x^n).$$

Hence

$$\frac{\partial \mathcal{L}}{\partial x^i} = \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial g_{jk}}{\partial x^i} \dot{x}^j \dot{x}^k - \frac{\partial \phi}{\partial x^i},$$

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^i} = \sum_{j=1}^{n} g_{ij} \dot{x}^j, \qquad \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) = \sum_{j,k=1}^{n} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^j \dot{x}^k + \sum_{j=1}^{n} g_{ij} \ddot{x}^j,$$

where $\ddot{x}^j = d^2 x^j / dt^2$. Thus the Euler-Lagrange equations become

$$\sum_{j=1}^{n} g_{ij} \ddot{x}^j + \sum_{j,k=1}^{n} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^j \dot{x}^k = \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial g_{jk}}{\partial x^i} \dot{x}^j \dot{x}^k - \frac{\partial \phi}{\partial x^i}$$

or

$$\sum_{j=1}^{n} g_{ij} \ddot{x}^j + \frac{1}{2} \sum_{j,k=1}^{n} \left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i} \right) \dot{x}^j \dot{x}^k = -\frac{\partial \phi}{\partial x^i}.$$

We multiply through by the matrix (g^{ij}) which is inverse to (g_{ij}) to obtain

$$\ddot{x}^l + \sum_{j,k=1}^{n} \Gamma^l_{jk} \dot{x}^j \dot{x}^k = -\sum_{i=1}^{n} g^{li} \frac{\partial \phi}{\partial x^i}, \tag{1.7}$$

where

$$\Gamma^l_{ij} = \sum_{i=1}^{n} g^{li} \left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i} \right). \tag{1.8}$$

The expressions Γ^l_{ij} are called the *Christoffel symbols*. Note that if $(x^1, \dots, x^n)$ are rectangular cartesian coordinates in Euclidean space, the Christoffel symbols vanish.

We can interpret the two sides of (1.7) as follows:

$$\ddot{x}^l + \sum_{j,k=1}^{n} \Gamma^l_{jk} \dot{x}^j \dot{x}^k = (\text{acceleration})^l,$$

$$-\sum_{i=1}^{n} g^{li} \frac{\partial \phi}{\partial x^i} = (\text{force per unit mass})^l.$$

Hence equation (1.7) is just the statement of Newton's second law, force equals mass times acceleration, for simple mechanical systems.

In the case where $\phi = 0$, we obtain the differential equations for geodesics,

$$\ddot{x}^i + \sum_{j,k=1}^{n} \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0, \tag{1.9}$$

where the Γ^i_{jk}'s are the Christoffel symbols.

Example. In the case of Euclidean space $\mathbb{E}^n$ with the standard Euclidean metric, $g_{ij} = \delta_{ij}$, the Christoffel symbols vanish and the equations for geodesics become

$$\frac{d^2x^i}{dt^2} = 0.$$

The solutions are

$$x^i = a^i t + b^i,$$

the straight lines parametrized with constant speed.

Note that the Euler-Lagrange equations can be written as follows:

$$\begin{cases} \frac{dx^l}{dt} = \dot{x}^l, \\ \frac{d\dot{x}^l}{dt} = -\sum_{j,k=1}^n \Gamma^l_{jk}\dot{x}^j\dot{x}^k - \sum_{i=1}^n g^{li}\frac{\partial\phi}{\partial x^i}. \end{cases} \tag{1.10}$$

This is a first-order system in canonical form, and hence it follows from the fundamental existence and uniqueness theorem from the theory of ordinary differential equations ([5], Chapter IV, §4) that given an element $v \in T_pM$, there is a unique solution to this system, defined for $t \in (-\epsilon, \epsilon)$ for some $\epsilon > 0$, which satisfies the initial conditions

$$x^i(0) = x^i(p), \qquad \dot{x}^i(0) = \dot{x}^i(v).$$

In the special case where $\phi = 0$, we can restate this as:

Theorem 2. *Given $p \in M$ and $v \in T_pM$, there is a unique geodesic $\gamma : (-\epsilon, \epsilon) \to M$ for some $\epsilon > 0$ such that $\gamma(0) = p$ and $\gamma'(0) = v$.*

Exercise II. Consider the upper half-plane $\mathbb{H}^2 = \{(x, y) \in \mathbb{R}^2 : y > 0\}$, with Riemannian metric

$$\langle \cdot, \cdot \rangle = \frac{1}{y^2}(dx \otimes dx + dy \otimes dy),$$

the so-called Poincaré upper half plane.

a. Calculate the Christoffel symbols Γ^k_{ij}.

b. Write down the equations for the geodesics, obtaining two equations

$$\frac{d^2x}{dt^2} = \cdots, \quad \frac{d^2y}{dt^2} = \cdots.$$

c. Assume that $y = y(x)$ and eliminate t from these two equations by using the relation

$$\frac{d^2y}{dt^2} = \frac{d}{dt}\left(\frac{dy}{dx}\frac{dx}{dt}\right) = \frac{d^2y}{dx^2}\left(\frac{dx}{dt}\right)^2 + \frac{dy}{dx}\frac{d^2x}{dt^2}.$$

Solve the resulting differential equation to determine the paths traced by the geodesics in the Poincaré upper half plane.

1.5 The Levi-Civita connection

In modern differential geometry, the Christoffel symbols Γ^k_{ij} are regarded as the components of a connection. We now describe how that goes.

You may recall from Math 240A that a *smooth vector field* on the manifold M is a smooth map

$$X : M \to TM \quad \text{such that} \quad \pi \circ X = \mathrm{id}_M,$$

where $\pi : TM \to M$ is the usual projection, or equivalently a smooth map

$$X : M \to TM \quad \text{such that} \quad X(p) \in T_pM.$$

The restriction of a vector field to the domain U of a smooth coordinate system $(x^1, \ldots, x^n)$ can be written as

$$X|U = \sum_{i=1}^{n} f^i \frac{\partial}{\partial x^i}, \quad \text{where} \quad f^i : U \to \mathbb{R}.$$

If we evaluate at a given point $p \in U$ this specializes to

$$X(p) = \sum_{i=1}^{n} f^i(p) \left. \frac{\partial}{\partial x^i} \right|_p .$$

A vector field X can be regarded as a first-order differential operator. Thus, if $g : M \to \mathbb{R}$ is a smooth function, we can operate on g by X, thereby obtaining a new smooth function $Xg : M \to \mathbb{R}$ by $(Xg)(p) = X(p)(g)$.

We let $\mathcal{X}(M)$ denote the space of all smooth vector fields on M. It can be regarded as a real vector space or as an $\mathcal{F}(M)$-module, where $\mathcal{F}(M)$ is the space of all smooth real-valued functions on M, where the multiplication $\mathcal{F}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ is defined by $(fX)(p) = f(p)X(p)$.

Definition. A *connection* on the tangent bundle TM is an operator

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M)$$

that satisfies the following axioms (where we write $\nabla_X Y$ for $\nabla(X, Y)$:

$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z, \tag{1.11}$$

$$\nabla_Z(fX + gY) = (Zf)X + f\nabla_Z X + (Zg)Y + g\nabla_Z Y, \tag{1.12}$$

for $f, g \in \mathcal{F}(M)$ and $X, Y, Z \in \mathcal{X}(M)$.

Note that (4.19) is the usual "Leibniz rule" for differentiation. We often call $\nabla_X Y$ the *covariant derivative* of Y in the direction of X.

Lemma 1. *Any connection ∇ is local; that is, if U is an open subset of M,*

$$X|U \equiv 0 \quad \Rightarrow \quad (\nabla_X Y)|U \equiv 0 \quad \text{and} \quad (\nabla_Y X)|U \equiv 0,$$

for any $Y \in \mathcal{X}(M)$.

Proof: Let p be a point of U and choose a smooth function $f : M \to \mathbb{R}$ such that $f \equiv 0$ on a neighborhood of p and $f \equiv 1$ outside U. Then

$$X|U \equiv 0 \Rightarrow fX \equiv X.$$

Hence

$$\begin{aligned}(\nabla_X Y)(p) = \nabla_{fX} Y(p) = f(p)(\nabla_X Y)(p) = 0,\\ (\nabla_Y X)(p) = \nabla_Y (fX)(p) = f(p)(\nabla_Y X)(p) + (Yf)(p)X(p) = 0.\end{aligned}$$

Since p was an arbitrary point of U, we conclude that $(\nabla_X Y)|U \equiv 0$ and $(\nabla_Y X)|U \equiv 0$.

This lemma implies that if U is an arbitrary open subset of M a connection ∇ on TM will restrict to a unique well-defined connection ∇ on TU.

Thus we can restrict to the domain U of a local coordinate system $(x^1, \ldots, x^n)$ and define the components $\Gamma^k_{ij} : U \to \mathbb{R}$ of the connection by

$$\nabla_{\partial/\partial x^i} \frac{\partial}{\partial x^j} = \sum_{k=1}^n \Gamma^k_{ij} \frac{\partial}{\partial x^k}.$$

Then if X and Y are smooth vector fields on U, say

$$X = \sum_{i=1}^n f^i \frac{\partial}{\partial x^i}, \quad Y = \sum_{j=1}^n g^j \frac{\partial}{\partial x^j},$$

we can use the connection axioms and the components of the connection to calculate $\nabla_X Y$:

$$\nabla_X Y = \sum_{i=1}^n \left[\sum_{j=1}^n f^j \frac{\partial g^i}{\partial x^j} + \sum_{j,k=1}^n \Gamma^i_{jk} f^j g^k \right] \frac{\partial}{\partial x^i}. \tag{1.13}$$

Lemma 2. $(\nabla_X Y)(p)$ *depends only on* $X(p)$ *and on the values of* Y *along some curve tangent to* $X(p)$.

Proof: This follows immediately from (1.13).

Because of the previous lemma, we can $\nabla_v X \in T_p M$, whenever $v \in T_p M$ and X is a vector field defined along some curve tangent to v at p, by setting

$$\nabla_v X = (\nabla_{\tilde{V}} \tilde{X})(p),$$

for any choice of extensions $\tilde{V}$ of v and $\tilde{X}$ of X. In particular, if $\gamma : [a, b] \to M$ is a smooth curve, we can define the vector field $\nabla_{\gamma'} \gamma'$ along γ. Recall that we define the *Lie bracket* of two vector fields X and Y by $[X, Y](f) =$

$X(Y(f)) - Y(X(f))$. If X and Y are smooth vector fields on the domain U of local corrdinates $(x^1, \dots, x^n)$, say

$$X = \sum_{i=1}^{n} f^i \frac{\partial}{\partial x^i}, \quad Y = \sum_{j=1}^{n} g^j \frac{\partial}{\partial x^j},$$

then

$$[X, Y] = \sum_{i,j=1}^{n} f^i \frac{\partial g^j}{\partial x^i} \frac{\partial}{\partial x^j} - \sum_{i,j=1}^{n} g^j \frac{\partial f^i}{\partial x^j} \frac{\partial}{\partial x^i}.$$

Fundamental Theorem of Riemannian Geometry. *If $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold, there is a unique connection on TM such that*

1. *∇ is symmetric, that is, $\nabla_X Y - \nabla_Y X = [X, Y]$, for $X, Y \in \mathcal{X}(M)$,*
2. *∇ is metric, that is, $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$, for $X, Y, Z \in \mathcal{X}(M)$.*

This connection is called the *Levi-Civita connection* of the Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$.

To prove the theorem we express the two conditions in terms of local coordinates $(x^1, \dots, x^n)$ defined on an open subset U of M. In terms of the components of ∇, defined by the formula

$$\nabla_{\partial/\partial x^i} \frac{\partial}{\partial x^j} = \sum_{k=1}^{n} \Gamma_{ij}^k \frac{\partial}{\partial x^k}, \tag{1.14}$$

the first condition becomes

$$\Gamma_{ij}^k = \Gamma_{ji}^k, \quad \text{since} \quad \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0.$$

Thus the Γ_{ij}^k's are symmetric in the lower pair of indices. If we write

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j,$$

then the second condition yields

$$\begin{aligned} \frac{\partial g_{ij}}{\partial x^k} &= \frac{\partial}{\partial x^k} \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = \left\langle \nabla_{\partial/\partial x^k} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle + \left\langle \frac{\partial}{\partial x^i}, \nabla_{\partial/\partial x^k} \frac{\partial}{\partial x^j} \right\rangle \\ &= \left\langle \sum_{l=1}^{n} \Gamma_{ki}^l \frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^j} \right\rangle + \left\langle \frac{\partial}{\partial x^i}, \sum_{l=1}^{n} \Gamma_{kj}^l \frac{\partial}{\partial x^l} \right\rangle = \sum_{l=1}^{n} g_{lj} \Gamma_{ki}^l + \sum_{l=1}^{n} g_{il} \Gamma_{kj}^l. \end{aligned}$$

In fact, the second condition is equivalent to

$$\frac{\partial g_{ij}}{\partial x^k} = \sum_{l=1}^{n} g_{lj} \Gamma_{ki}^l + \sum_{l=1}^{n} g_{il} \Gamma_{kj}^l. \tag{1.15}$$

We can permute the indices in (1.15), obtaining

$$\frac{\partial g_{jk}}{\partial x^i} = \sum_{l=1}^{n} g_{lk}\Gamma^l_{ij} + \sum_{l=1}^{n} g_{jl}\Gamma^l_{ik}. \tag{1.16}$$

and

$$\frac{\partial g_{ki}}{\partial x^j} = \sum_{l=1}^{n} g_{li}\Gamma^l_{jk} + \sum_{l=1}^{n} g_{kl}\Gamma^l_{ji}. \tag{1.17}$$

Subtracting (1.15) from the sum of (1.16) and (1.17), and using the symmetry of Γ^l_{ij} in the lower indices, yields

$$\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} = 2\sum_{l=1}^{n} g_{lk}\Gamma^l_{ij}.$$

Thus if we let (g^{ij}) denote the matrix inverse to $g_{ij})$, we find that

$$\Gamma^l_{ij} = \frac{1}{2}\sum_{i=1}^{n} g^{li}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right), \tag{1.18}$$

which is exactly the formula (1.8) we obtained before by means of Hamilton's principle.

This proves uniqueness of the connection which is both symmetric and metric. For existence, we define the connection locally by (1.14), where the Γ^l_{ij}'s are defined by (1.18) and check that the resulting connection is both symmetric and metric. (Note that by uniqueness, the locally defined connections fit together on overlaps.)

In the special case where the Riemannian manifolds is Euclidean space $\mathbb{E}^N$ the Levi-Civita connection is easy to describe. In this case, we have global rectangular cartesian coordinates $(x^1, \ldots, x^N)$ on $\mathbb{E}^N$ and any vector field Y on $\mathbb{E}^N$ can be written as

$$Y = \sum_{i=1}^{N} f^i \frac{\partial}{\partial x^i}, \quad \text{where} \quad f^i : \mathbb{E}^N \to \mathbb{R}.$$

In this case, the Levi-Civita connection ∇^E has components $\Gamma^k_{ij} = 0$, and therefore the operator ∇^E satisfies the formula

$$\nabla^E_X Y = \sum_{i=1}^{N} (Xf^i)\frac{\partial}{\partial x^i}.$$

It is easy to check that this connection which is symmetric and metric for the Euclidean metric.

If M is an imbedded submanifold of $\mathbb{E}^N$ with the induced metric, then one can define a connection $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ by

$$(\nabla_X Y)(p) = (\nabla^E_X Y(p))^\top,$$

where $(\cdot)^\top$ is the orthogonal projection into the tangent space. (Use Lemma 5.2 to justify this formula.) It is a straightforward exercise to show that ∇ is symmetric and metric for the induced connection, and hence ∇ is the Levi-Civita connection for M. Note that if $\gamma : [a, b] \to M$ is a smooth curve, then

$$(\nabla_{\gamma'}\gamma')(t) = (\gamma''(t))^\top,$$

so a smooth curve in $M \subseteq \mathbb{E}^N$ is a geodesic if and only if $\nabla_{\gamma'}\gamma' \equiv 0$. If we want to develop the subject independent of Nash's imbedding theorem, we can make the

Definition. If $(M, \langle\cdot,\cdot\rangle)$ is a Riemannian manifold, a smooth path $\gamma : [a, b] \to M$ is a *geodesic* if it satisfies the equation $\nabla_{\gamma'}\gamma' \equiv 0$, where ∇ is the Levi-Civita connection.

In terms of local coordinates, if

$$\gamma' = \sum_{i=1}^n \frac{d(x^i \circ \gamma)}{dt} \frac{\partial}{\partial x^i},$$

then a straightforward calculation yields

$$\nabla_{\gamma'}\gamma' = \sum_{i=1}^n \left[\frac{d^2(x^i \circ \gamma)}{dt^2} + \sum_{j,k=1}^n \Gamma^i_{jk} \frac{d(x^j \circ \gamma)}{dt} \frac{d(x^k \circ \gamma)}{dt} \right] \frac{\partial}{\partial x^i}. \tag{1.19}$$

This reduces to the equation (1.9) we obtained before from Hamilton's principle. Note that

$$\frac{d}{dt}\langle\gamma', \gamma'\rangle = \gamma'\langle\gamma', \gamma'\rangle = 2\langle\nabla_{\gamma'}\gamma', \gamma'\rangle = 0,$$

so geodesics automatically have constant speed.

More generally, if $\gamma : [a, b] \to M$ is a smooth curve, we call $\nabla_{\gamma'}\gamma'$ the *acceleration* of γ. Thus if $(M, \langle\cdot,\cdot\rangle, \phi)$ is a simple mechanical system, its equations of motion can be written as

$$\nabla_{\gamma'}\gamma' = -\mathrm{grad}(\phi), \tag{1.20}$$

where in local coordinates, $\mathrm{grad}(\phi) = \sum g^{ji}(\partial V/\partial x^i)(\partial/\partial x^j)$.

1.6 First variation of J

Now that we have the notion of connection available, it might be helpful to review the argument that the function

$$J : \Omega_{[a,b]}(M; p, q), \to \mathbb{R} \quad \text{defined by} \quad J(\gamma) = \frac{1}{2}\int_a^b \langle\gamma'(t), \gamma'(t)\rangle dt,$$

has geodesics as its critical points, and recast the argument in a form that is independent of choice of isometric imbedding.

In fact, the argument we gave before goes through with only one minor change, namely given a variation

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega \quad \text{with corresponding} \quad \alpha : (-\epsilon, \epsilon) \times [a, b] \to M,$$

we must make sense of the partial derivatives

$$\frac{\partial \alpha}{\partial s}, \quad \frac{\partial \alpha}{\partial t}, \quad \ldots,$$

since we can no longer regard α as a vector valued function.

But these is a simple remedy. We look at the first partial derivatives as maps

$$\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t} : (-\epsilon, \epsilon) \times [a, b] \to TM$$
$$\text{such that} \quad \pi \circ \left(\frac{\partial \alpha}{\partial s} \right) = \alpha, \quad \pi \circ \left(\frac{\partial \alpha}{\partial t} \right) = \alpha.$$

In terms of local coordinates, these maps are defined by

$$\left(\frac{\partial \alpha}{\partial s} \right)(s, t) = \sum_{i=1}^{n} \frac{\partial (x^i \circ \alpha)}{\partial s}(s, t) \left. \frac{\partial}{\partial x^i} \right|_{\alpha(s,t)},$$
$$\left(\frac{\partial \alpha}{\partial t} \right)(s, t) = \sum_{i=1}^{n} \frac{\partial (x^i \circ \alpha)}{\partial t}(s, t) \left. \frac{\partial}{\partial x^i} \right|_{\alpha(s,t)}.$$

We define higher order derivatives via the Levi-Civita connection. Thus for example, in terms of local coordinates, we set

$$\nabla_{\partial/\partial s} \left(\frac{\partial \alpha}{\partial t} \right) = \sum_{k=1}^{n} \left[\frac{\partial^2 (x^k \circ \alpha)}{\partial s \partial t} + \sum_{i,j=1}^{n} (\Gamma^k_{ij} \circ \alpha) \frac{\partial (x^i \circ \alpha)}{\partial s} \frac{\partial (x^i \circ \alpha)}{\partial t} \right] \left. \frac{\partial}{\partial x^k} \right|_{\alpha},$$

thereby obtaining a map

$$\nabla_{\partial/\partial s} \left(\frac{\partial \alpha}{\partial t} \right) : (-\epsilon, \epsilon) \times [a, b] \to TM \quad \text{such that} \quad \pi \circ \nabla_{\partial/\partial s} \left(\frac{\partial \alpha}{\partial t} \right) = \alpha.$$

Similarly, we define

$$\nabla_{\partial/\partial t} \left(\frac{\partial \alpha}{\partial t} \right), \qquad \nabla_{\partial/\partial t} \left(\frac{\partial \alpha}{\partial s} \right),$$

and so forth. In short, we replace higher order derivatives by covariant derivatives using the Levi-Civita connection for the Riemmannian metric.

The properties of the Levi-Civita connection imply that

$$\nabla_{\partial/\partial s} \left(\frac{\partial \alpha}{\partial t} \right) = \nabla_{\partial/\partial t} \left(\frac{\partial \alpha}{\partial s} \right)$$

and

$$\frac{\partial}{\partial t}\left\langle \frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\right\rangle = \left\langle \nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial s}\right), \frac{\partial \alpha}{\partial t}\right\rangle + \left\langle \frac{\partial \alpha}{\partial s}, \nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial t}\right)\right\rangle.$$

With these preparations out of the way, we can now proceed as before and let

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega_{[a,b]}(M; p, q)$$

be a smooth path with $\bar{\alpha}(0) = \gamma$ and

$$\frac{\partial \alpha}{\partial s}(0, t) = V(t),$$

where V is an element of the tangent space

$$\begin{aligned} T_\gamma\Omega = \{ & \text{ smooth maps } V : [a, b] \to TM \\ & \text{such that } \pi \circ V(t) = \gamma(t) \text{ for } t \in [a, b], \text{ and } V(a) = 0 = V(b) \ \}. \end{aligned}$$

Then just as before,

$$\begin{aligned} \frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} &= \frac{d}{ds}\left[\frac{1}{2}\int_a^b \left\langle \frac{\partial \alpha}{\partial t}(s,t), \frac{\partial \alpha}{\partial t}(s,t) dt\right\rangle\right]\Bigg|_{s=0} \\ &= \int_a^b \left\langle \nabla_{\partial/\partial s}\left(\frac{\partial \alpha}{\partial t}\right)(0,t), \frac{\partial \alpha}{\partial t}(0,t)\right\rangle dt \\ &= \int_a^b \left\langle \nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial s}\right)(0,t), \frac{\partial \alpha}{\partial t}(0,t)\right\rangle dt \\ &= \int_a^b \left[\frac{\partial}{\partial t}\left\langle \frac{\partial \alpha}{\partial s}(0,t), \frac{\partial \alpha}{\partial t}(0,t)\right\rangle - \left\langle \frac{\partial \alpha}{\partial s}(0,t), \nabla_{\partial/\partial t}\frac{\partial \alpha}{\partial t}(0,t)\right\rangle\right] dt. \end{aligned}$$

Since

$$\frac{\partial \alpha}{\partial s}(0, b) = 0 = \frac{\partial \alpha}{\partial s}(0, a),$$

we obtain

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} = -\int_a^b \langle V(t), (\nabla_{\gamma'}\gamma')(t)\rangle dt.$$

We call this the *first variation* of J in the direction of V, and write

$$dJ(\gamma)(V) = -\int_a^b \langle V(t), (\nabla_{\gamma'}\gamma')(t)\rangle dt. \tag{1.21}$$

A critical point for J is a point $\gamma \in \Omega_{[a,b]}(M; p, q)$ at which $dJ(\gamma) = 0$, and the above argument shows that the critical points for J are exactly the geodesics for the Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$.

Of course, we could modify the above derivation to determine the first variation of the action

$$J(\gamma) = \frac{1}{2}\int_a^b \langle \gamma'(t), \gamma'(t)\rangle dt - \int_a^b \phi(\gamma(t)) dt$$

for a simple mechanical system $(M, \langle \cdot, \cdot \rangle, \phi)$. We would find after a short calculation that

$$dJ(\gamma)(V) = -\int_a^b \langle V(t), (\nabla_{\gamma'}\gamma')(t)\rangle dt - \int_a^b d\phi(V)(\gamma(t)dt$$
$$= -\int_a^b \langle V(t), (\nabla_{\gamma'}\gamma')(t) - \mathrm{grad}(\phi)(\gamma(t))\rangle \, dt.$$

Once again, the critical points would be solutions to Newton's equation (1.20).

1.7 Lorentz manifolds

The notion of Riemmannian manifold has a generalization which is extremely useful in Einstein's theory of general relativity, as described for example is the standard texts [27] or [35].

Definition. Let M be a smooth manifold. A *pseudo-Riemannian metric* on M is a function which assigns to each $p \in M$ a nondegenerate symmetric bilinear map

$$\langle \cdot, \cdot \rangle_p : T_pM \times T_pM \longrightarrow \mathbb{R}$$

which which varies smoothly with $p \in M$. As before , varying smoothly with $p \in M$ means that if $\phi = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a smooth coordinate system on M, then for $p \in U$,

$$\langle \cdot, \cdot \rangle_p = \sum_{i,j=1}^n g_{ij}(p) dx^i|_p \otimes dx^j|_p,$$

where the functions $g_{ij} : U \to \mathbb{R}$ are smooth. The conditions that $\langle \cdot, \cdot \rangle_p$ be symmetric and nondegenerate are expressed in terms of the matrix (g_{ij}) by saying that (g_{ij}) is a symmetric matrix and has nonzero determinant.

It follows from linear algebra that for any choice of $p \in M$, local coordinates $(x^1, \ldots, x^n)$ can be chosen so that

$$(g_{ij}(p)) = \begin{pmatrix} -I_{p\times p} & 0 \\ 0 & I_{q\times q} \end{pmatrix},$$

where $I_{p\times p}$ and $I_{q\times ql}$ are $p \times p$ and $q \times q$ identity matrices with $p + q = n$. The pair (p, q) is called the *signature* of the pseudo-Riemannian metric.

Note that a pseudo-Riemannian metric of signature $(0, n)$ is just a Riemannian metric. A pseudo-Riemannian metric of signature $(1, n-1)$ is called a *Lorentz metric*.

A *pseudo-Riemannian manifold* is a pair $(M, \langle \cdot, \cdot \rangle)$ where M is a smooth manifold and $\langle \cdot, \cdot \rangle$ is a pseudo-Riemannian metric on M. Similarly, a *Lorentz manifold* is a pair $(M, \langle \cdot, \cdot \rangle)$ where M is a smooth manifold and $\langle \cdot, \cdot \rangle$ is a Lorentz metric on M.

Example. Let $\mathbb{R}^{n+1}$ be given coordinates $(t, x^1, \dots, x^n)$, with t being regarded as time and $(x^1, \dots, x^n)$ being regarded as Euclidean coordinates in space, and consider the Lorentz metric

$$\langle \cdot, \cdot \rangle = -c^2 dt \otimes dt + \sum_{i=1}^{n} dx^i \otimes dx^i,$$

where the constant c is regarded as the speed of light. When endowed with this metric, $\mathbb{R}^{n+1}$ is called *Minkowski space-time* and is denoted by $\mathbb{L}^{n+1}$. Four-dimensional Minkowski space-time is the arena for *special relativity*.

The arena for *general relativity* is a more general four-dimensional Lorentz manifold $(M, \langle \cdot, \cdot \rangle)$, also called space-time. In the case of general relativity, the components g_{ij} of the metric are regarded as potentials for the gravitational forces.

In either case, points of space-time can be thought of as *events* that happen at a given place in space and at a given time. The trajectory of a moving particle can be regarded as curve of events, called its *world line*.

If p is an event in a Lorentz manifold $(M, \langle \cdot, \cdot \rangle)$, the tangent space T_pM inherits a Lorentz inner product

$$\langle \cdot, \cdot \rangle_p : T_pM \times T_pM \longrightarrow \mathbb{R}.$$

We say that an element $v \in T_pM$ is

1. *timelike* if $\langle v, v \rangle < 0$,
2. *spacelike* if $\langle v, v \rangle > 0$, and
3. *lightlike* if $\langle v, v \rangle = 0$.

A parametrized curve $\gamma : [a, b] \to M$ into a Lorentz manifold $(M, \langle \cdot, \cdot \rangle)$ is said to be timelike if $\gamma'(u)$ is timelike for all $u \in [a, b]$. If a parametrized curve $\gamma : [a, b] \to M$ represents the world line of a massive object, it is timelike and the integral

$$L(\gamma) = \frac{1}{c} \int_a^b \sqrt{-\langle \gamma'(u) \gamma'(u) \rangle} du \tag{1.22}$$

is the *elapsed time* measured by a clock moving along the world line γ. We call $L(]gamma)$ the *proper time* of γ.

The Twin Paradox. The fact that elapsed time is measured by the integral (1.22) has counterintuitive consequences. Suppose that $\gamma : [a, b] \to \mathbb{L}^4$ is a timelike curve in four-dimensional Minkowski space-time, parametrized so that

$$\gamma(t) = (t, x^1(t), x^2(t), x^3(t)).$$

Then

$$\gamma'(t) = \frac{\partial}{\partial t} + \sum_{i=1}^{3} \frac{dx^i}{dt} \frac{\partial}{\partial x^i}, \quad \text{so} \quad \langle \gamma'(t), \gamma'(t) \rangle = -c^2 + \sum_{i=1}^{3} \left(\frac{dx^i}{dt} \right)^2,$$

and hence

$$L(\gamma) = \int_a^b \frac{1}{c}\sqrt{c^2 - \sum_{i=1}^3 \left(\frac{dx^i}{dt}\right)^2}dt = \int_a^b \sqrt{1 - \frac{1}{c^2}\sum_{i=1}^3 \left(\frac{dx^i}{dt}\right)^2}dt. \tag{1.23}$$

Thus if a clock is at rest with respect to the coordinates, that is $dx^i/dt \equiv 0$, it will measure the time interval $b - a$, while if it is in motion it will measure a somewhat shorter time interval. This failure of clocks to synchronize is what is called the twin paradox.

Equation (1.23) states that in Minkowski space-time, straight lines maximize L among all timelike world lines from an event p to an event q. When given an affine parametrization such curves have zero acceleration. One might hope that in general relativity, the world line of a massive body, not subject to any forces other than gravity, would also maximize L, and if it was appropriately parametrized, would have zero acceleration in terms of the Lorentz metric $\langle \cdot, \cdot \rangle$. Just as in the Riemannian case, it is easier to describe the critical point behavior of the closely related *action*

$$J : \Omega_{[a,b]}(M : p, q) \to \mathbb{R}, \quad \text{defined by} \quad J(\gamma) = \frac{1}{2}\int_a^b \langle \gamma'(t), \gamma'(t) \rangle dt.$$

The critical points of J are called *geodesics*.

How does one determine the geodesics in a Lorentz manifold? Fortunately, the fundamental theorem of Riemannian geometry generalizes immediately to pseudo-Riemannian metrics;

Fundamental Theorem of pseudo-Riemannian Geometry. *If $\langle \cdot, \cdot \rangle$ is a pesudo-Riemannian metric on a smooth manifold M, there is a unique connection on TM such that*

1. *∇ is symmetric, that is, $\nabla_X Y - \nabla_Y X = [X, Y]$, for $X, Y \in \mathcal{X}(M)$,*
2. *∇ is metric, that is, $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$, for $X, Y, Z \in \mathcal{X}(M)$.*

The proof is identical to the proof we gave before. Moreover, just as before, we can define the Christoffel symbols for local coordinates, and they are given by exactly the same formula (1.18). Finally, by the first variation formula, one shows that a smooth parametrized curve $\gamma : [a, b] \to M$ is a geodesics if and only if it satisfies the equation $\nabla_{\gamma'}\gamma' \equiv 0$.

As described in more detail in [35], there are two main components to general relativity: The Einstein field equations describe how the distribution of matter in the universe determines a Lorentz metric on space-time, while timelike geodesics are exactly the world lines of massive objects which are subjected to no forces other than gravity. Lightlike geodesics are the trajectories of light rays.

1.8 The Riemann-Christoffel curvature tensor

Let $(M, \langle \cdot, \cdot \rangle)$ be a Riemannian manifold (or more generally a pseudo-Riemannian manifold) with Levi-Civita connection ∇. If $\mathcal{X}(M)$ denotes the space of smooth vector fields on M, we define

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M)$$

by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

We call R the *Riemann-Christoffel curvature tensor* of $(M, \langle \cdot, \cdot \rangle)$.

Proposition 1. *The operator R is multilinear over functions, that is,*

$$R(fX,Y)Z = R(X,fY)Z = R(X,Y)fZ = fR(X,Y)Z.$$

Proof: We prove only the equality $R(X,Y)fZ = fR(X,Y)Z$, leaving the others as easy exercises:

$$\begin{aligned} R(X,Y)fZ &= \nabla_X \nabla_Y (fZ) - \nabla_Y \nabla_X (fZ) - \nabla_{[X,Y]}(fZ) \\ &= \nabla_X ((Yf)Z + f\nabla_Y Z) - \nabla_Y ((Xf)Z + f\nabla_X Z) - [X,Y](f)Z - f\nabla_{[X,Y]}(Z) \\ &= XY(f)Z + (Yf)\nabla_X Z + (Xf)\nabla_Y Z + f\nabla_X \nabla_Y Z \\ &\quad - YX(f)Z - (Xf)\nabla_Y Z - (Yf)\nabla_X Z - f\nabla_Y \nabla_Y Z \\ &\quad - [X,Y](f)Z - f\nabla_{[X,Y]}(Z) \\ &= f(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) = fR(X,Y)Z. \end{aligned}$$

Since the connection ∇ can be localized by Lemma 5.1, so can the curvature; that is, if U is an open subset of M, $(R(X,Y)Z)|U$ depends only $X|U$, $Y|U$ and $Z|U$. Thus the curvature tensor is determined in local coordinates by its component functions $R^l_{ijk} : U \to \mathbb{R}$, defined by the equations

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} = \sum_{l=1}^{n} R^l_{ijk} \frac{\partial}{\partial x^l}.$$

Proposition 2. *The components R^l_{ijk} of the Riemann-Christoffel curvature tensor are determined from the Christoffel symbols Γ^l_{jk} by the equations*

$$R^l_{ijk} = \frac{\partial}{\partial x^i}(\Gamma^l_{jk}) - \frac{\partial}{\partial x^j}(\Gamma^l_{ik}) + \sum_{m=1}^{n} \Gamma^l_{mi}\Gamma^m_{jk} - \sum_{m=1}^{n} \Gamma^l_{mj}\Gamma^m_{ik}.$$

The proof is a straightforward computation.

The simplest example of course is Euclidean space $\mathbb{E}^N$. In this case, the metric coefficients (g_{ij} are constant, and hence it follows from (1.18) that the Christoffel symbols $\Gamma^k_{ij} = 0$. Thus it follows from Proposition 8.2 that the curvature tensor R is identically zero. Recall that in this case, the Levi-Civita connection ∇^E on $\mathbb{E}^N$ is given by the simple formula

$$\nabla^E_X \left(\sum_{i=1}^N f^i \frac{\partial}{\partial x^i} \right) = \sum_{i=1}^N X(f^i) \frac{\partial}{\partial x^i}.$$

It is often easy to calculate the curvature of submanifolds on $\mathbb{E}^N$ with the induced Riemannian metric by means of the so-called Gauss equation, as we now explain. Thus suppose that $\iota : M \to \mathbb{E}^N$ is an imbedding and agree to identify $p \in M$ with $\iota(p) \in \mathbb{E}^N$ and $v \in T_pM$ with its image $\iota_*(v) \in T_p\mathbb{E}^N$. If $p \in M$ and $v \in T_p\mathbb{E}^N$, we let

$$v = v^\top + v^\perp, \quad \text{where} \quad v^\top \in T_pM \quad \text{and} \quad v^\perp \perp T_pM.$$

Thus $(\cdot)^\top$ is the orthogonal projection into the tangent space and $(\cdot)^\perp$ is the orthogonal projection into the normal space, the orthogonal complement to the tangent space. We have already noted we can define the Levi-Civita connection $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ by the formula

$$(\nabla_X Y)(p) = (\nabla^E_X Y(p))^\top.$$

If we let $\mathcal{X}^\perp(M)$ denote the space of vector fields in $\mathbb{E}^N$ which are defined at points of M and are perpendicular to M, then we can define

$$\alpha : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}^\perp(M) \quad \text{by} \quad \alpha(X, Y) = (\nabla^E_X Y(p))^\perp.$$

We call α the *second fundamental form* of M in $\mathbb{E}^N$.

Proposition 3. *The second fundamental form satisfies the identities:*

$$\alpha(fX, Y) = \alpha(X, fY) = f\alpha(X, Y), \quad \alpha(X, Y) = \alpha(Y, X).$$

Indeed,

$$\alpha(fX, Y) = (\nabla^E_{fX} Y)^\perp = f(\nabla^E_X Y)^\perp = f\alpha(X, Y),$$

$$\alpha(X, fY) = (\nabla^E_X)(fY))^\perp = ((Xf)Y + f\nabla^E_X Y)^\perp = f\alpha(X, Y),$$

so α is bilinear over functions. It therefore suffices to establish $\alpha(X, Y) = \alpha(Y, X)$ in the case where $[X, Y] = 0$, but in this case

$$\alpha(X, Y) - \alpha(Y, X) = (\nabla^E_X Y - \nabla^E_Y X)^\perp = 0.$$

There is some special terminology that is used in the case where $\gamma : (a, b) \to M \subseteq \mathbb{E}^N$ is a unit speed curve. In this case, we say that the acceleration $\gamma''(t) \in T_{\gamma(t)}\mathbb{E}^N$ is the *curvature* of γ, while

$$(\gamma''(t))^\top = (\nabla_{\gamma'}\gamma')(t) = (\textit{geodesic curvature} \text{ of } \gamma \text{ at } t),$$

$$(\gamma''(t))^\perp = \alpha(\gamma'(t), \gamma'(t)) = (\textit{normal curvature} \text{ of } \gamma \text{ at } t).$$

Thus if $x \in T_pM$ is a unit length vector, $\alpha(x, x)$ can be interpreted as the normal curvature of some curve tangent to x at p.

Gauss Theorem. The curvature tensor R of a submanifold $M \subseteq \mathbb{E}^N$ is given by the *Gauss equation*

$$\langle R(X,Y)W, Z\rangle = \alpha(X,Z) \cdot \alpha(Y,W) - \alpha(X,W) \cdot \alpha(Y,Z), \tag{1.24}$$

where X, Y, Z and W are elements of $\mathcal{X}(M)$, and the dot on the right denotes the Euclidean metric in the ambient space $\mathbb{E}^N$.

Proof: Since Euclidean space has zero curvature,

$$\nabla^E_X \nabla^E_Y W - \nabla^E_Y \nabla^E_X W - \nabla^E_{[X,Y]} W = 0,$$

and hence

$$\begin{aligned}
0 &= (\nabla^E_X \nabla^E_Y W) \cdot Z - (\nabla^E_Y \nabla^E_X W) \cdot Z - \nabla^E_{[X,Y]} W \cdot Z \\
&= X(\nabla^E_Y W \cdot Z) - \nabla^E_Y W \cdot \nabla^E_X Z - Y(\nabla^E_X W \cdot Z) + \nabla^E_X W \cdot \nabla^E_Y Z - \nabla^E_{[X,Y]} W \cdot Z \\
&= X\langle \nabla_Y W, Z\rangle - \langle \nabla_Y W, \nabla_X Z\rangle - \alpha(Y,W) \cdot \alpha(X,Z) \\
&\quad - Y\langle \nabla_X W, Z\rangle - \langle \nabla_X W, \nabla_Y Z\rangle - \alpha(X,W) \cdot \alpha(Y,Z) - \langle \nabla_{[X,Y]} W, Z\rangle.
\end{aligned}$$

Thus we find that

$$\begin{aligned}
0 = \langle \nabla_X \nabla_Y W, Z\rangle &- \alpha(Y,W) \cdot \alpha(X,W) \\
&- \langle \nabla_Y \nabla_X W, Z\rangle + \alpha(X,W) \cdot \alpha(Y,W) - \langle \nabla_{[X,Y]} W, Z\rangle.
\end{aligned}$$

This yields

$$\langle \nabla_X \nabla_Y W - \nabla_Y \nabla_X W - \nabla_{[X,Y]} W, Z\rangle = \alpha(Y,W) \cdot \alpha(X,W) - \alpha(X,W) \cdot \alpha(Y,W),$$

which is exactly (1.24).

For example, we can consider the sphere of radius a about the origin in $\mathbb{E}^{n+1}$:

$$\mathbb{S}^n(a) = \{(x^1, \ldots, x^{n+1}) \in \mathbb{E}^{n+1} : (x^1)^2 + \cdots + (x^{n+1})^2 = a^2\}.$$

If $\gamma : (-\epsilon, \epsilon) \to \mathbb{S}^n \subseteq \mathbb{E}^{n+1}$ is a unit speed great circle, say

$$\gamma(t) = a\cos((1/a)t)\mathbf{e}_1 + a\sin((1/a)t)\mathbf{e}_2,$$

where $(\mathbf{e}_1, \mathbf{e}_2)$ are orthonormal vectors located at the origin in $\mathbb{E}^{n+1}$, then a direct calculation shows that

$$\gamma''(t) = -\frac{1}{a}\mathbf{N}(\gamma(t)),$$

where $\mathbf{N}(p)$ is the outward pointing unit normal to $\mathbb{S}^n(a)$ at the point $p \in \mathbb{S}^n(a)$. Thus the second fundamental form of $\mathbb{S}^n(a)$ in $\mathbb{E}^{n+1}$ satisfies

$$\alpha(x,x) = -\frac{1}{a}\mathbf{N}(p), \quad \text{for all unit length } x \in T_p\mathbb{S}^n(a).$$

If x does not have unit length, then

$$\alpha\left(\frac{x}{\|x\|}, \frac{x}{\|x\|}\right) = -\frac{1}{a}\mathbf{N}(p) \quad \Rightarrow \quad \alpha(x,x) = -\frac{1}{a}\langle x, x\rangle \mathbf{N}(p).$$

By polarization, we obtain

$$\alpha(x,y) = \frac{-1}{a}\langle x, y\rangle \mathbf{N}(p), \quad \text{for all } x, y \in T_p\mathbb{S}^n(a).$$

Thus substitution into the Gauss equation yields

$$\begin{aligned}\langle R(x,y)w, z\rangle = {} & \left(\frac{-1}{a}\langle x, z\rangle \mathbf{N}(p)\right) \cdot \left(\frac{-1}{a}\langle y, w\rangle \mathbf{N}(p)\right) \\ & - \left(\frac{-1}{a}\langle x, w\rangle \mathbf{N}(p)\right) \cdot \left(\frac{-1}{a}\langle y, x\rangle \mathbf{N}(p)\right).\end{aligned}$$

Thus we finally obtain a formula for the curvature of $\mathbb{S}^n(a)$:

$$\langle R(x,y)w, z\rangle = \frac{1}{a^2}(\langle x, z\rangle\langle y, w\rangle - \langle x, w\rangle\langle y, z\rangle).$$

In a similar fashion, one can sometimes calculate the curvature of spacelike hypersurfaces in Minkowski space-time. The metric coefficients for Minkowski space time $\mathbb{L}^{n+1}$ are constant, so once again $\Gamma^k_{ij} = 0$ and the curvature of Minkowski space-time is zero. In this case, the Levi-Civita connection ∇^L is defined by

$$\nabla^L_X\left(f^0\frac{\partial}{\partial t} + \sum_{i=1}^N f^i \frac{\partial}{\partial x^i}\right) = X(f^0)\frac{\partial}{\partial t} + \sum_{i=1}^N X(f^i)\frac{\partial}{\partial x^i}.$$

Suppose that M is an n-dimensional manifold and $\iota : M \to \mathbb{L}^{n+1}$ is an imbedding. We say that $\iota(M)$ is a *spacelike hypersurface* if the standard Lorentz metric on $\mathbb{L}^{n+1}$ induces a positive-definite Riemannian metric on M. For simplicity, let us set the speed of light $c = 1$ so that the Lorentz metric on $\mathbb{L}^{n+1}$ is simply

$$\langle \cdot, \cdot \rangle_L = -dt \otimes dt + \sum_{i=1}^n dx^i \otimes dx^i.$$

Just as in the case where the ambient space is Euclidean space, we find that the Levi-Civita connection $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ on TM is given by the formula

$$(\nabla_X Y)(p) = (\nabla^L_X Y(p))^\top,$$

where $(\cdot)^\top$ is the orthogonal projection into the tangent space. If we let $\mathcal{X}^\perp(M)$ denote the vector field in $\mathbb{L}^N$ which are defined at points of M and are perpendicular to M, we can define the second fundamental form of M in $\mathbb{L}^{n+1}$ by

$$\alpha : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}^\perp(M) \quad \text{by} \quad \alpha(X, Y) = (\nabla^E_X Y(p))^\perp,$$

where $(\cdot)^\perp$ is the orthogonal projection to the orthogonal complement to the tangent space. Moreover, the curvature of the spacelike hypersurface is given by the Gauss equation

$$\langle R(X, Y)W, Z\rangle = \langle \alpha(X, Z), \alpha(Y, W)\rangle_L - \langle \alpha(X, W), \alpha(Y, Z)\rangle_L, \tag{1.25}$$

where X, Y, Z and W are elements of $\mathcal{X}(M)$.

As a key example, we can set

$$\mathbb{H}^n(a) = \{(t, x^1 \ldots, x^n) \in \mathbb{L}^{n+1} : t^2 - (x^1)^2 - \cdots - (x^n)^2 = a^2,\ t > 0\},$$

the set of future-pointed unit timelike vectors situated at the origin in $\mathbb{L}^{n+1}$. Clearly $\mathbb{H}^n(a)$ is an imbedded submanifold of $\mathbb{L}^{n+1}$ and we claim that the induced metric on $\mathbb{H}^n(a)$ is positive-definite.

To prove this, we could consider $(x^1, \ldots, x^n)$ as global coordinates on $\mathbb{H}^n(a)$, so that

$$t = \sqrt{a^2 + (x^1)^2 + \cdots + (x^n)^2}.$$

Then

$$dt = \frac{\sum_{i=1}^n x^i dx^i}{\sqrt{a^2 + (x^1)^2 - \cdots + (x^n)^2}},$$

and the induced metric on $\mathbb{H}^n(a)$ is

$$\langle \cdot, \cdot \rangle = -\frac{\sum x^i x^j dx^i \otimes dx^j}{a^2 + (x^1)^2 - \cdots + (x^n)^2} + dx^1 \otimes dx^1 + \cdots + dx^n \otimes dx^n.$$

Thus

$$g_{ij} = \delta_{ij} - \frac{\sum x^i x^j}{a^2 + (x^1)^2 - \cdots + (x^n)^2},$$

and from this expression we immediately see that the induced metric on $H^n(a)$ is indeed positive-definite.

Of course, $\mathbb{H}^n(a)$ is nothing other than the upper sheet of a hyperboloid of two sheets. Suppose that $p \in \mathbb{H}^n(a)$, that $\mathbf{e}_0$ is a future-pointing unit length timelike vector such that $p = a\mathbf{e}_0$ and Π is a two-dimensional plane that passes through the origin and contains $\mathbf{e}_0$. Using elementary linear algebra, Π must also contain a unit length spacelike vector $\mathbf{e}_1$ such that $\langle \mathbf{e}_0, \mathbf{e}_1\rangle_L = 0$. Then the smooth curve

$$\gamma : (-\epsilon, \epsilon) \to \mathbb{H}^n(a) \quad \text{defined by} \quad \gamma(t) = a\cosh(t/a)\mathbf{e}_0 + a\sinh(t/a)\mathbf{e}_1$$

is spacelike and direct calculation shows that

$$\langle \gamma'(t), \gamma'(t)\rangle_L = -(\sinh(t/a))^2 + (\cosh(t/a))^2 = 1.$$

Moreover,

$$\gamma''(t) = \frac{1}{a}(\cosh(t/a)\mathbf{e}_0 + \sinh(t/a)\mathbf{e}_1) = \frac{1}{a}\mathbf{N}(\gamma(t)),$$

where $\mathbf{N}(p)$ is the unit normal to $\mathbb{H}^n(a)$ at p. Thus

$$(\nabla_{\gamma'}\gamma')(t) = (\gamma''(t))^\top = 0,$$

so γ is a geodesic and

$$\alpha(\gamma'(t), \gamma'(t)) = (\gamma''(t))^\perp = \frac{1}{a}\mathbf{N}(\gamma(t)).$$

Note that we can construct a unit speed geodesic γ in M as above with $\gamma(0) = p$ for any $p \in H^n(a)$ and $\gamma'(0) = \mathbf{e}_1$ for any unit length $\mathbf{e}_1 \in T_p\mathbb{H}^n(a)$.

Thus just as in the case of the sphere, we can use the Gauss equation (1.25) to determine the curvature of $\mathbb{H}^n(a)$. Thus the second fundamental form of $\mathbb{H}^n(a)$ in $\mathbb{L}^{n+1}$ satisfies

$$\alpha(x, x) = -\frac{1}{a}\mathbf{N}(p), \quad \text{for all unit length } x \in T_p\mathbb{H}^n(a),$$

where $\mathbf{N}(p)$ is the future-pointing unit normal to M. If x does not have unit length, then

$$\alpha\left(\frac{x}{\|x\|}, \frac{x}{\|x\|}\right) = \frac{1}{a}\mathbf{N}(p) \quad \Rightarrow \quad \alpha(x, x) = \frac{1}{a}\langle x, x\rangle\mathbf{N}(p).$$

By polarization, we obtain

$$\alpha(x, y) = \frac{1}{a}\langle x, y\rangle\mathbf{N}(p), \quad \text{for all } x, y \in T_p\mathbb{N}^n(a).$$

Thus substitution into the Gauss equation yields

$$\begin{aligned}\langle R(x, y)w, z\rangle = &\left(\frac{1}{a}\langle x, z\rangle\mathbf{N}(p)\right) \cdot \left(\frac{1}{a}\langle y, w\rangle\mathbf{N}(p)\right) \\ &\qquad - \left(\frac{1}{a}\langle x, w\rangle\mathbf{N}(p)\right) \cdot \left(\frac{1}{a}\langle y, x\rangle\mathbf{N}(p)\right).\end{aligned}$$

Since $\mathbf{N}(p)$ is timelike and hence $\langle\mathbf{N}(p), \mathbf{N}(p)\rangle_L = -1$, we finally obtain a formula for the curvature of $\mathbb{S}^n(a)$:

$$\langle R(x, y)w, z\rangle = \frac{-1}{a^2}(\langle x, z\rangle\langle y, w\rangle - \langle x, w\rangle\langle y, z\rangle).$$

The Riemannian manifold $\mathbb{H}^n(a)$ is called *hyperbolic space*, and its geometry is called *hyperbolic geometry*. We have constructed a model for hyperbolic geometry, the upper sheet of the hyperboloid of two sheets in $\mathbb{L}^{n+1}$, and have seen that the geodesics in this model are just the intersections with two-planes passing through the origin in $\mathbb{L}^{n+1}$.

1.9 Curvature symmetries; sectional curvature

The Riemann-Christoffel curvature tensor is the basic local invariant of a pseudo-Riemannian manifold. If M has dimension n, one would expect R to have n^4 independent components R^l_{ijk}, but the number of independent components is cut down considerably because of the curvature symmetries:

Proposition 1. *The curvature tensor R of a pseudo-Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ satisfies the identities:*

1. $R(X, Y) = -R(Y, X)$,
2. $R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$,
3. $\langle R(X, Y)W, Z \rangle = -\langle R(X, Y)Z, W \rangle$, *and*
4. $\langle R(X, Y)W, Z \rangle = -\langle R(W, Z)X, Y \rangle$.

Remark 1. If we assumed the Nash imbedding theorem (in the Riemannian case), we could derive these identities immediately from the Gauss equation (1.24).

Remark 2. We can write the above curvature symmetries in terms of the components R^l_{ijk} of the curvature tensor. Actually, it is easier to express these symmetries if we lower the index and write

$$R_{ijlk} = \sum_{p=1}^{n} g_{lp} R^l_{ijk}.$$

This lowering of the index into the third position is consistent with regarding the R_{ijlk}'s as the components of the map

$$R : T_pM \times T_pM \times T_pM \times T_pM \to \mathbb{R} \quad \text{by} \quad R(X, Y, Z, W) = \langle R(X, Y)W, Z \rangle.$$

In terms of these components, the curvature symmetries are

$$R_{ijlk} = -R_{jilk}, \quad R_{ijlk} + R_{jkli} + R_{kilj} = 0,$$
$$R_{ijlk} = -R_{ijkl}, \quad R_{ijlk} = R_{lkij}.$$

Proof of proposition: Note first that since R is a tensor, we can assume without loss of generality that all brackets of X, Y, Z and W are zero. Then

$$R(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X = -(\nabla_Y \nabla_X - \nabla_X \nabla_Y) = -R(Y, X),$$

establishing the first identity. Next,

$$\begin{aligned} R(X, Y)Z + R(Y, Z)X + R(Z, X)Y &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z \\ &\quad + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X + \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y \\ &= \nabla_X(\nabla_Y Z - \nabla_Z Y) + \nabla_Y(\nabla_Z X - \nabla_X Z) + \nabla_Z(\nabla_X Y - \nabla_Y X) = 0, \end{aligned}$$

the last equality holding because ∇ is symmetric. For the third identity, we calculate

$$\begin{aligned}\langle R(X,Y)Z,Z\rangle &= \langle \nabla_X\nabla_Y Z, Z\rangle - \langle \nabla_Y\nabla_X Z, Z\rangle \\ &= X\langle \nabla_Y Z, Z\rangle - \langle \nabla_Y Z, \nabla_X Z\rangle - Y\langle \nabla_X Z, Z\rangle + \langle \nabla_X Z, \nabla_Y Z\rangle \\ &= \frac{1}{2}XY\langle Z,Z\rangle - \frac{1}{2}YX\langle Z,Z\rangle = \frac{1}{2}[X,Y]\langle Z,Z\rangle = 0.\end{aligned}$$

Hence the symmetric part of the bilinear form

$$(W,Z) \mapsto \langle R(X,Y)W,Z\rangle$$

is zero, from which the third identity follows. Finally, it follows from the first and second identities that

$$\langle R(X,Y)W,Z\rangle = -\langle R(Y,X)W,Z\rangle = \langle R(X,W)Y,Z\rangle + \langle R(W,Y)X,Z\rangle,$$

and from the third and second that

$$\langle R(X,Y)W,Z\rangle = -\langle R(X,Y)Z,W\rangle = \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle.$$

Adding the last two expressions yields

$$\begin{aligned}2\langle R(X,Y)W,Z\rangle &= \langle R(X,W)Y,Z\rangle \\ &\quad + \langle R(W,Y)X,Z\rangle + \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle. \end{aligned} \tag{1.26}$$

Exchanging the pair (X,Y) with (W,Z) yields

$$\begin{aligned}2\langle R(W,Z)X,Y\rangle &= \langle R(W,X)Z,Y\rangle \\ &\quad + \langle R(X,Z)W,Y\rangle + \langle R(Z,Y)W,X\rangle + \langle R(Y,W)Z,X\rangle. \end{aligned} \tag{1.27}$$

Each term on the right of (4.29) equals one of the terms on the right of (1.27), so

$$\langle R(X,Y)W,Z\rangle = \langle R(W,Z)X,Y\rangle,$$

finishing the proof of the proposition.

Proposition 2. *Let*

$$R, S : T_pM \times T_pM \times T_pM \times T_pM \to \mathbb{R}$$

be two quadrilinear functions which satisfy the curvature symmetries. If

$$R(x,y,x,y) = S(x,y,x,y), \quad \textit{for all } x, y \in T_pM,$$

then $R = S$.

Proof: Let $T = R - S$. Then T satisfies the curvature symmetries and

$$T(x,y,x,y) = 0, \quad \text{for all } x, y \in T_pM.$$

Hence

$$\begin{aligned} 0 = T(x, y+z, x, y+z) & \\ = T(x,y,x,y) + T(x,y,x,z) + T(x,z,x,y) + T(x,z,x,z) & \\ = 2T(x,y,x,z), & \end{aligned}$$

so $T(x,y,x,z) = 0$. Similarly,

$$0 = T(x+z, y, x+z, w) = T(x,y,z,w) + T(z,y,x,w),$$

$$0 = T(x+w, y, z, x+w) = T(x,y,z,w) + T(w,y,z,x).$$

Finally,

$$\begin{aligned} 0 &= 2T(x,y,z,w) + T(z,y,x,w) + T(w,y,z,x) \\ &= 2T(x,y,z,w) - T(y,z,x,w) - T(z,y,x,w) = 3T(x,y,z,w). \end{aligned}$$

So $T = 0$ and $R = S$.

This proposition shows that the curvature is completely determined by the sectional curvatures, defined as follows:

Definition. Suppose that σ is a two-dimensional subspace of T_pM such that the restriction of $\langle \cdot, \cdot \rangle$ to σ is nondegenerate. Then the *sectional curvature* of σ is

$$K(\sigma) = \frac{\langle R(x,y)y, x\rangle}{\langle x,x\rangle\langle y,y\rangle - \langle x,y\rangle^2},$$

whenever (x,y) is a basis for σ. The curvature symmetries imply that $K(\sigma)$ is independent of the choice of basis.

Recall our key three examples, the so-called *spaces of constant curvature*:

$$\begin{cases} \text{If } M = \mathbb{E}^n, \text{ then } K(\sigma) \equiv 0 \text{ for all two-planes } \sigma \subseteq T_pM. \\ \text{If } M = \mathbb{S}^n(a), \text{ then } K(\sigma) \equiv 1/a^2 \text{for all two-planes } \sigma \subseteq T_pM. \\ \text{If } M = \mathbb{H}^n(a), \text{ then } K(\sigma) \equiv -1/a^2 \text{ for all two-planes } \sigma \subseteq T_pM. \end{cases}$$

The spaces of constant curvature are the most symmetric Riemannian manifolds possible.

Definition. If $(M, \langle \cdot, \cdot \rangle)$ is a pseudo-Riemannian manifold, a diffeomorphism $\phi : M \to M$ is said to be an *isometry* if

$$\langle (\phi_*)_p(v), (\phi_*)_p(w) \rangle = \langle v, w \rangle, \quad \text{for all } v, w \in T_pM \text{ and all } p \in M. \tag{1.28}$$

Of course, we can rewrite (1.28) as $\phi^* \langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle$, where

$$\phi^* \langle v, w \rangle = \langle (\phi_*)_p(v), (\phi_*)_p(w) \rangle, \quad \text{for } v, w \in T_pM.$$

Note that the orthogonal group $O(n+1)$ acts as a group of isometries on $\mathbb{S}^n(a)$. In this case, we have an isometry group of dimension $(1/2)n(n+1)$. Similarly, the group of isometries of $\mathbb{E}^n$, the group of Euclidean motions, is a Lie group of dimension $(1/2)n(n+1)$. The group of isometries on $\mathbb{H}^n(a)$ also has dimension $(1/2)n(n+1)$; it is called the *Lorentz group*.

In each of the three cases there is an isometry ϕ which takes any point p of M to any other point q and any orthonormal basis of T_pM to any orthonormal basis of T_qM. This allows us to construct non-Euclidean geometries for $\mathbb{S}^n(a)$ and $\mathbb{H}^n(a)$ which are quite similar to Euclidean geometry. In the case of $\mathbb{H}^n(a)$ all the postulates of Euclidean geometry are satisfied except for the parallel postulate.

1.10 Gaussian curvature of surfaces

We now make contact with the theory of surfaces in $\mathbb{E}^3$ as described in undergraduate texts such as [29]. If $(M, \langle \cdot, \cdot \rangle)$ is a two-dimensional Riemannian manifold, then there is only one two-plane at each point p, namely T_pM. In this case, we can define a smooth function $K : M \to \mathbb{R}$ by

$$K(p) = K(T_pM) = (\text{sectional curvature of } T_pM).$$

The function K is called the *Gaussian curvature* of M.

An important special case is that of a two-dimensional smooth surface M^2 imbedded in $\mathbb{R}^3$, with M^2 given the induced Riemannian metric. We assume that it is possible to choose a smooth unit normal $\mathbf{N}$ to M,

$$\mathbf{N} : M^2 \to \mathbb{S}^2, \quad \text{with } \mathbf{N}(p) \text{ perpendicular to } T_pM.$$

Such a choice of unit normal determines an orientation of M^2.

If N_pM is the orthogonal complement to T_PM, then the second fundamental form $\alpha : T_pM \times T_pM \to N_pM$ determined a symmetric bilinear form $h : T_pM \times T_pM \to \mathbb{R}$ by the formula

$$h(x, y) = \alpha(x, y) \cdot \mathbf{N}(p), \quad \text{for } x, y \in T_pM,$$

which is also called the *second fundamental form* in the theory of surfaces.

Recall that if (x^1, x^2) is a smooth coordinate system on M, we can define the components of the induced Riemannian metric on M^2 by the formulae

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle, \quad \text{for } i, j = 1, 2.$$

If $F : M^2 \to \mathbb{E}^3$ is the imbedding than the components of the induced Riemannian metric (also called the first fundamental form) are given by the formula

$$g_{ij} = \frac{\partial F}{\partial x^i} \cdot \frac{\partial F}{\partial x^j}.$$

Similarly, we can define the components of the second fundamental form by

$$h_{ij} = h\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right), \quad \text{for } i, j = 1, 2.$$

These components can be found by the explicit formula

$$h_{ij} = \left(\nabla^E_{\partial/\partial x^i} \frac{\partial}{\partial x^j}\right) \cdot \mathbf{N} = \frac{\partial^2 F}{\partial x^i \partial x^j} \cdot \mathbf{N}.$$

Let

$$X = \frac{\partial}{\partial x^1}, \quad Y = \frac{\partial}{\partial x^2}.$$

Then it follows from the definition of Gaussian curvature and the Gauss equation that

$$\begin{aligned} K &= \frac{\langle RX, Y)Y, X\rangle}{\langle X, X\rangle\langle Y, Y\rangle - \langle X, Y\rangle^2} \\ &= \frac{\alpha(X, X) \cdot \alpha(Y, Y) - \alpha(X, Y) \cdot \alpha(X, Y)}{\langle X, X\rangle\langle Y, Y\rangle - \langle X, Y\rangle^2} \\ &= \frac{h_{11}h_{22} - h_{12}^2}{g_{11}g_{22} - g_{12}^2} = \frac{\begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix}}{\begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix}}. \end{aligned}$$

Example. Let us consider the *catenoid*, the submanifold of $\mathbb{R}^3$ defined by the equation

$$r = \sqrt{x^2 + y^2} = \cosh z,$$

where (r, θ, z) are cylindrical coordinates. This is obtained by rotating the catenary around the z-axis. As parametrization, we can take $M^2 = \mathbb{R} \times S^1$ and

$$F : \mathbb{R} \times S^1 \to S \quad \text{by} \quad \mathbf{x}(u, v) = \begin{pmatrix} \cosh u \cos v \\ \cosh u \sin v \\ u \end{pmatrix}.$$

Here v is the coordinate on S^1 which is just the quotient group $\mathbb{R}/\mathbb{Z}$, where $\mathbb{Z}$ is the cyclic group generated by 2π. Then

$$\frac{\partial F}{\partial u} = \begin{pmatrix} \sinh u \cos v \\ \sinh u \sin v \\ 1 \end{pmatrix} \quad \text{and} \quad \frac{\partial F}{\partial v} = \begin{pmatrix} -\cosh u \sin v \\ \cosh u \cos v \\ 0 \end{pmatrix},$$

and hence the coefficients of the first fundamental form in this case are

$$g_{11} = 1 + \sinh^2 u = \cosh^2 u, \quad g_{12} = 0, \quad \text{and} \quad g_{22} = \cosh^2 u.$$

The induced Riemannian metric (or first fundamental form) in this case is

$$\langle \cdot, \cdot \rangle = \cosh^2 u (du \otimes du + dv \otimes dv).$$

To find a unit normal, we first calculate

$$\frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} = \begin{vmatrix} \mathbf{i} & \sinh u \cos v & -\cosh u \sin v \\ \mathbf{j} & \sinh u \sin v & \cosh u \cos v \\ \mathbf{k} & 1 & 0 \end{vmatrix} = \begin{pmatrix} -\cosh u \cos v \\ -\cosh u \sin v \\ \cosh u \sinh u \end{pmatrix}.$$

Thus a unit normal to S can be given by the formula

$$\mathbf{N} = \frac{\frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v}}{\left| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right|} = \frac{1}{\cosh u} \begin{pmatrix} -\cos v \\ -\sin v \\ \sinh u \end{pmatrix}.$$

To calculate the second fundamental form, we need the second order partial derivatives,

$$\frac{\partial^2 F}{\partial u^2} = \begin{pmatrix} \cosh u \cos v \\ \cosh u \sin v \\ 0 \end{pmatrix}, \quad \frac{\partial^2 F}{\partial u \partial v} = \begin{pmatrix} -\sinh u \sin v \\ \sinh u \cos v \\ 0 \end{pmatrix},$$

and

$$\frac{\partial^2 \mathbf{x}}{\partial v^2} = \begin{pmatrix} -\cosh u \cos v \\ -\cosh u \sin v \\ 0 \end{pmatrix}.$$

These give the coefficients of the second fundamental form

$$h_{11} = \frac{\partial^2 F}{\partial u^2} \cdot \mathbf{N} = 1, \qquad h_{12} = h_{21} = \frac{\partial^2 F}{\partial u \partial v} \cdot \mathbf{N} = 0,$$

and

$$h_{22} = \frac{\partial^2 F}{\partial v^2} \cdot \mathbf{N} = -1.$$

$$K = \frac{-1}{(\cosh u)^4}.$$

Exercise III. Consider the torus $M^2 = S^1 \times S^1$ with imbedding

$$F : U \to S \quad \text{by} \quad \mathbf{x}(u, v) = \begin{pmatrix} (2 + \cos u) \cos v \\ (2 + \cos u) \sin v \\ \sin u \end{pmatrix},$$

where u and v are the angular coordinates on the two S^1 factors, with $u+2\pi = u$, $v + 2\pi = v$.

a. Calculate the components g_{ij} of the induced Riemannian metric on M^2.

b. Calculate a continuously varying unit normal $\mathbf{N}$ and the components h_{ij} of the second fundamental form of M^2.

c. Determine the Gaussian curvature K.

1.11 Review of Lie groups

In addition to the spaces of constant curvature, there is another class of manifolds for which the geodesics and curvature can be computed relatively easily, the compact Lie groups with biinvariant Riemannian metrics. Before discussing this class of examples, we provide a brief review of Lie groups and Lie algebras, following Chapters 3 and 4 of [5].

Suppose now that G is a Lie group and $\sigma \in G$. We can then define the *left translation* by σ,

$$L_\sigma : G \to G \quad \text{by} \quad L_\sigma(\tau) = \sigma\tau,$$

a map which is clearly a diffeomorphism. Similarly, we can define *right translation*

$$R_\sigma : G \to G \quad \text{by} \quad R_\sigma(\tau) = \tau\sigma.$$

A vector field X on G is said to be *left invariant* if $(L_\sigma)_*(X) = X$ for all $\sigma \in G$, where

$$(L_\sigma)_*(X)(f) = X(f \circ L_\sigma) \circ L_\sigma^{-1}.$$

A straightforward calculation shows that if X and Y are left invariant vector fields on G, then so is their bracket $[X, Y]$. (See Theorem 7.9 in Chapter 4 of [5].) Thus the space

$$\mathfrak{g} = \{X \in \mathcal{X}(G) : \ (L_\sigma)_*(X) = X \text{ for all } \sigma \in G\ \}$$

is closed under Lie bracket, and the real bilinear map

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

is skew-symmetric (that is, $[X, Y] = -[Y, X]$), and satisfies the *Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Thus $\mathfrak{g}$ is a Lie algebra and we call it the *Lie algebra* of G. If e is the identity of the Lie group, restriction to T_eG yields an isomorphism $\alpha : \mathfrak{g} \to T_eG$. The inverse $\beta : T_eG \to \mathfrak{g}$ is defined by $\beta(v)(\sigma) = (L_\sigma)_*(v)$.

The most important examples of Lie groups are the *general linear group*

$$GL(n, \mathbb{R}) = \{\ n \times n \text{ matrices } A \text{ with real entries } : \det A \neq 0\},$$

and its subgroups. For $1 \leq i, j \leq n$, we can define coordinates

$$x^i_j : GL(n, \mathbb{R}) \to \mathbb{R} \quad \text{by} \quad x^i_j((a^i_j)) = a^i_j.$$

Of course, these are just the rectangular cartesian coordinates on an ambient Euclidean space in which $GL(n, \mathbb{R})$ sits as an open subset. If $X = (x^i_j) \in GL(n, \mathbb{R})$, left translation by X is a linear map, so is its own differential. Thus

$$(L_X)_* \left(\sum_{i,j=1}^n a^i_j \frac{\partial}{\partial x^i_j} \right) = \sum_{i,j,k=1}^n x^i_k a^k_j \frac{\partial}{\partial x^i_j}.$$

If we allow X to vary over $GL(n,\mathbb{R})$ we obtain a left invariant vector field

$$X_A = \sum_{i,j,k=1}^{n} a^i_j x^k_i \frac{\partial}{\partial x^k_j}$$

which is defined on $GL(n,\mathbb{R})$. It is the unique left invariant vector field on $GL(n,\mathbb{R})$ which satisfies the condition

$$X_A(I) = \sum_{i,j=1}^{n} a^i_j \left.\frac{\partial}{\partial x^i_j}\right|_I ,$$

where I is the identity matrix, the identity of the Lie group $GL(n,\mathbb{R})$. Every left invariant vector field on $GL(n,\mathbb{R})$ is obtained in this way, for some choice of $n \times n$ matrix $A = (a^i_j)$. A direct calculation yields

$$[X_A, X_B] = X_{[A,B]}, \quad \text{where} \quad [A,B] = AB - BA, \tag{1.29}$$

which gives an alternate proof that left invariant vector fields are closed under Lie brackets in this case. Thus the Lie algebra of $GL(n,\mathbb{R})$ is isomorphic to

$$\mathfrak{gl}(n,\mathbb{R}) \cong T_I G = \{\ n \times n \text{ matrices } A \text{ with real entries }\},$$

with the usual bracket of matrices as Lie bracket.

Exercise IV. Prove equation (1.29).

For a general Lie group G, if $X \in \mathfrak{g}$, the integral curve θ_X for X such that $\theta_X(0) = e$ satisfies the identity $\theta_X(s+t) = \theta_X(s) \cdot \theta_X(t)$ because the derivatives at $t = 0$ for fixed s are the same. From this fact, one easily concludes that θ_X extends to a Lie group homomorphism

$$\theta_X : \mathbb{R} \longrightarrow G.$$

We call θ_X the *one-parameter group* which corresponds to $X \in \mathfrak{g}$. Since the vector field X is left invariant, the curve

$$t \mapsto L_\sigma(\theta_X(t)) = \sigma\theta_X(t) = R_{\theta_X(t)}(\sigma)$$

is the integral curve for X which passes through σ at $t = 0$, and therefore the one-parameter group of diffeomorphisms on G corresponding to $X \in \mathfrak{g}$ is

$$\phi_t = R_{\theta_X(t)}, \quad \text{for } t \in \mathbb{R}.$$

In the case where $G = GL(n,\mathbb{R})$ the one-parameter groups are easy to describe. In this case, if $A \in \mathfrak{gl}(n,\mathbb{R})$, we claim that the corresponding one-parameter group is

$$\theta_A(t) = e^{tA} = I + tA + \frac{1}{2!}t^2A^2 + \frac{1}{2!}t^3A^3 + \cdots .$$

Indeed, it follows from the identity

$$\frac{d}{dt}(e^{tA}) = Ae^{tA} = e^{tA}A$$

that θ_A is an integral curve for the left invariant vector field determined by A and that $\theta'_A(0) = A$.

If G is a Lie subgroup of $GL(n,\mathbb{R})$, then its left invariant vector fields are defined by taking elements of $T_IG \subseteq T_IGL(n,\mathbb{R})$ and spreading them out over G by left translations of G. Thus the left invariant vector fields on G are just the restrictions of the elements of $\mathfrak{g}l(n,\mathbb{R})$ which are tangent to G.

We can use the one-parameter groups to determine which elements of $\mathfrak{g}l(n,\mathbb{R})$ are tangent to G at I. Consider, for example, the *orthogonal group*,

$$O(n) = \{A \in GL(n,\mathbb{R}) : A^T A = I\},$$

where $(\cdot)^T$ denotes transpose. Its Lie algebra is

$$\mathfrak{o}(n) = \{A \in \mathfrak{g}l(n,\mathbb{R}) : e^{tA} \in O(n) \text{ for all } t \in \mathbb{R}\ \}.$$

Differentiating the equation

$$(e^{tA})^T e^{tA} = I \quad \text{yields} \quad (e^{tA})^T A^T e^{tA} + (e^{tA})^T A e^{tA} = 0,$$

and evaluating at $t = 0$ yields a formula for the Lie algebra of the orthogonal group,

$$\mathfrak{o}(n) = \{A \in \mathfrak{g}l(n,\mathbb{R}) : A^T + A = 0\},$$

the Lie algebra of skew-symmetric matrices.

The complex general linear group,

$$GL(n,\mathbb{C}) = \{\ n \times n \text{ matrices } A \text{ with complex entries } : \det A \neq 0\},$$

is also frequently encountered, and its Lie algebra is

$$\mathfrak{g}l(n,\mathbb{C}) \cong T_eG = \{\ n \times n \text{ matrices } A \text{ with complex entries }\},$$

with the usual bracket of matrices as Lie bracket. It can be regarded as a Lie subgroup of $GL(2n,\mathbb{R})$. The *unitary group* is

$$U(n) = \{A \in GL(n,\mathbb{C}) : \bar{A}^T A = I\},$$

and its Lie algebra is

$$\mathfrak{u}(n) = \{A \in \mathfrak{g}l(n,\mathbb{C}) : \bar{A}^T + A = 0\},$$

the Lie algebra of skew-Hermitian matrices.

With these basic ideas it should be easy to calculate the Lie algebras of most other commonly encountered Lie groups.

If G and H are Lie groups and $h : G \to H$ is a Lie group homomorphism, we can define a map

$$h_* : \mathfrak{g} \to \mathfrak{h} \quad \text{by} \quad h_*(X) = \beta[(h_*)_e(X(e))].$$

One can check that this is a Lie algebra homomorphism; see Corollary 7.10 in Chapter 4 of [5]. This gives rise to a "covariant functor" from the category of Lie groups and Lie group homomorphisms to the category of Lie algebras and Lie algebra homomorphisms. A somewhat deeper theorem shows that for any Lie algebra $\mathfrak{g}$ there is a unique simply connected Lie group G with Lie algebra $\mathfrak{g}$. This correspondence between Lie groups and Lie algebras often reduces problems regarding Lie groups to Lie algebras, which are much simpler objects that can be studied via techniques of linear algebra.

1.12 Lie groups with biinvariant metrics

Definition. Suppose that G is a Lie group. A pseudo-Riemannian metric on G is *biinvariant* if the diffeomorphisms L_σ and R_σ are isometries for every $\sigma \in G$.

Example 1. For can define a Riemannian metric on $GL(n, \mathbb{R})$ by

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} dx^i_j \otimes dx^i_j. \tag{1.30}$$

This is just the Euclidean metric that $GL(n, \mathbb{R})$ inherits as an open subset of $\mathbb{E}^{n^2}$. The metric on $GL(n, \mathbb{R})$ is not biinvariant, but we claim that the metric it induces on the subgroup $O(n)$ is biinvariant.

To prove this, it suffices to show that the metric (1.30) is invariant under L_A and R_A, when $A \in O(n)$. If $A = (a^i_j) \in O(n)$ and $B = (b^i_j) \in GL(n, \mathbb{R})$, then

$$(x^i_j \circ L_A)(B) = x^i_j(AB) = \sum_{k=1}^{n} x^i_k(A) x^k_j(B) = \sum_{k=1}^{n} a^i_k x^k_j(B),$$

so that

$$L_A^*(x^i_j) = x^i_j \circ L_A = \sum_{k=1}^{n} a^i_k x^k_j.$$

It follows that

$$L_A^*(dx^i_j) = \sum_{k=1}^{n} a^i_k dx^k_j,$$

and hence

$$\begin{aligned} L_A^* \langle \cdot, \cdot \rangle &= \sum_{i,j=1}^{n} L_A^*(dx^i_j) \otimes L_A^*(dx^i_j) \\ &= \sum_{i,j,k,l=1}^{n} a^i_k dx^k_j \otimes a^i_l dx^l_j = \sum_{i,j,k,l=1}^{n} (a^i_k a^i_l) dx^k_j \otimes dx^l_j. \end{aligned}$$

Since $A^T A = I$, $\sum_{i=1}^n a_k^i a_l^i = \delta_{il}$, and hence

$$L_A^* \langle \cdot, \cdot \rangle = \sum_{j,k,l=1}^n \delta_{kl} dx_j^k \otimes dx_j^l = \langle \cdot, \cdot \rangle.$$

By a quite similar computation, one shows that

$$R_A^* \langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle, \quad \text{for } A \in O(n).$$

Hence the Riemannian metric defined by (1.30) is indeed invariant under right and left translations by elements of the compact group $O(n)$. Thus (1.30) induces a biinvariant Riemannian metric on $O(n)$, as claimed. Note that if we identify $T_I O(n)$ with the Lie algebra $\mathfrak{o}(n)$ of skew-symmetric matrices, this Riemannian metric is given by

$$\langle X, Y \rangle = \text{Trace}(X^T Y), \quad \text{for} \quad X, Y \in \mathfrak{o}(n).$$

Example 2. The unitary group $U(n)$ is an imbedded subgroup of $GL(2n, \mathbb{R})$ which lies inside $O(n)$, and hence if $\langle \cdot, \cdot \rangle_E$ is the Euclidean metric induced on $GL(2n, \mathbb{R})$,

$$L_A^* \langle \cdot, \cdot \rangle_E = \langle \cdot, \cdot \rangle_E = R_A^* \langle \cdot, \cdot \rangle_E, \quad \text{for } A \in U(n).$$

Thus the Euclidean metric on $GL(2n, \mathbb{R})$ induces a biinvariant Riemannian metric on $U(n)$. If we identify $T_I U(n)$ with the Lie algebra $\mathfrak{u}(n)$ of skew-Hermitian matrices, one can check that this Riemannian metric is given by

$$\langle X, Y \rangle = 2\text{Re}\left(\text{Trace}(X^T \bar{Y})\right), \quad \text{for} \quad X, Y \in \mathfrak{u}(n). \tag{1.31}$$

Remark. Once we have integration at our disposal, we will be able to prove that any compact Lie group has a biinvariant Riemannian metric. (See § 2.3.)

Proposition 1. *Suppose that G is a Lie group with a biinvariant pseudo-Riemannian metric $\langle \cdot, \cdot \rangle$. Then*

1. *geodesics passing through the identity $e \in G$ are just the one-parameter subgroups of G,*

2. *the Levi-Civita connection on TG is defined by*

$$\nabla_X Y = \frac{1}{2}[X, Y], \quad \textit{for } X, Y \in \mathfrak{g},$$

3. *the curvature tensor is given by*

$$\langle R(X, Y)W, Z \rangle = \frac{1}{4}\langle [X, Y], [Z, W] \rangle, \quad \textit{for } X, Y, Z, W \in \mathfrak{g}. \tag{1.32}$$

Before proving this, we need to some facts about the Lie bracket that are proven in [5]. Recall that if X is a vector field on a smooth manifold M with one-parameter group of local diffeomorphisms $\{\phi_t : t \in \mathbb{R}\}$ and Y is a second smooth vector field on M, then the Lie bracket $[X, Y]$ is determined by the formula

$$[X,Y](p) = -\frac{d}{dt}\left((\phi_t)_*(Y)(p)\right)\Big|_{t=0}. \tag{1.33}$$

(See the discussion surrounding Theorem 7.8 in Chapter 4 of [5].)

Definition. A vector field X on a pseudo-Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ is said to be *Killing* if its one-parameter group of local diffeomorphisms $\{\phi_t : t \in \mathbb{R}\}$ consists of isometries.

The formula (1.33) for the Lie bracket has the following consequence needed in the proof of the theorem:

Lemma 2. *If X is a Killing field, then*

$$\langle \nabla_Y X, Z\rangle + \langle Y, \nabla_Z X\rangle = 0, \quad \textit{for } Y, Z \in \mathcal{X}(M).$$

Proof: Note first that if X is fixed

$$\langle \nabla_Y X, Z\rangle(p) \quad \text{and} \quad \langle X, \nabla_Y Z\rangle(p)$$

depend only on $X(p)$ and $Y(p)$. Thus we can assume without loss of generality that $\langle Y, Z\rangle$ is constant. Then, since X is Killing, $\langle (\phi_t)_*(Y), (\phi_t)_*(Z)\rangle$ is constant, and

$$0 = \left\langle \frac{d}{dt}\left((\phi_t)_*(Y)\right)\Big|_{t=0}, Z \right\rangle + \left\langle Y, \frac{d}{dt}\left((\phi_t)_*(Z)\right)\Big|_{t=0} \right\rangle$$
$$= -\langle [X,Y], Z\rangle - \langle Y, [X,Z]\rangle.$$

On the other hand, since ∇ is the Levi-Civita connection,

$$0 = X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

Adding the last two equations yields the statement of the lemma.

Application. If X is a Killing field on the pseudo-Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ and $\gamma : (a,b) \to M$ is a geodesic, then since $\langle \nabla_Y X, Y\rangle = 0$,

$$\frac{d}{dt}\langle \gamma', X\rangle = \langle \nabla_{\gamma'}\gamma', X\rangle + \langle \gamma', \nabla_{\gamma'} X\rangle = 0.$$

Thus $\langle \gamma', X\rangle$ is constant along the geodesic. This often gives very useful constraints on geodesic flow.

We now turn to the proof of Theorem 1: First note that since the metric $\langle \cdot, \cdot \rangle$ is left invariant,

$$X, Y \in \mathfrak{g} \quad \Rightarrow \quad \langle X, Y\rangle \quad \text{is constant.}$$

Since the metric is right invariant, each $R_{\theta_X(t)}$ is an isometry, and hence X is a Killing field. Thus

$$\langle \nabla_Y X, Z \rangle + \langle \nabla_Z X, Y \rangle = 0, \quad \text{for } X, Y, Z \in \mathfrak{g}.$$

In particular,

$$\langle \nabla_X X, Y \rangle = -\langle \nabla_Y X, X \rangle = -\frac{1}{2} Y \langle X, X \rangle = 0.$$

Thus $\nabla_X X = 0$ for $X \in \mathfrak{g}$ and the integral curves of X must be geodesics.

Next note that

$$0 = \nabla_{X+Y}(X+Y) = \nabla_X X + \nabla_X Y + \nabla_Y X + \nabla_Y Y = \nabla_X Y + \nabla_Y X.$$

Averaging the equations

$$\nabla_X Y + \nabla_Y X = 0, \qquad \nabla_X Y - \nabla_Y X = [X, Y]$$

yields the second assertion of the proposition.

Finally, if $X, Y, Z \in \mathfrak{g}$, use of the Jacobi identity yields

$$\begin{aligned} R(X,Y)Z &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ &= \frac{1}{4}[X,[Y,Z]] - \frac{1}{4}[Y,[X,Z]] - \frac{1}{2}[[X,Y],Z] = -\frac{1}{4}[[X,Y],Z]. \end{aligned}$$

On the other hand, if $X, Y, Z \in \mathfrak{g}$,

$$0 = 2X\langle Y, Z \rangle = 2\langle \nabla_X Y, Z \rangle + 2\langle Y, \nabla_X Z \rangle = \langle [X,Y], Z \rangle + \langle Y, [X,Z] \rangle.$$

Thus we conclude that

$$\langle R(X,Y)W, Z \rangle = -\frac{1}{4}\langle [[X,Y],W], Z \rangle = \frac{1}{4}\langle [X,Y],[Z,W] \rangle,$$

finishing the proof of the third assertion.

Remark. If G is a Lie group with a biinvariant pseudo-Riemannian metric, the map

$$\nu : G \to G \quad \text{defined by} \quad \nu(\sigma) = \sigma^{-1},$$

is an isometry. Indeed, it is immediate that $(\nu_*)_e = -\mathrm{id}$ is an isometry, and the identity

$$\nu = R_{\sigma^{-1}} \circ \nu \circ L_{\sigma^{-1}}$$

shows that $(\nu_*)_\sigma$ is an isometry for each $\sigma \in G$. Thus ν is an isometry of G which reverses geodesics through the identity e. More generally, the map $I_\sigma = L_{\sigma^{-1}} \circ \nu \circ L_\sigma$ is an isometry which reverses geodesics through σ

A *Riemannian symmetric space* is a Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ such that for each $p \in M$ there is an isometry $I_p : M \to M$ which reverses geodesics through p. Examples include not just the Lie groups with biinvariant Riemannian metrics and the spaces of constant curvature, but many other important examples, including the Grassmann manifolds to be described in the next section.

1.13 Grassmann manifolds

If G is a compact Lie group with biinvariant Riemannian metric $\langle \cdot, \cdot \rangle$, certain submanifolds $M \subseteq G$ inherit Riemannian metrics for which geodesics and curvature can be easily computed. These include the complex projective space with its "Fubini-Study" metric, a space which occurs in algebraic geometry and other contexts.

To explain these examples, we assume as known the basic theory of homogeneous spaces. As described in Chapter 4, §9 of [5], if G is a Lie group and H is a compact subgroup, the *homogeneous space* of left cosets G/H is a smooth manifold and the projection $\pi : G \to G/H$ is a smooth submersion. Moreover, the map $G \times G/H \to G/H$, defined by $(\sigma, \tau H) \to \sigma\tau H$, is smooth.

Suppose now that H is a compact subgroup of G and that $s : G \to G$ is a group homomorphism such that:

1. $s^2 = \mathrm{id}$, and
2. $H = \{\sigma \in G : s(\sigma) = \sigma\}$.

Given such a triple (G, H, s), the group homomorphism s induces a Lie algebra homomorphism $s_* : \mathfrak{g} \to \mathfrak{g}$ such that $s_*^2 = \mathrm{id}$. We let

$$\mathfrak{h} = \{X \in \mathfrak{g} : s_*(X) = X\}, \quad \mathfrak{p} = \{X \in \mathfrak{g} : s_*(X) = -X\}.$$

Moreover, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is a direct sum decomposition, and the fact that s_* is a Lie algebra homomorphism implies that

$$[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}, \quad [\mathfrak{h}, \mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{h}.$$

Finally, note that $\mathfrak{h}$ is the Lie algebra of H and hence is isomorphic to the tangent space to H at the identity e, while $\mathfrak{p}$ is the tangent space to G/K at eK.

Under these conditions we can define a map

$$f : G/H \to G \quad \text{defined by} \quad f(\sigma H) = \sigma s(\sigma^{-1}).$$

Indeed, if $h \in H$ then $f(\sigma h) = \sigma h s(h^{-1}\sigma^{-1}) = \sigma s(\sigma^{-1})$, so f is a well-defined map on the homogeneous space G/H. Moreover,

$$\sigma s(\sigma^{-1}) = \tau s(\tau^{-1}) \quad \Leftrightarrow \quad \tau^{-1}\sigma = s(\tau^{-1}\sigma) \quad \Leftrightarrow \quad \tau^{-1}\sigma \in H,$$

so f is injective. Finally, one checks that

$$X \in \mathfrak{p} \quad \Rightarrow \quad t \mapsto s(e^{-tX})$$

is a one-parameter group and checking the derivative at $t = 0$ shows that $s(e^{-tX}) = e^{tX}$ and hence $f(e^{tX}) = e^{2tX}$. Moreover,

$$f(\sigma e^{tX}) = \sigma e^{tX} s(e^{-tX}) s(\sigma^{-1}) = L_\sigma \circ R_{\sigma^{-1}}(e^{2tX}). \tag{1.34}$$

From these facts it follows that f is a one-to-one immersion from G/H into G, and hence an imbedding which exhibits G/H as an imbedded submanifold of G.

We can therefore define an induced Riemannian metric $\langle\cdot,\cdot\rangle_{G/H} = f^*\langle\cdot,\cdot\rangle_G$ such that

$$\langle X, Y\rangle_{G/H} = 4\langle X, Y\rangle_G, \quad \text{for} \quad X, Y \in \mathfrak{p}. \tag{1.35}$$

Since $L_\sigma \circ R_{\sigma^{-1}}$ is an isometry, the geodesics in the induced submanifold metric on G/H are just the curves $t \mapsto f(\sigma e^{tX})$. It follows that G acts as a group of isometries on G/H when G/H is given the induced metric.

We mention two examples:

Example 1. Suppose $G = O(n)$ and s is conjugation with the element

$$I_{p,q} = \begin{pmatrix} -I_{p\times p} & 0 \\ 0 & I_{q\times q} \end{pmatrix}, \quad \text{where} \quad p + q = n.$$

Thus

$$s(A) = I_{p,q} A I_{p,q}, \quad \text{for } A \in O(n),$$

and it is easily verified that s preserves the biinvariant metric and is a group homomorphism. In this case $H = O(p) \times O(q)$ and the quotient $O(n)/O(p) \times O(q)$ is the *Grassmann manifold* of real p-planes in n-space.

Example 2. Suppose $G = U(n)$ and s is conjugation with the element

$$I_{p,q} = \begin{pmatrix} -I_{p\times p} & 0 \\ 0 & I_{q\times q} \end{pmatrix}, \quad \text{where} \quad p + q = n.$$

In this case $H = U(p) \times U(q)$ and the quotient $U(n)/U(p) \times U(q)$ is the *Grassmann manifold* of complex p-planes in n-space. The special case $U(n)/U(1) \times U(n-1)$ of complex one-dimensional subspaces of $U(n)$ is also known as complex projective space $\mathbb{C}P^{n-1}$.

Theorem. *Given a triple (G, H, s) satisfying the above conditions, the curvature of $f(G//H$ is given by the formula*

$$\langle R(X,Y)W, Z\rangle = 4\langle [X,Y], [Z,W]\rangle, \quad \textit{for } X, Y, Z, W \in T_{eK}(G/H) \cong \mathfrak{p}.$$

Sketch of proof: The curvature formula follows from the Gauss equation for a submanifold M of a Riemannian manifold $(N, \langle\cdot,\cdot\rangle)$, when M is given the induced submanifold metric. To prove such an equation one follows the discussion already given in §1.8, except that we replace the ambient Euclidean space $\mathbb{E}^N$ with a general Riemannian manifold $(N, \langle\cdot,\cdot\rangle)$.

Thus if $p \in M \subseteq N$ and $v \in T_pN$, we let

$$v = v^\top + v^\perp, \quad \text{where} \quad v^\top \in T_pM \quad \text{and} \quad v^\perp \perp T_pM,$$

$(\cdot)^\top$ and $(\cdot)^\perp$ being the orthogonal projection into the tangent space and normal space. The Levi-Civita connection ∇^M on M is then defined by the formula

$$(\nabla^M_X Y)(p) = (\nabla^N_X Y(p))^\top,$$

where ∇^N is the Levi-Civita connection on N. If we let $\mathcal{X}^\perp(M)$ denote the vector fields in N which are defined at points of M and are perpendicular to M, then we can define the second fundamental form

$$\alpha : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}^\perp(M) \quad \text{by} \quad \alpha(X,Y) = (\nabla^N_X Y(p))^\perp.$$

As before, it satisfied the identities:

$$\alpha(fX, Y) = \alpha(X, fY) = f\alpha(X,Y), \quad \alpha(X,Y) = \alpha(Y,X).$$

If $\gamma : (a,b) \to M \subseteq \mathbb{E}^N$ is a unit speed curve, we call $(\nabla^N_{\gamma'}\gamma')$ the geodesic curvature of γ in N, while

$$\left(\nabla^N_{\gamma'}\gamma'\right)^\top = \nabla^M_{\gamma'}\gamma' = (\text{geodesic curvature of } \gamma \text{ in } M),$$

$$\left(\nabla^N_{\gamma'}\gamma'\right)^\perp = \alpha(\gamma, \gamma') = (\text{normal curvature of } \gamma).$$

Under these circumstances, one can show that the curvature tensor R^M of $M \subseteq \mathbb{E}^N$ is given by the *Gauss equation*

$$\langle R^M(X,Y)W, Z\rangle = \langle R^N(X,Y)W,Z\rangle + \langle \alpha(X,Z), \alpha(Y,W)\rangle - \langle \alpha(X,W), \alpha(Y,Z)\rangle, \tag{1.36}$$

whenever X, Y, Z and W are elements of $\mathcal{X}(M)$. The proof of (1.36) is identical to the proof of the Gauss equation we gave before in §1.8.

In our application, since geodesics in $f(G/H)$ are geodesics in the ambient manifold G, $\alpha = 0$ and the theorem follows directly from (1.36), together with (1.32) and the fact that the differential of the map $f : G/H \to G$ multiplies every element of $\mathfrak{p} = T_{eK}(G/H)$ by two.

Example. We consider the special case in which $G = U(n)$ and s is conjugation with

$$I_{1,n-1} = \begin{pmatrix} -1 & 0 \\ 0 & I_{(n-1)\times(n-1)} \end{pmatrix},$$

so that the fixed point set of the automorphism s is $H = U(1) \times U(n-1)$ and $G/H = \mathbb{C}P^{n-1}$.

Recall that the Lie algebra $\mathfrak{u}(n)$ divides into a direct sum $\mathfrak{u}(n) = \mathfrak{h} \oplus \mathfrak{p}$, where

$$\mathfrak{h} = \{X \in \mathfrak{g} : s_*(X) = X\}, \quad \mathfrak{p} = \{X \in \mathfrak{g} : s_*(X) = -X\},$$

where $\mathfrak{h}$ is the Lie algebra of $U(1) \times U(n-1)$. We consider two elements

$$X = \begin{pmatrix} 0 & -\bar{\xi}_2 & \cdots & -\bar{\xi}_n \\ \xi_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdot \\ \xi_n & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 0 & -\bar{\eta}_2 & \cdots & -\bar{\eta}_n \\ \eta_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdot \\ \eta_n & 0 & \cdots & 0 \end{pmatrix}$$

of $\mathfrak{p}$, and determine their Lie bracket $[X,Y] \in \mathfrak{h}$.

It is easier to do this by carrying out the multiplication in matrix terms. To simplify notation, we write

$$X = \begin{pmatrix} 0 & -\bar{\xi}^T \\ \xi & 0 \end{pmatrix} \quad \text{and} \qquad Y = \begin{pmatrix} 0 & -\bar{\eta}^T \\ \eta & 0 \end{pmatrix}, \tag{1.37}$$

so that

$$XY = \begin{pmatrix} -\bar{\xi}^T\eta & 0 \\ 0 & -\xi\bar{\eta}^T \end{pmatrix}, \quad YX = \begin{pmatrix} -\bar{\eta}^T\xi & 0 \\ 0 & -\eta\bar{\xi}^T \end{pmatrix}$$

and

$$[X,Y] = \begin{pmatrix} -\bar{\xi}^T\eta + \bar{\eta}^T\xi & 0 \\ 0 & -\xi\bar{\eta}^T + \eta\bar{\xi}^T \end{pmatrix}.$$

We next use the formula for the curvature of G/H to show that the sectional curvatures $K(\sigma)$ for $\mathbb{C}P^{n-1}$ satisfy the inequalities $a^2 \leq K(\sigma) \leq 4a^2$ for some $a^2 > 0$. As inner product on $T_I U(n)$, we use

$$\langle A, B\rangle = \frac{1}{2}\mathrm{Re}\left(\mathrm{Trace}(A^T\bar{B})\right), \quad \text{for} \quad A, B \in \mathfrak{u}(n).$$

This differs by a factor of four from the Riemannian metric induced by the natural imbedding into $\mathbb{E}^{(2n)^2}$, but with the rescaled metric

$$\langle X, Y\rangle = \mathrm{Re}(\xi^T\bar{\eta}),$$

when X and Y are given by (1.37). To simplify the calculations, assume that

$$\langle X, X\rangle = \langle Y, Y\rangle = 1, \quad \text{and} \quad \langle X, Y\rangle = 0.$$

Then

$$|\xi|^2 = |\eta|^2 = 1 \quad \text{and} \quad \xi^T\bar{\eta} = -\eta^T\bar{\xi},$$

the latter since $\xi^T\bar{\eta}$ is purely imaginary. Then

$$\begin{aligned}
&\langle [X,Y],[X,Y]\rangle \\
&\quad = \frac{1}{2}\mathrm{Trace}\begin{pmatrix} -\eta^T\bar{\xi} + \xi^T\bar{\eta} & 0 \\ 0 & -\bar{\eta}\xi^T + \bar{\xi}\eta^T \end{pmatrix}\begin{pmatrix} -\xi^T\bar{\eta} + \eta^T\bar{\xi} & 0 \\ 0 & -\bar{\xi}\eta^T + \bar{\eta}\xi^T \end{pmatrix} \\
&\qquad\quad = \frac{1}{2}\mathrm{Trace}\begin{pmatrix} 4\left|\mathrm{Im}(\xi^T\bar{\eta})\right|^2 & 0 \\ 0 & (-\bar{\eta}\xi^T + \bar{\xi}\eta^T)(-\bar{\xi}\eta^T + \bar{\eta}\xi^T) \end{pmatrix} \\
&\qquad\qquad = 2\left|\mathrm{Im}(\xi^T\bar{\eta})\right|^2 + |\xi|^|\eta|^2 + \left|\mathrm{Im}(\xi^T\bar{\eta})\right|^2 \\
&\qquad\qquad\qquad = |\xi|^|\eta|^2 + 3\left|\mathrm{Im}(\xi^T\bar{\eta})\right|^2.
\end{aligned}$$

The last expression ranges between 1 and 4, and it follows from the Cauchy-Schwarz inequality that it achieves its maximum when $\eta = i\xi$. Thus if σ is the two-plane spanned by X and Y,

$$K(\sigma) = \frac{4\langle [X,Y],[X,Y]\rangle}{\langle X,X\rangle\langle Y,Y\rangle - \langle X,Y\rangle^2} = 4\left[|\xi|^|\eta|^2 + 3\left|\mathrm{Im}(\xi^T\bar{\eta})\right|^2\right]$$

lies in the interval $[4, 16]$, achieving both extreme values when $n - 1 \geq 2$.

The Riemannian metric we have defined on $G/H = \mathbb{C}P^{n-1}$ is called the *Fubini-Study metric*. It occurs frequently in algebraic geometry.

1.14 The exponential map

Our next goal is to develop a system of local coordinates centered at a given point p in a Riemannian manifold which are as Euclidean as possible.

Proposition 1. *Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a pseudo-Riemannian manifold and $p \in M$. Then there is an open neighborhood V of 0 in T_pM such that if $v \in T_pM$ the unique geodesic γ_v in M which satisfies the initial conditions $\gamma_v(0) = p$ and $\gamma_v'(0) = v$ is defined on the interval $[0, 1]$.*

Proof: According to ODE theory applied to the second-order system of differential equations

$$\frac{d^2x^i}{dt^2} + \sum_{j,k=1}^{n} \Gamma^i_{jk} \frac{dx^j}{dt} dx^k dt = 0,$$

there is a neighborhood W of 0 in T_pM and an $\epsilon > 0$ such that the geodesic γ_w is defined on $[0, \epsilon]$ for all $w \in W$. Let $V = \epsilon W$. Then if $v \in V$, $v = \epsilon w$ for some $w \in W$, and since $\gamma_v(t) = \gamma_w(\epsilon t)$, γ_v is defined on $[0, 1]$, proving the proposition.

Definition. Define the *exponential map*

$$\exp_p V \to M \quad \text{by} \quad \exp_p(v) = \gamma_v(1).$$

Remark. Note that if $G = O(n)$ with the standard biinvariant metric $\langle \cdot, \cdot \rangle$ which we constructed in §1.12,

$$\exp_I A = e^{tA}, \quad \text{for } A \in T_I O(n).$$

This explains the origin of the term "exponential map."

Note that if $v \in V$, $t \mapsto \exp_p(tv)$ is a geodesic (because $\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t)$), and hence $\exp_p$ takes straight line segments through the origin in T_pM to geodesic segments through p in M.

Proposition 2. *There is an open neighborhood $\tilde{U}$ of 0 in T_pM which $\exp_p$ maps diffeomorphically onto an open neighborhood U of p in M.*

Proof: By the inverse function theorem, it will suffice to show that

$$((\exp_p)_*)_0 : T_0(T_pM) \longrightarrow T_pM$$

is an isomorphism. We identify $T_0(T_pM)$ with T_pM. If $v \in T_pM$, define

$$\lambda_v : \mathbb{R} \to T_pM \quad \text{by} \quad \lambda_v(t) = tv.$$

Then $\lambda_v'(0) = v$ and

$$
\begin{aligned}
((\exp_p)_*)_0(v) = ((\exp_p)_*)_0(\lambda_v'(0)) &= (\exp_p) \circ \lambda_v)'(0) \\
&= \frac{d}{dt}(\exp_p(tv))\bigg|_{t=0} = \frac{d}{dt}(\gamma_v(t))\bigg|_{t=0} = v,
\end{aligned}
$$

so $((\exp_p)_*)_0$ is indeed an isomorphism.

It will sometimes be useful to have a stronger version of the above proposition, proven by the same method, but making use of the map

$$\exp : (\text{neighborhood of 0-section in } TM) \longrightarrow M \times M,$$

defined by

$$\exp(v) = (p, \exp_p(v)), \quad \text{for } v \in T_p M.$$

Proposition 3. *Given a point $p_0 \in M$ there is an open neighborhood $\tilde{W}$ of the zero vector 0 of $T_{p_0}M$ which exp maps diffeomorphically onto an open neighborhood W of (p_0, p_0) in $M \times M$.*

Proof: If 0 denotes the zero vector in $T_{p_0}M$, it suffices to show that

$$(\exp_*)_0 : T_0(TM) \longrightarrow T_{(p_0,p_0)}(M \times M)$$

is an isomorphism. Since both vector spaces have the same dimension it suffices to show that $(\exp_*)_0$ is an epimorphism. Let

$$\pi_1 : M \times M \to M, \qquad \pi_2 : M \times M \to M$$

denote the projections on the first and second factors, respectively. Then $\pi_i \circ \exp : TM \to M$ is the bundle projection $TM \to M$ and hence $((\pi_1 \circ \exp)_*)_0$ is an epimorphism. On the other hand, the composition

$$T_{p_0}M \subseteq TM \xrightarrow{\exp} M \times M \xrightarrow{\pi_2} M$$

is just $\exp_{p_0}$ and hence $((\pi_2 \circ \exp)_*)_0$ is an epimorphism by the previous proposition. Hence $(\exp_*)_0$ is indeed an epimorphism as claimed.

Corollary 4. *Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold and $p_0 \in M$. Then there is an open neighborhood U of p_0 and an $\epsilon > 0$ such that $\exp_p$ maps*

$$\{v \in T_p M : \langle v, v \rangle < \epsilon^2\}$$

diffeomorphically onto an open subset of M for all $p \in U$.

If $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold and $p \in M$. If we choose a basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ for $T_p M$, orthonormal with respect to the inner product $\langle \cdots, \cdot \rangle_p$, we can define "Euclidean" coordinates $(\dot{x}^1, \ldots, \dot{x}^n)$ on $T_p M$ by

$$\dot{x}^i(v) = a^i \quad \Leftrightarrow \quad v = \sum_{i+1}^{n} a^i \mathbf{e}_i.$$

If U is an open neighborhood of $p \in M$ such that $\exp_p$ maps an open neighborhood $\tilde{U}$ of $0 \in T_pM$ diffeomorphically onto U, we can define coordinates

$$(x^1, \ldots, x^n) : U \to \mathbb{R}^n \quad \text{by} \quad x^i \circ \exp_p = \dot{x}^i.$$

The coordinate $(x^1, \ldots, x^n)$ are called *Riemannian normal coordinates* on M centered at p, or simply *normal coordinates*. These normal coordinates are the coordinates which are as Euclidean as possible near p.

Suppose that in terms of the normal coordinates

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j.$$

It is interesting to determine the Taylor series expansion of the g_{ij}'s in normal coordinates. Of course, we have $g_{ij}(p) = \delta_{ij}$.

To evaluate the first order derivatives, we note that whenever $a^1, \ldots, a^n$ are constants, the curve γ defined by

$$x^i \circ \gamma(t) = a^i t$$

is a geodesic in M by definition of the exponential map. Thus the functions $x^i = x^i \circ \gamma$ must satisfy the geodesic equation

$$\ddot{x}^k + \sum_{i,j=1}^{n} \Gamma^k_{ij} \dot{x}^i \dot{x}^j = 0.$$

Substitution into this equation yields

$$\sum_{i,j=1}^{n} \Gamma^k_{ij}(p) a^i a^j = 0.$$

Since this holds for all choices of the constants $(a^1, \ldots, a^n)$ we conclude that $\Gamma^k_{ij}(p) = 0$. It then follows from (1.15) that

$$\frac{\partial g_{ij}}{\partial x^k}(p) = 0.$$

Later we will see that the Taylor series for the Riemannian metric in normal coordinates centered at p is given by

$$g_{ij} = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^{n} R_{ikjl}(p) x^k x^l + (\text{higher order terms}).$$

This formula gives a very explicit formula for how much the Riemannian metric differs from the Euclidean metric near a given point p. Before proving this, we will need the so-called Gauss lemma.

1.15 The Gauss Lemma

We now suppose that $(x^1, \ldots, x^n)$ are normal coordinates centered at a point p in a Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$, and defined on an open neighborhood U of p. We can then define a radial function

$$r : U \to \mathbb{R} \quad \text{by} \quad r = \sqrt{(x^1)^2 + \cdots + (x^n)^2},$$

and a radial vector field S on $U - \{p\}$ by

$$S = \sum_{i=1}^{n} \frac{x^i}{r} \frac{\partial}{\partial x^i}.$$

For $1 \leq i, j \leq n$, let E_{ij} be the rotation vector field on U defined by

$$E_{ij} = x^i \frac{\partial}{\partial x^j} - x^j \frac{\partial}{\partial x^i}.$$

Lemma 1. $[E_{ij}, S] = 0$.

Proof: This can be verified by direct calculation. For a more conceptual argument, one can note that the one-parameter group of local diffeomorphisms $\{\phi_t : t \in \mathbb{R}\}$ on U induced by E_{ij} consists of rotations in terms of the normal coordinates, so $(\phi_t)_*(R) = R$, so

$$[E_{ij}, R] = -\frac{d}{dt} \left((\phi_t)_*(R)\right)\Big|_{t=0} = 0.$$

Lemma 2. *If ∇ is the Levi-Civita connection on M, then $\nabla_S S = 0$.*

Proof: If $(a^q, \ldots, a^n)$ are real numbers such that $\sum (a^i)^2 = 1$, then the curve γ defined by

$$x^i(\gamma(t) = a^i t$$

is an integral curve for S. On the other hand,

$$\gamma(t) = \exp_p \left(\sum_{i=1}^{n} a^i t \, \frac{\partial}{\partial x^i}\Big|_p \right),$$

and hence γ is a geodesic. We conclude that all integral curves for S are geodesics and hence $\nabla_S S = 0$.

Lemma 3. $\langle S, S \rangle \equiv 1$.

Proof: If γ is as in the preceding lemma,

$$\frac{d}{dt} \langle \gamma'(t), \gamma'(t) \rangle = 2 \langle \nabla_{\gamma'(t)} \gamma'(t), \gamma'(t) \rangle = 0,$$

so $\gamma'(t)$ must have constant length. But

$$\langle \gamma'(0), \gamma'(0) \rangle = \sum_{i=1}^{n} (a^i)^2 = 1,$$

so we conclude that $\langle S, S \rangle \equiv 1$.

Lemma 4. $\langle S, E_{ij} \rangle \equiv 0$.

Proof: We calculate the derivative of $\langle S, E_{ij} \rangle$ in the radial direction:

$$\begin{aligned} S\langle S, E_{ij} \rangle = \langle \nabla_S S, E_{ij} \rangle + \langle S, \nabla_S E_{ij} \rangle = \langle S, \nabla_S E_{ij} \rangle \\ = \langle S, \nabla_{E_{ij}} S \rangle = \frac{1}{2} E_{ij} \langle S, S \rangle = 0. \end{aligned}$$

Thus $\langle S, E_{ij} \rangle$ is constant along the geodesic rays emanating from p. let $\|X\| = \sqrt{\langle X, X \rangle}$. Then as $(x^1, \dots, x^n) \to (0, \dots 0)$,

$$|\langle S, E_{ij} \rangle| \leq \|S\| \|E_{ij}\| = \|E_{ij}\| \to 0.$$

If follows that the constant $\langle S, E_{ij} \rangle$ must be zero.

Before proving the next lemma, we observe that

$$S(r) = 1, \qquad E_{ij}(r) = 0.$$

These fact can be verified by direct computation.

Lemma 5. *$dr = \langle S, \cdot \rangle$; in other words, $dr(X) = \langle S, X \rangle$, whenever X is a smooth vector field on $U - \{p\}$.*

Proof: It clearly suffices to prove this when either $X = S$ or $X = E_{ij}$. In the first case,

$$dr(S) = S(r) = 1 = \langle S, S \rangle,$$

while in the second,

$$dr(E_{ij}) = E_{ij}(r) = 0 = \langle S, E_{ij} \rangle.$$

Remark. It is Lemma 4 which is often called the *Gauss Lemma*.

1.16 Curvature in normal coordinates

Our next goal is to prove the following theorem, which explains how the curvature of a Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ measures deviation from the Euclidean metric.

Taylor Series Theorem. *The Taylor series for the Riemannian metric (g_{ij}) normal coordinates centered at a point p is given by*

$$g_{ij} = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^{n} R_{ikjl}(p) x^k x^l + (\textit{higher order terms}).$$

To prove this, we make use of "constant extensions" of vectors in T_pM, relative to the normal coordinates $(x^1, \ldots, x^n)$. Suppose that $w \in T_pM$ and

$$w = \sum_{i=1}^{n} a^i \left. \frac{\partial}{\partial x^i} \right|_p .$$

Then the *constant extension* of w is the vector field

$$W = \sum_{i=1}^{n} a^i \frac{\partial}{\partial x^i}.$$

Since there is a genuine constant vector field in T_pM which is $\exp_p$-related to W, W depends only on w, not on the choice of normal coordinates.

We define a quadrilinear map

$$G : T_pM \times T_pM \times T_pM \times T_pM \longrightarrow \mathbb{R}$$

as follows:

$$G(x, y, z, w) = XY\langle Z, W\rangle(p),$$

where X, Y, Z and W and the constant extensions of x, y, z and w. Thus the components of G will be the second order derivatives of the metric tensor.

Lemma. *The quadralinear form G satisfies the following symmetries:*

1. $G(x, y, z, w) = G(y, x, z, w)$,
2. $G(x, y, z, w) = G(x, y, w, z)$,
3. $G(x, x, x, x) = 0$,
4. $G(x, x, x, y) = 0$,
5. $G(x, y, z, w) = G(z, w, x, y)$, *and*
6. $G(x, y, z, w) + G(x, z, w, y) + G(x, w, y, z) = 0$.

Proof: The second of these identities is immediate and the first follows from equality of mixed partials. The other identities require more work.

For the identity $G(w, w, w, w) = 0$, we let $W = \sum a^i(\partial/\partial x^i)$; then the curve γ defined by

$$x^i(\gamma(t)) = a^i t$$

is an integral curve for W such that $\gamma(0) = p$. It is also a constant speed geodesic and hence

$$WW\langle W, W\rangle(p) = 0.$$

We next check that $G(w, w, w, z) = 0$. It clearly suffices to prove this when z is unit length and perpendicular to a unit length w. We can choose our normal coordinates so that

$$W = \frac{\partial}{\partial x^1}, \qquad Z = \frac{\partial}{\partial x^2}.$$

We consider the curve γ in M defined by

$$x^1 \circ \gamma(t) = t, \qquad x^i \circ \gamma(t) = 0, \quad \text{for } i > 1.$$

Along γ we have $W = S$ and $Z = (1/x^1)E_{12}$, so it follows from Lemma 12.4 that $\langle W, Z\rangle \equiv 0$ along γ, and hence

$$WW\langle W, Z\rangle(p) = 0.$$

it follows from the first two symmetries that whenever $u, v \in T_pM$ and $t \in \mathbb{R}$,

$$\begin{aligned}0 = G(&u + tv, u + tv, u + tv, u - tv)\\ &= t^3(\text{something}) + t^2[G(v, v, u, u) - G(u, u, v, v)] + t(\text{something}).\end{aligned}$$

Since this identity must hold for all t, the coefficient of t^2 must be zero, so

$$G(u, u, v, v) = G(v, v, u, u),$$

which yields the fifth symmetry.

To obtain the final identity, we let

$$v_1, v_2, v_3, v_4 \in T_pM \quad \text{and} \quad t_1, t_2, t_3, t_4 \in \mathbb{R},$$

and note that

$$G\left(\sum t_i v_i, \sum t_j v_j, \sum t_k v_k, \sum t_l v_l\right) = 0.$$

The coefficient of $t_1t_2t_3t_4$ must vanish, and hence

$$\sum_{\sigma \in S_4} G\left(v_{\sigma(1)}, v_{\sigma(2)}, v_{\sigma(3)}, v_{\sigma(4)}\right) = 0.$$

This, together with the earlier symmetries, yields the last symmetry.

Now we let

$$g_{ij,kl} = G\left(\left.\frac{\partial}{\partial x^k}\right|_p, \left.\frac{\partial}{\partial x^l}\right|_p, \left.\frac{\partial}{\partial x^i}\right|_p, \left.\frac{\partial}{\partial x^j}\right|_p\right).$$

Lemma . $R_{iljk}(p) = g_{ij,lk} - g_{ik,lj}$.

Proof: Since the Christoffel symbols Γ^k_{ij} vanish at p, it follows that

$$\frac{\partial}{\partial x^i}(\Gamma^l_{jk})(p) = \frac{1}{2}\frac{\partial}{\partial x^i}\left[\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l}\right](p),$$

$$\frac{\partial}{\partial x^j}(\Gamma^l_{ik})(p) = \frac{1}{2}\frac{\partial}{\partial x^j}\left[\frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l}\right](p),$$

and hence we conclude from Proposition 2 from §1.8 that

$$\begin{aligned} R_{ijlk}(p) = \frac{1}{2}&\left[\frac{\partial^2 g_{jl}}{\partial x^i \partial x^k} - \frac{\partial^2 g_{jk}}{\partial x^i \partial x^l} + \frac{\partial^2 g_{ik}}{\partial x^j \partial x^l} - \frac{\partial^2 g_{il}}{\partial x^j \partial x^k}\right](p) \\ &= \frac{1}{2}\left[g_{jl,ik} + g_{lk,jl} - g_{jk,il} - g_{il,jk}\right] = g_{ik,jl} - g_{il,jk}, \end{aligned}$$

the last step following from the third symmetry of G.

From the last two lemmas, we now conclude that

$$R_{ikjl}(p) + R_{iljk}(p) = g_{il,jk} - g_{ij,lk} + g_{ik,lj} - g_{ij,lk} = -3g_{ij,kl}.$$

We therefore conclude that

$$\frac{\partial^2 g_{ij}}{\partial x^k \partial x^l}(p) = -\frac{1}{3}[R_{ikjl}(p) + R_{iljk}(p)].$$

Substitution into the Taylor expansion

$$g_{ij} = \delta_{ij} + \frac{1}{2}\sum_{k,l=1}^{n} \frac{\partial^2 g_{ij}}{\partial x^k \partial x^l}(p) x^k x^l + (\text{higher order terms})$$

now yields the Taylor Series Theorem.

1.17 Riemannian manifolds as metric spaces

We can use the normal coordinates constructed in the previous sections to establish the following important result:

Local Minimization Theorem. *Suppose that $(M^n, \langle\cdot,\cdot\rangle)$ is a Riemannian manifold and that $\tilde{U}$ is an open ball of radius $\epsilon > 0$ centered at $0 \in T_pM$ which $\exp_p$ maps diffeomorphically onto an open neighborhood U of p in M. Suppose that $v \in \tilde{U}$ and that $\gamma : [0,1] \to M$ is the geodesic defined by $\gamma(t) = \exp_p(tv)$. Let $q = \exp_p(v)$. If $\lambda : [0,1] \to M$ is any smooth curve with $\lambda(0) = p$ and $\lambda(1) = q$, then*

1. *$L(\lambda) \geq L(\gamma)$, with equality holding only if λ is a reparametrization of γ, and*

2. *$J(\lambda) \geq J(\gamma)$, with equality holding only if $\lambda = \gamma$.*

To prove the first of these assertions, we use normal coordinates $(x^1, \dots, x^n)$ defined on U. Note that $L(\gamma) = r(q)$. Suppose that $\lambda : [0,1] \to M$ is any smooth curve with $\lambda(0) = p$ and $\lambda(1) = q$.

Case I. Suppose that λ does not leave U. Then

$$\begin{aligned} L(\lambda) = \int_0^1 \sqrt{\langle \lambda'(t), \lambda'(t) \rangle} dt = \int_0^1 \|\lambda'(t)\| dt \geq \int_0^1 \langle \lambda'(t), R(\lambda(t) \rangle dt \\ \geq \int_0^1 dr(\lambda'(t)) dt = (r \circ \lambda)(1) - (r \circ \lambda)(0) = L(\gamma). \end{aligned}$$

Moreover, equality holds only if $\lambda'(t)$ is a nonnegative multiple of $R(\lambda(t))$ which holds only if λ is a reparametrization of γ.

Case II. Suppose that λ leaves U at some time $t_0 \in (0,1)$. Then

$$\begin{aligned} L(\lambda) = \int_0^1 \sqrt{\langle \lambda'(t), \lambda'(t) \rangle} dt > \int_0^{t_0} \|\lambda'(t)\| dt \geq \int_0^1 \langle \lambda'(t), R(\lambda(t) \rangle dt \\ \geq \int_0^{t_0} dr(\lambda'(t)) dt = (r \circ \lambda)(t_0) - (r \circ \lambda)(0) = \epsilon > L(\gamma). \end{aligned}$$

The second assertion is proven in a similar fashion.

If $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold, we can define a distance function

$$d : M \times M \longrightarrow \mathbb{R}$$

by setting

$$\begin{aligned} d(p,q) = \inf\{\ L(\gamma) \text{ such that } \gamma : [0,1] \to M \text{ is a smooth path} \\ \text{with } \gamma(0) = p \text{ and } \gamma(1) = q\ \}. \end{aligned}$$

Then the previous theorem shos that $d(p,q) = 0$ implies that $p = q$. Hence

1. $d(p,q) \geq 0$, with equality holding if and only if $p = q$,
2. $d(p,q) = d(q,p)$, and
3. $d(p,r) \leq d(p,q) + d(q,r)$.

Thus (M,d) is a metric space. It is relatively straightforward to show that the metric topology on M agrees with the usual topology of M.

Definition. If p and q are points in a Riemannian manifold M, a *minimal geodesic* from p to q is a geodesic $\gamma : [a,b] \to M$ such that

$$\gamma(a) = p, \quad \gamma(b) = q \quad \text{and} \quad L(\gamma) = d(p,q).$$

An open set $U \subset M$ is said to be *geodesically convex* if whenever p and q are elements of U, there is a unique minimal geodesic from p to q and moreover, that minimal geodesic lies entirely within U.

Geodesic Convexity Theorem. *Suppose that $(M^n, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold. Then M has an open cover by geodesically convex open sets.*

A proof could be constructed based upon the preceding arguments, but we omit the details. (One proof is outlined in Problem 6.4 from [20].)

1.18 Completeness

We return now to the variational problem with which we started this chapter. Given two points p and q in a Riemannian manifold M, does there exist a minimal geodesic from p to q? For this variational problem to have a solution we need an hypothesis on the Riemannian metric.

Definition. A pseudo-Riemannian manifold $(M, \langle \cdot, \cdot)$ is said to be *geodesically complete* if geodesics in M can be extended indefinitely without running off the manifold. Equivalently, $(M, \langle \cdot, \cdot)$ is geodesically complete if $\exp_p$ is globally defined for all $p \in M$.

Examples: The spaces of constant curvature $\mathbb{E}^n$, $\mathbb{S}^n(a)$ and $\mathbb{H}^n(a)$ are all geodesically complete, as are the compact Lie groups with biinvariant metrics and the Grassmann manifolds. On the other hand, nonempty proper open subsets of any of these spaces are not geodesically complete.

Minimal Geodesic Theorem I. *Suppose that $(M^n, \langle \cdot, \cdot \rangle)$ is connected and geodesically complete. Then any two points p and q of M can be connected by a minimal geodesic.*

The idea behind the proof is extremely simple. Given $p \in M$, the geodesic completeness assumption implies that $\exp_p$ is globally defined. Let $a = d(p, q)$, then we should have $q = \exp_p(bv)$, where v is a unit length vector in T_pM which "points in the direction" of q.

More precisely, let $\tilde{B}_\epsilon$ be a closed ball of radius ϵ centered at 0 in T_pM, and suppose that $\tilde{B}_\epsilon$ is contained in a an open set which is mapped diffeomorphically by $\exp_p$ onto an open neighborhood of p in M. Let $\tilde{S}_\epsilon$ be the boundary of $\tilde{B}_\epsilon$ and let S be the image of $\tilde{S}$ under $\exp_p$. Since S is a compact subset of M there is a point $m \in S$ of minimal distance from q. We can write $m = \exp_p(\epsilon v)$ for some unit length $v \in T_pM$. Finally, we define

$$\gamma : [0, a] \to M \quad \text{by} \quad \gamma(t) = \exp_p(tv).$$

Then γ is a candidate for the minimal geodesic from p to q.

To finish the proof, we need to show that $\gamma(a) = q$. It will suffice to show that

$$d(\gamma(t), q) = a - t, \tag{1.38}$$

for all $t \in [0,a]$. Note that $d(\gamma(t),q) \geq a-t$, because if $d(\gamma(t),q) < a-t$, then

$$d(p,q) \leq d(p,\gamma(t)) + d(\gamma(t),q) < t + (a-t) = a.$$

Moreover, if (1.38) holds for $t_0 \in [0,a]$, it also holds for all $t \in [0,t_0]$, because if $t \in [0,t_0]$, then

$$d(\gamma(t),q) \leq d(\gamma(t),\gamma(t_0)) + d(\gamma(t_0),q) \leq (t_0 - t) + (a - t_0) = a - t.$$

We let

$$t_0 = \sup\{t \in [0,a] : d(\gamma(t),q) = a - t\},$$

and note that $d(\gamma(t_0),q) = a - t_0$ by continuity. We will show that:

1. $t_0 \geq \epsilon$, and
2. $0 < t_0 < a$ leads to a contradiction.

To establish the first of these assertions, we note that by the Theorem from §1.17,

$$d(p,q) = \inf\{d(p,r) + d(r,q) : r \in S\} = \epsilon + \inf\{d(r,q) : r \in S\} = \epsilon + d(m,q),$$

and hence $a = \epsilon + d(m,q) = \epsilon + d(\gamma(\epsilon),q)$.

To prove the second assertion, we construct a sphere S about $\gamma(t_0)$ as we did for p, and let m be the point on S of minimal distance from q. Then

$$d(\gamma(t_0),q) = \inf\{d(\gamma(t_0),r) + d(r,q) : r \in S\} = \epsilon + d(m,q),$$

and hence

$$a - t_0 = \epsilon + d(m,q), \quad \text{so} \quad a - (t_0 + \epsilon) = d(m,q).$$

Note that $d(p,m) \geq t_0 + \epsilon$ because otherwise

$$d(p,q) \leq d(p,m) + d(m,q) < t_0 + \epsilon + a - (t_0 + \epsilon) = a,$$

so the broken geodesic from p to $\gamma(t_0)$ to m has length $t_0 + \epsilon = d(p,m)$. If the broken geodesic had a corner it could be shortened by rounding off the corner. Hence m must lie on the image of γ, so $\gamma(t_0 + \epsilon) = m$, contradicting the maximality of t_0.

It follows that $t_0 = a$, $d(\gamma(a),q) = 0$ and $\gamma(a) = q$, finishing the proof of the theorem.

For a Riemannian manifold, we also have a notion of completeness in terms of metric spaces. Fortunately, the two notions of completeness coincide:

Hopf-Rinow Theorem. *Suppose that $(M^n, \langle \cdot,\cdot \rangle)$ is a connected Riemannian manifold. Then (M,d) is complete as a metric space if and only if $(M, \langle \cdot,\cdot \rangle$ is geodesically complete.*

To prove this theorem, suppose first that $(M^n, \langle \cdot,\cdot \rangle)$ is complete as a metric space but not geodesically complete. Then there is some unit speed geodesic

$\gamma : [0, b) \to M$ which extends to no interval $[0, b + \delta)$ for $\delta > 0$. Let (t_i) be a sequence from $[0, b)$ such that $t_i \to b$. If $p_i = \gamma(t_i)$, then $d(p_i, p_j) \leq |t_i - t_j|$, so (p_i) is a Cauchy sequence in (M, d). Let p_0 be the limit of (p_i). Then by Corollary 4 from §1.14, we see that there is some fixed $\epsilon > 0$ such that $\exp_{p_i}(v)$ is defined for all $|v| < \epsilon$ when i is sufficiently large. This implies γ can be extended a distance ϵ beyond p_i when i is sufficiently large, yielding a contradiction.

Thus we need only show that when $(M, \langle \cdot, \cdot \rangle$ is geodesically complete, (M, d) is complete as a metric space. Let p be a fixed point in M and (q_i) a Cauchy sequence in (M, d). We need to show that (q_i) converges to a point $q \in M$. We can assume that $d(q_i, q_j) < \epsilon$ for some $\epsilon > 0$, and let $K = d(p, q_1)$. Then $d(p, q_i) \leq K + \epsilon$ for all i, and hence $q_i = \exp_p(v_i)$ where $\|v_i\| \leq K + \epsilon$. It follows that (v_i) has a convergent subsequence, which converges to some point $v \in T_pM$. Then $q = \exp_p(v)$ is a limit of the Cauchy sequence (q_i), and (M, d) is indeed a complete metric space.

1.19 Smooth closed geodesics

If we are willing to strengthen completeness to compactness, we can give another proof of the Minimal Geodesic Theorem, which is quite intuitive and illustrates techniques that are commonly used for calculus of variations problems. Moreover, this approach is easily modified to give a proof that a compact Riemannian manifold which is not simply connected must possess a nonconstant smooth closed geodesic.

Simplifying notation a little, we let

$$\Omega(M; p, q) = \{ \text{ smooth maps } \gamma : [0, 1] \to M \text{ such that } \gamma(0) = p \text{ and } \gamma(1) = q \ \}$$

$$\text{and let} \quad \Omega(M; p, q)^a = \{\gamma \in \Omega(M; p, q) : J(\gamma) < a\}.$$

Assuming that M is compact, we can conclude that there is a $\delta > 0$ such that any p and q in M with $d(p, q) < \delta$ are connected by a unique minimal geodesic

$$\gamma_{p,q} : [0, 1] \to M \quad \text{with} \quad L(\gamma_{p,q}) = d(p, q).$$

Moreover, if $\delta > 0$ is sufficiently small, the ball of radius δ about any point is geodesically convex and $\gamma_{p,q}$ depends smoothly on p and q. If $\gamma : [a, a + \epsilon] \to M$ is a smooth path and

$$\epsilon < \frac{\delta^2}{2a}, \quad \text{then} \quad J(\gamma) \leq a \quad \Rightarrow L(\gamma) \leq \sqrt{2a\epsilon} < \delta.$$

as we see from (1.2).

Choose $N \in \mathbb{N}$ such that $1/N < \epsilon$, and if $\gamma \in \Omega(M; p, q)^a$, let $p_i = \gamma(i/N)$, for $0 \leq i \leq N$. Then γ is approximated by the map $\tilde{\gamma} : [0, 1] \to M$ such that

$$\tilde{\gamma}(t) = \gamma_{p_{i-1}p_i}\left(\frac{(i-1) + t}{N}\right), \quad \text{for} \quad t \in \left[\frac{i-1}{N}, \frac{i}{N}\right].$$

Thus $\tilde{\gamma}$ lies in the space of "broken geodesics,"

$$\begin{aligned} BG_N(M;p,q) = \{ &\text{ maps } \gamma : [0,1] \to M \text{ such that} \\ &\qquad \gamma| \left[\frac{i-1}{N}, \frac{i}{N}\right] \text{is a constant speed geodesic } \}, \end{aligned}$$

and $\Omega(M;p,q)^a$ is approximated by

$$BG_N(M;p,q)^a = \{\gamma \in BG_N(M;p,q) : J(\gamma) < a\}.$$

Suppose that γ is an element of $BG_N(M;p,q)^a$. Then if

$$p_i = \gamma\left(\frac{i}{N}\right), \quad \text{then} \quad d(p_{i-1}, p_i) \le \sqrt{\frac{2a}{N}} < \sqrt{2a\epsilon} < \delta,$$

so γ is completely determined by

$$(p_0, p_1, \dots, p_i, \dots, p_N), \quad \text{where} \quad p_0 = p, \quad p_N = q.$$

Thus we have an injection

$$\begin{aligned} j : BG_N(M;p,q)^a \to &\overbrace{M \times M \times \cdots \times M}^{N-1}, \\ &\qquad j(\gamma) = \left(\gamma\left(\frac{1}{N}\right), \dots, \gamma\left(\frac{N-1}{N}\right)\right). \end{aligned}$$

We also have a map $r : \Omega(M;p,q)^a \to BG_N(M;p,q)^a$ defined as follows: If $\gamma \in \Omega(M;p,q)^a$, let $r(\gamma)$ be the broken geodesic from

$$p = p_0 \text{ to } p_1 = \gamma\left(\frac{1}{N}\right) \text{ to } \cdots \text{ to } p_{N-1} = \gamma\left(\frac{N-1}{N}\right) \text{ to } q.$$

We can regard $r(\gamma)$ as the closest approximation to γ in the space of broken geodesics.

Minimal Geodesic Theorem II. *Suppose that $(M^n, \langle\cdot,\cdot\rangle)$ is a compact connected Riemannian manifold. Then any two points p and q of M can be connected by a minimal geodesic.*

To prove this, let

$$\mu = \inf\{J(\gamma) : \gamma \in \Omega(M;p,q)\}.$$

Choose $a > \mu$, so that $\Omega(M;p,q)^a$ is nonempty, and let (γ_j) be a sequence in $\Omega(M;p,q)^a$ such that $J(\gamma_j) \to \mu$. Let $\tilde{\gamma}_j = r(\gamma_j)$, the corresponding broken geodesic from

$$p = p_{0j} \text{ to } p_{1j} = \gamma\left(\frac{1}{N}\right) \text{ to } \cdots \text{ to } p_{(N-1)j} = \gamma\left(\frac{N-1}{N}\right) \text{ to } q,$$

and note that $J(\tilde{\gamma}_j) \leq J(\gamma_j)$.

Since M is compact, we can choose a subsequence (j_k) such that (p_{ij_k}) converges to some point $p_i \in M$ for each i. Hence a subsequence of $(\tilde{\gamma}_j)$ converges to an element $\tilde{\gamma} \in BG_N(M;p,q)^a$. Moreover,

$$J(\tilde{\gamma}_j) \leq \lim_{j\to\infty} J(\tilde{\gamma}_j) \leq \lim_{j\to\infty} J(\gamma_j) = \mu.$$

The curve $\tilde{\gamma}$ must be of constant speed, because otherwise we could decrease J be reparametrizing $\tilde{\gamma}$. Hence $\tilde{\gamma}$ must also minimize length L on $BG_N(M;p,q)^a$.

Finally, $\tilde{\gamma}$ cannot have any corners, because if it did, we could decrease length by rounding corners. (This follows from the first variation formula for piecewise smooth curves given in §1.3.2.) We conclude that $\tilde{\gamma} : [0,1] \to M$ is a smooth geodesic with $L(\tilde{\gamma}) = d(p,q)$, that is, $\tilde{\gamma}$ is a minimal geodesic from p to q, finishing the proof of the theorem.

Remark. Note that $BG_N(M;p,q)^a$ can be regarded as a finite-dimensional manifold which approximates the infinite-dimensional space $\Omega(M;p,q)^a$. This is a powerful idea which Marston Morse used in his critical point theory for geodesics. (See [25] for a thorough working out of this approach.)

Although the preceding theorem is weaker than the one presented in the previous section, the technique of proof can be extended to other contexts. We say that two smooth curves

$$\gamma_1 : S^1 \to M \quad \text{and} \quad \gamma_2 : S^1 \to M$$

are *freely homotopic* if there is a continuous path

$$\Gamma : [0,1] \times S^1 \to M \quad \text{such that} \quad \Gamma(0,t) = \gamma_1(t) \quad \text{and} \quad \Gamma(1,t) = \gamma_2(t).$$

We say that M is *simply connected* if any smooth path $\gamma : S^1 \to M$ is freely homotopic to a constant path. Thus M is simply connected if and only if its fundamental group, as defined in [14], is zero.

As before, we can approximate the space $\mathrm{Map}(S^1, M)$ of smooth maps $\gamma : S^1 \to M$ by a finite-dimensional space, where S^1 is regarded as the interval $[0,1]$ with the points 0 and 1 identified. This time the finite-dimensional space is the space of "broken geodesics,"

$$\begin{aligned} BG_N(S^1, M) = \{ &\text{ maps } \gamma : [0,1] \to M \text{ such that} \\ &\gamma| \left[\frac{i-1}{N}, \frac{i}{N}\right] \text{is a constant speed geodesic and } \gamma(0) = \gamma(1) \ \}. \end{aligned}$$

Just as before, when a is sufficiently small, then

$$\mathrm{Map}(S^1, M)^a = \{\gamma \in \mathrm{Map}(S^1, M) : J(\gamma) < a\}$$

is approximated by

$$BG_N(S^1, M)^a = \{\gamma \in BG_N(S^1, M) : J(\gamma) < a\}.$$

Moreover, if $p_i = \gamma(i/N)$, then γ is completely determined by

$$(p_1, p_2, \ldots, p_i, \ldots, p_N).$$

Thus we have an injection

$$j : BG_N(S^1, M)^a \to \overbrace{M \times M \times \cdots \times M}^{N},$$
$$j(\gamma) = \left(\gamma\left(\frac{1}{N}\right), \ldots, \gamma\left(\frac{N-1}{N}\right), \gamma(1)\right).$$

We also have a map $r : \mathrm{Map}(S^1, M)^a \to BG_N(S^1, M)^a$ defined as follows: If $\gamma \in \mathrm{Map}(S^1, M)^a$, let $r(\gamma)$ be the broken geodesic from

$$\gamma(0) \text{ to } p_1 = \gamma\left(\frac{1}{N}\right) \text{ to } \cdots \text{ to } p_{N-1} = \gamma\left(\frac{N-1}{N}\right) \text{ to } p_N = \gamma(1).$$

Closed Geodesic Theorem. *Suppose that $(M^n, \langle\cdot,\cdot\rangle)$ is a compact connected Riemannian manifold which is not simply connected. Then there is a nonconstant smooth closed geodesic in M which minimizes length among all nonconstant smooth closed curves in M^n.*

The proof is virtually identical to that for the Minimal Geodesic Theorem II except for a minor change in notation. We note that since M is not simply connected, the space

$$\mathcal{F} = \{\gamma \in \mathrm{Map}(S^1, M) : \ \gamma \text{ is not freely homotopic to a constant } \}$$

is nonempty, and we let

$$\mu = \inf\{J(\gamma) : \gamma \in \mathcal{F}\}.$$

Choose $a > \mu$, so that $\mathcal{F}^a = \{\gamma \in \mathcal{F} : J(\gamma) < a\}$ is nonempty, and let (γ_j) be a sequence in $\mathcal{F}^a$ such that $J(\gamma_j) \to \mu$. Let $\tilde{\gamma}_j = r(\gamma_j)$, the corresponding broken geodesic and from

$$p_{Nj} = \gamma(0) \text{ to } p_{1j} = \gamma\left(\frac{1}{N}\right) \text{ to } \cdots \text{ to } p_{(N-1)j} = \gamma\left(\frac{N-1}{N}\right) \text{ to } p_{Nj} = \gamma(1),$$

and note that $J(\tilde{\gamma}_j) \leq J(\gamma_j)$.

Since M is compact, we can choose a subsequence (j_k) such that (p_{ij_k}) converges to some point $p_i \in M$ for each i. Hence a subsequence of $(\tilde{\gamma}_j)$ converges to an element $\tilde{\gamma} \in BG_N(S^1, M)^a$. Moreover,

$$J(\tilde{\gamma}_j) \leq \lim_{j\to\infty} J(\tilde{\gamma}_j) \leq \lim_{j\to\infty} J(\gamma_j) = \mu.$$

The curve $\tilde{\gamma}$ must be of constant speed, because otherwise we could decrease J be reparametrizing $\tilde{\gamma}$. Hence $\tilde{\gamma}$ must also minimize length L on $BG_N(S^1, M)^a$.

Finally, $\tilde{\gamma}$ cannot have any corners, because if it did, we could decrease length by rounding corners. (This follows again from the first variation formula for piecewise smooth curves given in §1.3.2.) We conclude that $\tilde{\gamma} : S^1 \to M$ is a smooth geodesic which is not constant since it cannot even be freely homotopic to a constant.

Chapter 2

Differential forms

2.1 Tensor algebra

The key advantage of differential forms over more general tensor fields is that they pull back under smooth maps. In the next several sections, we explain how this leads to one of the simplest ways of constructing a topological invariant of smooth manifolds, namely the de Rham cohomology.

Recall that T_p^*M is the vector space of linear maps $\alpha : T_pM \to \mathbb{R}$. We define the k-fold *tensor product* $\otimes^k T_p^*M$ to be the vector space of $\mathbb{R}$-multilinear maps

$$\phi : \overbrace{T_pM \times T_pM \times \cdots \times T_pM}^{k} \longrightarrow \mathbb{R}.$$

Thus $\otimes^1 T_p^*M$ is just the space of linear functionals on T_pM which is T_p^*M itself, while by convention $\otimes^0 T_P^*M = \mathbb{R}$.

We can define a product on it as follows. If $\phi \in \otimes^k T_p^*M$ and $\psi \in \otimes^l T_p^*M$, we define $\phi \otimes \psi \in \otimes^{k+l} T_p^*M$ by

$$(\phi \otimes \psi)(v_1, \ldots, v_{k+l}) = \phi(v_1, \ldots, v_k)\psi(v_{k+1}, \ldots, v_{k+l}).$$

This multiplication is called the *tensor product* and is bilinear,

$$(a\phi + \tilde{\phi}) \otimes \psi = a\phi \otimes \psi + \tilde{\phi} \otimes \psi, \qquad \phi \otimes (a\psi + \tilde{\psi}) = a\phi \otimes \psi + \phi \otimes \tilde{\psi},$$

as well as associative,

$$(\phi \otimes \psi) \otimes \omega = \phi \otimes (\psi \otimes \omega).$$

Hence we can write $\phi \otimes \psi \otimes \omega$ with no danger of confusion. The tensor product makes the direct sum

$$\otimes^* T_p^*M = \sum_{i=0}^{\infty} \otimes^k T_p^*M$$

into a graded algebra over $\mathbb{R}$, called the *tensor algebra* of T_p^*M.

Proposition 1. *If $(x^1, \dots, x^n)$ are smooth coordinates defined on an open neighborhood of $p \in M$, then*

$$\{dx^{i_1}|_p \otimes \cdots \otimes dx^{i_k}|_p : 1 \leq i_1 \leq n, \dots, 1 \leq i_k \leq n\}$$

is a basis for $\otimes^k T_p^* M$. Thus $\otimes^k T_p^* M$ has dimension n^k.

Sketch of proof: For linear independence, suppose that

$$\sum a_{i_1 \cdots i_k} dx^{i_1}|_p \otimes \cdots \otimes dx^{i_k}|_p = 0.$$

Then

$$\begin{aligned} 0 &= \sum a_{i_1 \cdots i_k} dx^{i_1}|_p \otimes \cdots \otimes dx^{i_k}|_p \left(\left.\frac{\partial}{\partial x^{j_1}}\right|_p, \dots, \left.\frac{\partial}{\partial x^{j_k}}\right|_p \right) \\ &= \sum a_{i_1 \cdots i_k} dx^{i_1}|_p \left(\left.\frac{\partial}{\partial x^{j_1}}\right|_p \right) dx^{i_k}|_p \left(\left.\frac{\partial}{\partial x^{j_k}}\right|_p \right) = a_{j_1 \cdots j_k}. \end{aligned}$$

To show that the elements span, suppose that $\phi \in \otimes^k T_p^* M$, and show that

$$\phi = \sum a_{i_1 \cdots i_k} dx^{i_1}|_p \otimes \cdots \otimes dx^{i_k}|_p,$$
$$\text{where} \quad \sum a_{i_1 \cdots i_k} = \phi \left(\left.\frac{\partial}{\partial x^{i_1}}\right|_p, \dots, \left.\frac{\partial}{\partial x^{i_k}}\right|_p \right).$$

We let $\Lambda^k T_p^* M$ denote the space of skew-symmetric elements of $\phi \in \otimes^k T_p^* M$. By skew-symmetric, we mean that the value of ϕ changes sign when two distinct arguments are interchanged,

$$\phi(v_1, \dots, v_i, \cdots, v_j, \dots, v_k) = -\phi(v_1, \dots, v_j, \cdots, v_i, \dots, v_k),$$

whenever $i \neq j$.

This can be expressed in terms of the symmetric group S_k on k letters. Recall that by definition, S_k is the group of bijections from the set $\{1, \dots, k\}$ onto itself, with composition being the group operation. We define a function

$$\mathrm{sgn} : S_k \to \{\pm 1\} \quad \text{by} \quad \mathrm{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(i) - \sigma(j)}{i - j},$$

and check that it is a group homomorphism. We say that an element $\sigma \in S_k$ is *even* if $\mathrm{sgn}(\sigma) = 1$, *odd* if $\mathrm{sgn}(\sigma) = -1$. Then a multlinear map

$$\phi : \overbrace{T_p M \times T_p M \times \cdots \times T_p M}^{k} \longrightarrow \mathbb{R}$$

is skew-symmetric if

$$\phi(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = (\mathrm{sgn}\sigma)\phi(v_1, \dots v_k),$$

for all $\sigma \in S_k$.

Note that $\Lambda^k T_p^* M$ is a linear subspace of $\otimes^k T_p^* M$. We define a projection

$$\mathrm{Alt} : \otimes^k T_p^* M \longrightarrow \Lambda^k T_p^* M$$

by

$$\mathrm{Alt}(\phi)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\mathrm{sgn}\sigma)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

It is a straightforward exercise to show that

$$\mathrm{Alt}(\mathrm{Alt}(\phi \otimes \psi) \otimes \omega) = \mathrm{Alt}(\phi \otimes \psi \otimes \omega) = \mathrm{Alt}(\phi \otimes \mathrm{Alt}(\psi \otimes \omega). \tag{2.1}$$

For details, one can check the argument for Lemma 6.6 of Chapter 5 in [5].

If $\phi \in \Lambda^k T_p^* M$ and $\psi \in \Lambda^l T_p^* M$, we can define $\phi \wedge \psi \in \Lambda^{k+l} T_p^* M$ by

$$\phi \wedge \psi = \frac{(k+l)!}{k!l!} \mathrm{Alt}(\phi \otimes \psi).$$

This multiplication is called the *wedge product*. It is bilinear,

$$(a\phi + \tilde{\phi}) \wedge \psi = a\phi \wedge \psi + \tilde{\phi} \wedge \psi, \qquad \phi \wedge (a\psi + \tilde{\psi}) = a\phi \wedge \psi + \phi \wedge \tilde{\psi},$$

skew-commutative

$$\phi \wedge \psi = (-1)^{kl} \psi \wedge \phi, \quad \text{for} \quad \phi \in \Lambda^k T_p^* M \quad \text{and} \quad \psi \in \Lambda^l T_p^* M,$$

and associative

$$(\phi \wedge \psi) \wedge \omega = \phi \wedge (\psi \wedge \omega).$$

Only the last fact is nontrivial, and it follows rather quickly from identity (2.1). This product makes the direct sum

$$\Lambda^* T_p^* M = \sum_{i=0}^{n} \Lambda^k T_p^* M$$

into a graded commutative algebra over $\mathbb{R}$, called the *exterior algebra* of $T_p^* M$.

Proposition 2. *If $(x^1, \ldots, x^n)$ are smooth coordinates defined on an open neighborhood of $p \in M$, then*

$$\{dx^{i_1}|_p \wedge \cdots \wedge dx^{i_k}|_p : 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$$

is a basis for $\otimes^k T_p^* M$. Thus $\otimes^k T_p^* M$ has dimension $\binom{n}{k}$.

the proof is quite similar to that of Proposition 1.

2.2 The exterior derivative

We now let $\Lambda^k T^* M = \bigcup\{\Lambda^k T_p^* M\}$, a disjoint union. Just as in the case of the tangent and cotangent bundles, $\Lambda^k T^* M$ has a smooth manifold structure, together with a projection $\pi : \Lambda^k T^* M \to M$ such that $\pi(\Lambda^k T_p^* M) = p$. We can describe the coordinates for the smooth structure on $\Lambda^k T^* M$ as follows: If $(x^1, \ldots, x^n)$ are smooth coordinates on an open set $U \subseteq M$, the corresponding smooth coordinates on $\pi^{-1}(U)$ are are the pullbacks of $(x^1, \ldots, x^n)$ to $\pi^{-1}(U)$, together with the additional coordinates $p_{i_1 \cdots i_k} : \pi^{-1}(U) \to \mathbb{R}$ defined by

$$p_{i_1 \cdots i_k}\left(\sum_{j_1 < \cdots < j_k} a_{j_1 \cdots j_k} dx^{j_1}|_p \wedge \cdots \wedge dx^{j_k}|_p \right) = a_{i_1 \cdots i_k}.$$

If U is an open subset of M, a *differential k-form* or a *differential form of degree k* on U is a smooth map $\omega : U \to \Lambda^k T_p^* M$ such that $\pi \circ \omega = \mathrm{id}_U$. Informally, we can say that a differential k-form on U is a function ω which assigns to each point $p \in U$ an element $\omega(p) \in \Lambda^k T_p^* M$ in such a way that $\omega(p)$ varies smoothly with p.

Let $\Omega^k(U)$ denote the real vector space of differential k-forms on U. If $(U, (x^1, \ldots, x^n))$ is a smooth coordinate system on M, we can define

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} \in \Omega^k(U) \quad \text{by} \quad (dx^{i_1} \wedge \cdots \wedge dx^{i_k})(p) = dx^{i_1}|_p \wedge \cdots \wedge dx^{i_k}|_p.$$

Then any element $\omega \in \Omega^k(U)$ can be written uniquely as a sum

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where $f_{i_1 \cdots i_k} : U \to \mathbb{R}$ is a smooth function. If $\omega \in \Omega^k(U)$ and $\phi \in \Omega^l(U)$, then we can define the wedge product $\omega \wedge \phi \in \Omega^{k+l}(U)$ by

$$(\omega \wedge \phi)(p) = \omega(p) \wedge \phi(p).$$

Note that if $f \in \Omega^0(M) = \mathcal{F}(M)$ and $\omega \in \Omega^k(M)$, then $f \wedge \omega = f\omega$.

Exterior Derivative Theorem. *There is a unique collection of linear maps of real vector spaces,*

$$d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M),$$

which satisfy the following conditions:

1. *If ω is a k-form, the value $d\omega(p)$ depends only on ω and its derivatives at p.*

2. *If f is a smooth real-valued function regarded as a differential 0-form, $d(f)$ is the differential of f defined before.*

3. *$d \circ d = 0$.*

4. *If ω is a k-form and ϕ is an l-form, then*

$$d(\omega \wedge \phi) = (d\omega) \wedge \phi + (-1)^k \omega \wedge (d\phi).$$

We call d the *exterior derivative*.

We begin the proof of the theorem by establishing uniqueness. By property 1, it suffices to prove uniqueness in the case where $M = U$, where U is the domain of a local coordinate system $(x^1, \ldots, x^n)$. If $\omega \in \Omega^k(U)$, we can write

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where $f_{i_1 \cdots i_k} : U \to \mathbb{R}$ is a smooth function. Using linearity and the axiom for products we now find that

$$\begin{aligned} d\omega &= \sum_{i_1 < \cdots < i_k} d\left(f_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) \\ &= \sum_{i_1 < \cdots < i_k} df_{i_1 \cdots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} + \sum_{i_1 < \cdots < i_k} f_{i_1 \cdots i_k} d\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right). \end{aligned}$$

Using the axiom for products, the fact that $d \circ d = 0$ and induction, one shows that

$$d\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) = 0.$$

Hence

$$d\omega = \sum_{i_1 < \cdots < i_k} df_{i_1 \cdots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \tag{2.2}$$

where $df_{i_1 \cdots i_k} \in \Omega^1(U)$ is the previously defined differential of a function. This formula establishes uniqueness.

We next prove local existence, existence on U where U is the domain of a local coordinate system $(x^1, \ldots, x^n)$. To do this, we can define $d\omega$ by (2.2) and check that it satisfies the axioms. The first two axioms are immediate. To establish the last axiom, we use the easily proven formula

$$d(fg) = g(df) + f(dg).$$

Suppose that

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \text{ and } \phi = \sum_{j_1 < \cdots < j_l} g_{j_1 \cdots j_l} dx^{j_1} \wedge \cdots \wedge dx^{j_l}.$$

Then

$$\omega \wedge \phi = \sum f_{i_1 \cdots i_k} g_{j_1 \cdots j_l} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l},$$

and hence

$$\begin{aligned}
d(\omega\wedge\phi) &= \sum d(f_{i_1\cdots i_k}g_{j_1\cdots j_l})\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_l}\\
&= \sum g_{j_1\cdots j_l}df_{i_1\cdots i_k}\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_l}\\
&+ \sum f_{i_1\cdots i_k}dg_{j_1\cdots j_l}\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\cdots\wedge dx^{j_l}\\
&= d\omega\wedge\phi+(-1)^k\omega\wedge d\phi.
\end{aligned}$$

For the third axiom, we use the equality of mixed partial derivatives. First, we note that if $f\in\Omega^0(M)$, then

$$\begin{aligned}
d(df) = d\left(\sum_{j=1}^n \frac{\partial f}{\partial x^j}dx^j\right) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x^i\partial x^j}dx^i\wedge dx^j\\
&= \sum_{i<j}\left[\frac{\partial^2 f}{\partial x^i\partial x^j}-\frac{\partial^2 f}{\partial x^j\partial x^i}\right]dx^i\wedge dx^j = 0.
\end{aligned}$$

In general, if $\omega\in\Omega^k(U)$, say

$$\omega = \sum_{i_1<\cdots<i_k} f_{i_1\cdots i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k},$$

we find that

$$\begin{aligned}
&d(d\omega) = d\left(\sum_{i_1<\cdots<i_k} df_{i_1\cdots i_k}\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}\right)\\
&= \sum_{i_1<\cdots<i_k} d(df_{i_1\cdots i_k})\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k} - \sum_{i_1<\cdots<i_k} df_{i_1\cdots i_k}\wedge d(dx^{i_1}\wedge\cdots\wedge dx^{i_k}) = 0.
\end{aligned}$$

This finishes the proof of local existence. To prove global existence, we note that the locally defined exterior derivative operators must fit together on overlaps due to uniqueness, and hence they fit together to form a globally defined exterior derivative operator on M.

Example. The exterior derivative is actually an extension of the gradient, divergence and curl operators one meets in several variable calculus. Thus suppose that $M=\mathbb{E}^3$ with the standard euclidean coordinates (x,y,z) and let

$$d\mathbf{x} = dx\mathbf{i}+dy\mathbf{j}+dz\mathbf{k},\quad \mathbf{N}dA = (dy\wedge dz)\mathbf{i}+(dz\wedge dx)\mathbf{j}+(dx\wedge dy)\mathbf{k},$$

where $(\mathbf{i},\mathbf{j},\mathbf{k})$ is the usual orthonormal basis. If $f\in\Omega^0(\mathbb{E}^3)$ is a smooth function,

$$df = \frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dx = (\text{gradient of } f)\cdot d\mathbf{x}.$$

If $\theta = \mathbf{F} \cdot d\mathbf{x}$ is an element of $\Omega^1(\mathbb{E}^3)$, where $\mathbf{F}$ is a vector field, say $\theta = Pdx + Qdy + Rdz$, then

$$\begin{aligned} d\theta = &\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy \wedge dz \\ &+ \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz \wedge dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy \\ &= (\text{curl of } \mathbf{F}) \cdot \mathbf{N} dA. \end{aligned}$$

Finally, if $\omega = \mathbf{F} \cdot \mathbf{N} dA$ is an element of $\Omega^2(\mathbb{E}^3)$, say

$$\omega = Pdy \wedge dz + Qdy \wedge dx + Rdx \wedge dy,$$

then

$$d\omega = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \wedge dy \wedge dz = (\text{divergence of } \mathbf{F}) dx \wedge dy \wedge dz.$$

The exterior derivative extends these familiar operations from calculus to arbitrary smooth manifolds in such a way that they are "natural under smooth maps."

We now explain what we mean by natural under smooth maps. Suppose that $F : M \to N$ is a smooth map. If $p \in M$, the linear map

$$(F_*)_p : T_pM \longrightarrow T_{F(p)}N$$

induces a linear map

$$F_p^* : \Lambda^k T^*_{F(p)} N \longrightarrow \Lambda^k T^*_p M$$

By

$$F_p^*(\phi)(v_1, \ldots, v_k) = \phi((F_*)_p(v_1), \ldots, (F_*)_p(v_k)).$$

This in turn induces a linear map

$$F^* : \Omega^k(N) \to \Omega^k(M) \quad \text{by} \quad F^*(\omega)(p) = F_p^*(\omega(F(p))).$$

If $f \in \Omega^0(M)$, we agree to let $F^*(f) = f \circ F$.

Proposition. *The map F^* preserves wedge products and exterior derivatives:*

1. $F^*(\omega \wedge \theta) = F^*(\omega) \wedge F^*(\theta)$.
2. $d(F^*(\omega)) = F^*(d\omega)$.

We leave the proof of the first of these facts as an easy exercise. We first check the second for the case of a function $f \in \Omega^0(M)$. In this case, if $v \in T_pM$,

$$\begin{aligned} F^*(df)(v) = df((F_*)_p(v)) &= (F_*)_p(v)(f) \\ &= v(f \circ F) = d(f \circ F)(v) = d(F^*(f))(v). \end{aligned}$$

We next check in the case where $N = U$, the domain of a local coordinate system $(x^1, \dots, x^n)$. In this case,

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

and

$$d\omega = \sum_{i_1 < \cdots < i_k} df_{i_1 \cdots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Using the first assertion of the proposition, we see that

$$\begin{aligned} F^*(d\omega) &= \sum_{i_1 < \cdots < i_k} F^*(df_{i_1 \cdots i_k}) \wedge F^*(dx^{i_1}) \wedge \cdots \wedge F^*(dx^{i_k}) \\ &= \sum_{i_1 < \cdots < i_k} d(f_{i_1 \cdots i_k} \circ F) \wedge d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F). \end{aligned} \tag{2.3}$$

On the other hand,

$$F^*\omega = \sum_{i_1 < \cdots < i_k} (f_{i_1 \cdots i_k} \circ F) d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F),$$

so

$$dF^*\omega = \sum_{i_1 < \cdots < i_k} d(f_{i_1 \cdots i_k} \circ F) \wedge d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F). \tag{2.4}$$

Comparing (2.3) and (2.4) we see that $F^* \circ d = d \circ F^*$, when N has a global coordinate system.

Since the operators are local, $F^* \circ d = d \circ F^*$ on any smooth manifold M.

If $\omega \in \Omega^k(M)$ and $X_1, \dots, X_k$ are smooth vector fields on M, we can define a smooth function $\omega(X_1, \dots, X_k)$ on M by

$$\omega(X_1, \dots, X_k)(p) = \omega(p)(X_1(p), \dots, X_k(p)).$$

Exercise V. Show that if X and Y are smooth vector fields on M and $\theta \in \Omega^1(M)$ is a smooth one-form, then

$$d\theta(X, Y) = X(\theta(Y)) - Y(\theta(X)) - \theta([X, Y]).$$

Hint: Use local coordinates.

2.3 Integration of differential forms

The way to think of differential forms of degree n is that they are integrands for multiple integrals over n-dimensional oriented manifolds.

Two smooth charts $(U, (x^1, \dots, x^n))$ and $V, (y^1, \dots, y^n))$ on an n-dimensional smooth manifold M are said to be *coherently oriented* if

$$\det \left(\frac{\partial y^i}{\partial x^j} \right) > 0,$$

where defined. We say that M is *orientable* if it possesses an atlas of coherently oriented charts. Such an atlas is called an *orientation* for M. An *oriented* smooth manifold is a smooth manifold together with a choice of orientation.

Suppose that M is a smooth manifold with orientation defined by the atlas $\mathcal{A}$ of coherently oriented charts. A chart $V, (y^1, \dots, y^n))$ on M is said to be *positively oriented* if

$$\det \left(\frac{\partial y^i}{\partial x^j} \right) > 0,$$

where defined, for every chart $(U, (x^1, \dots, x^n))$ in $\mathcal{A}$.

Local integration of n-forms: Suppose that ω is a smooth n-form with compact support in an open subset U of a smooth n-dimensional oriented manifold M. If $\phi = (x^1, \dots, x^n)$ are positively oriented coordinates on U, we can write

$$f dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n.$$

We can then define the integral of ω over U by the formula

$$\int_U \omega = \int_{\mathbb{R}}^n (f \circ \phi^{-1}) dx^1 \cdots dx^n.$$

Thus to integrate an n-form over an oriented n-manifolds, we essentially we just leave out the wedges and take the ordinary Riemann integral.

We need to check that this definition is independent of choice of positively oriented smooth coordinates. To do this, note that if $\psi = (y^1, \dots, y^n)$ is a second positively oriented coordinate system on U, then

$$dx^i = \sum_{j=1}^n \frac{\partial x^i}{\partial y^j} dy^j,$$

and hence

$$\begin{aligned} dx^1 \wedge \cdots \wedge dx^n &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \frac{\partial x^1}{\partial y^{j_1}} \cdots \frac{\partial x^n}{\partial y^{j_n}} dy^{j_1} \wedge \cdots \wedge y^{j_n} \\ &= \sum_{\sigma \in S_n} \frac{\partial x^1}{\partial y^{\sigma(1)}} \cdots \frac{\partial x^n}{\partial y^{\sigma(n)}} dy^{\sigma(1)} \wedge \cdots \wedge y^{\sigma(n)} \\ &= \sum_{\sigma \in S_n} (\mathrm{sgn} \sigma) \frac{\partial x^1}{\partial y^{\sigma(1)}} \cdots \frac{\partial x^n}{\partial y^{\sigma(n)}} dy^1 \wedge \cdots \wedge dy^n. \end{aligned}$$

Recall that if $A = (a^i_j)$ is an arbitrary $n \times n$ matrix,

$$\det A = \sum_{\sigma \in S_n} (\mathrm{sgn} \sigma) a^1_{\sigma(1)} \cdots a^n_{\sigma(n)},$$

and since the two coordinate systems are coherently oriented,

$$dx^1 \wedge \cdots \wedge dx^n = \det \left(\frac{\partial y^i}{\partial x^j} \right) dy^1 \wedge \cdots \wedge dy^n = \left| \det \left(\frac{\partial y^i}{\partial x^j} \right) \right| dy^1 \wedge \cdots \wedge dy^n.$$

Thus the two expressions for the integral of ω over U will agree if and only if

$$\int_{\mathbb{R}}^{n} (f \circ \phi^{-1}) dx^1 \cdots dx^n = \int_{\mathbb{R}}^{n} (f \circ \psi^{-1}) \left| \det \left(\frac{\partial y^i}{\partial x^j} \right) \right| dx^1 \cdots dx^n.$$

But this is just the formula for change of variable in a multiple integral familiar from several variable calculus. A proof can be found in any undergraduate text on real analysis, such as Rudin [31], Theorem 9.32.

Local integration of differential forms has the usual linearity property:

$$\int_U (c_1\omega_1 + c_1\omega_2) = c_1 \int_U \omega_1 + c_2 \int_U \omega_2,$$

where c_1 and c_2 are constants and ω_1 and ω_2 are n-forms with compact support within U.

Global integration of n-forms: Suppose now that ω is a smooth n-form with compact support on a smooth oriented n-dimensional manifold. Choose an open covering $\{U_\alpha : \alpha \in A\}$ such that each U_α is the domain of a positively oriented coordinate system, and let $\{\psi_\alpha : \alpha \in A\}$be a partition of unity subordinate to the open cover $\{U_\alpha : \alpha \in A\}$. Then we can apply the preceding construction to each differential form $\psi_\alpha\omega$ and define the integral of ω over M by the formula

$$\int_M \omega = \sum_{\alpha \in A} \int_{U_\alpha} \psi_\alpha \omega.$$

The sum is actually finite, since ω has compact support and the supports of $\{\psi_\alpha : \alpha \in A\}$ are locally finite.

We need to check that this definition is independent of choice of cover and partition of unity. Suppose that $\{V_\beta : \beta \in A\}$ is another open cover by domains of positively oriented coordinate systems and that $\{\eta_\beta : \beta \in B\}$ is a subordinate partition of unity. Then

$$\sum_{\alpha \in A} \int_{U_\alpha} \psi_\alpha \omega = \sum_{\alpha \in A} \sum_{\beta \in B} \int_{U_\alpha \cap V_\beta} \psi_\alpha \eta_\beta \omega = \sum_{\beta \in B} \int_{V_\beta} \eta_\beta \omega.$$

We can now make precise what we mean when we say that k-forms are integrands for integrals over k-dimensional manifolds. Suppose that M is an n-dimensional manifold and S is a k-dimensional oriented submanifold of M with inclusion $\iota : S \to M$. If $\omega \in \Omega^k(M)$, we can define the integral of ω over S,

$$\int_S \iota^* \omega.$$

In fact, if $F : S \to M$ is simply a smooth map, we can define the integral of ω over the *singular manifold* (S, F) as

$$\int_S F^* \omega.$$

A diffeomorphism $F : M \to N$ from an oriented manifold M to an oriented manifold N is *orientation preserving* if whenever $(x^1, \ldots, x^n)$ is a positively oriented coordinate system on $U \subseteq N$, then $(x^1 \circ F, \ldots, x^n \circ F)$ is a positively oriented coordinate system on $F^{-1}(U) \subseteq M$. It is easily verified that if M and N are compact oriented manifolds and $F : M \to N$ is an orientation-preserving diffeomorphism, then

$$\int_M F^*\omega = \int_N \omega, \quad \text{for } \omega \in \Omega^n(N).$$

(To see this, just note that if one performs the integration with respect to the pulled back coordinate systems one must get the same result.) Here is an important application of integration over compact oriented manifolds:

Theorem. *Let G be a compact Lie group. Then G possesses a biinvariant Riemannian metric.*

Sketch of proof: A one-form ω is left invariant if $L_\sigma^*(\omega) = \omega$ for every $\sigma \in G$. The left invariant one-forms on G form a vector space $\mathfrak{g}^*$ which is dual to the Lie algebra. Ir $(\omega_1, \ldots, \omega_n)$ is a basis for $\mathfrak{g}^*$, then $\omega_1 \wedge \cdots \wedge \omega^n$ is a left invariant n-form on G. We normalize the basis so that

$$\int_G \omega_1 \wedge \cdots \wedge \omega^n = 1,$$

and define the *Haar integral* of a function $f : G \to \mathbb{R}$ by

$$\int_G f(\sigma)d\sigma = \int_G f\omega_1 \wedge \cdots \wedge \omega^n.$$

The key feature of the Haar integral is that it is invariant under left translation,

$$\int_G f(\sigma)d\sigma = \int_G f(\tau\sigma)d\sigma, \quad \text{for } \tau \in G,$$

because

$$L_\tau^*(f\omega_1 \wedge \cdots \wedge \omega_n) = (f \circ L_\tau)\omega_1 \wedge \cdots \wedge \omega^n.$$

Any positive definite inner product

$$\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$$

defines a left invariant Riemannian metric on G. We want to construct a Riemannian metric which is also right invariant. For $\sigma \in G$, we can define a map

$$\mathrm{Ad}(\sigma)\mathfrak{g} \to \mathfrak{g} \quad \text{by} \quad \mathrm{Ad}(\sigma)(X) = (L_\sigma)_*(R_{\sigma-1})_*(X).$$

It is easily checked that Ad is a group homomorphism from G into the group $\mathrm{Aut}(\mathfrak{g})$ of linear automorphisms of $\mathfrak{g}$. The left invariant metric defined by a positive definite inner product $\langle \cdot, \cdot \rangle$ is also right invariant if

$$\langle \mathrm{Ad}(\sigma)(X), \mathrm{Ad}(\sigma)(X) \rangle = \langle X, X \rangle, \quad \text{for all } X \in \mathfrak{g} \text{ and all } \sigma \in G. \tag{2.5}$$

We can define such a metric

$$\langle\langle \cdot, \cdot \rangle\rangle : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$$

by setting

$$\langle\langle X, Y \rangle\rangle = \int_G \langle \mathrm{Ad}(\sigma)X, \mathrm{Ad}(\sigma)Y \rangle d\sigma,$$

thereby obtaining the desired biinvariant metric on G.

Remark: Note that this argument requires G to be compact. Most noncompact Lie groups do not possess biinvariant Riemannian metrics. However, if G is any Lie group with Lie algebra $\mathfrak{g}$ and

$$ad(X) : \mathfrak{g} \to \mathfrak{g} \quad \text{by} \quad ad(X)(Y) = [X, Y],$$

for $X \in \mathfrak{g}$, then the "Killing form"

$$\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R} \quad \text{defined by} \quad \langle X, Y \rangle = \mathrm{Trace}(ad(X)ad(Y))$$

satisfies (2.5). Hence if it is nondegenerate it defines a biinvariant pseudo-Riemannian metric on G and geodesics and curvature can be calculated exactly as in §1.12. A Lie group is said to be *semisimple* if its Killing form is nondegenerate. Many noncompact Lie groups are semisimple; for example, with some effort one could show that

$$SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R} : \det A = 1\}$$

is semisimple. Thus we can obtain many examples of biinvariant pseudo-Riemannian metrics this way.

Exercise VI. For $1 \le i, j \le n$, define functions $x^i_j, y^i_j : GL(n, \mathbb{R}) \to \mathbb{R}$ by

$$x^i_j \begin{pmatrix} a^1_1 & \cdots & a^1_n \\ \cdot & \cdots & \cdot \\ a^n_1 & \cdots & a^n_n \end{pmatrix} = a^i_j, \quad y^i_j = x^i_j(A^{-1}).$$

a. Show that the differential form

$$\omega^i_j = \sum_{k=1}^n y^i_k dx^k_j$$

is left invariant.

b. Establish the *Maurer-Cartan equations* for $GL(n, \mathbb{R})$:

$$d\omega^i_j = -\sum_{k=1}^n \omega^i_k \wedge \omega^k_j.$$

2.4 Theorem of Stokes

The Theorem of Stokes generalizes Green's Theorem from several variable calculus to manifolds of arbitrary dimension. To set up the context, we first define manifolds with boundary. Let

$$\mathbb{R}^n_- = \{(x^1, \ldots, x^n) : x_1 \leq 0\}, \quad \partial\mathbb{R}^n_- = \{(x^1, \ldots, x^n) : x_1 = 0\}.$$

In this section, we will identify $\partial\mathbb{R}^n_-$ with $\mathbb{R}^{n-1}$.

Definition. An n-dimensional *smooth manifold with boundary* is a metrizable space M together with a collection $\mathcal{A} = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$ such that:

1. Each ϕ_α is a homeomorphism from an open subset U_α of M onto an open subset of $\mathbb{R}^n_-$.
2. $\bigcup\{U_\alpha : \alpha \in A\} = M$.
3. $\phi_\beta \circ \phi_\alpha^{-1}$ is C^∞ where defined.

The *boundary* of M is

$$\partial M = \{p \in M : \phi_\alpha(p) \in \partial\mathbb{R}^n_- \text{ for some } \alpha \in A\ \}.$$

Lemma. *If $\phi_\alpha(p) \in \partial\mathbb{R}^n_-$ for some $\alpha \in A$, then $\phi_\beta(p) \in \partial\mathbb{R}^n_-$ for all $\beta \in A$ for which $\phi_\beta(p)$ is defined.*

Proof: Suppose on the contrary that $\phi_\beta(p)$ lies in the interior of $\mathbb{R}^n_-$ for some β. Note that $\phi_\alpha \circ \phi_\beta^{-1}$ has a nonsingular differential at $\phi_\beta(p)$. Hence by the inverse function theorem, $\phi_\alpha \circ \phi_\beta^{-1}$ maps some neighborhood V_β of $\phi_\beta(p)$ onto an open neighborhood V_α of $\phi_\alpha(p)$. Then V_α is open in $\mathbb{R}^n$ and $V_\alpha \subseteq \mathbb{R}^n_-$. Hence $\phi_\alpha(p)$ lies in the interior of $\mathbb{R}^n_-$, a contradiction.

If $\alpha \in A$, we let $V_\alpha = U_\alpha \cap \partial M$ and let $\psi_\alpha = \phi_\alpha | V_\alpha$. The lemma shows that ψ_α is $\mathbb{R}^{n-1}$-valued. Thus ∂M become a $(n-1)$-dimensional smooth manifold with smooth atlas $\{V_\alpha, \psi_\alpha) : \alpha \in A\}$.

Orientation: Suppose that M is an oriented n-dimensional manifold with boundary, so that M has an atlas $\mathcal{A}$ whose elements are coherently oriented. Thus is $(U, (x^1, \ldots, x^n))$ and $(V, (y^1, \ldots, y^n))$ are two elements of $\mathcal{A}$, then

$$\det\left(\frac{\partial y^i}{\partial x^j}\right) > 0,$$

where defined.

Then $(U \cap \partial M, (x^2, \ldots, x^n))$ and $(V \cap \partial M, (y^2, \ldots, y^n))$ are smooth coordinate systems on ∂M. We claim that they are coherently oriented. Indeed, if $p \in \partial M \cap (U \cap V)$,

$$\frac{\partial y^1}{\partial x^i}(p) = 0, \quad \text{for } 2 \leq i \leq n, \quad \text{since} \quad \left.\frac{\partial}{\partial x^i}\right|_p$$

is tangent to ∂M and y^1 is constant along ∂M. On the other hand, $(\partial/\partial x^1)|p$ points out of M, that is in the direction of increasing y^1. Hence

$$\det\begin{pmatrix} (\partial y^1/\partial x^1)(p) & 0 & \cdots & 0 \\ * & (\partial y^2/\partial x^2)(p) & \cdots & (\partial y^2/\partial x^n)(p) \\ * & \cdot & \cdots & \cdot \\ * & (\partial y^n/\partial x^2)(p) & \cdots & (\partial y^n/\partial x^n)(p) \end{pmatrix} > 0$$

implies that

$$\det\begin{pmatrix} (\partial y^2/\partial x^2)(p) & \cdots & (\partial y^2/\partial x^n)(p) \\ \cdot & \cdots & \cdot \\ (\partial y^n/\partial x^2)(p) & \cdots & (\partial y^n/\partial x^n)(p) \end{pmatrix} > 0,$$

and hence $(U \cap \partial M, (x^2, \ldots, x^n))$ and $(V \cap \partial M, (y^2, \ldots, y^n))$ are indeed coherently oriented.

We conclude that an orientation on a smooth manifold with boundary induces an orientation on its boundary.

Stokes' Theorem. *Let M be an oriented smooth manifold with boundary ∂M, and given ∂M the induced orientation. Let $\iota : \partial M \to M$ be the inclusion map. If θ is a smoth $(n-1)$-form on M with compact support, then*

$$\int_{\partial M} \iota^*\theta = \int_M d\theta. \tag{2.6}$$

Proof: Cover M by positively oriented charts $\{(U_\alpha, \phi_\alpha) : \alpha \in A\}$ such that if $\alpha in A$, either

$$\phi_\alpha(U_\alpha) = (a_\alpha^1, b_\alpha^1) \times \cdots \times (a_\alpha^n, b_\alpha^n), \quad \text{or} \tag{2.7}$$

$$\phi_\alpha(U_\alpha) = (a_\alpha^1, 0] \times \cdots \times (a_\alpha^n, b_\alpha^n). \tag{2.8}$$

Let $\{\psi_\alpha : \alpha \in A\}$ be a partition of unity subordinate to the open cover $\{U_\alpha : \alpha \in A\}$.

It will suffice to prove Stokes' Theorem for the special case where the support of θ is contained in some U_α for $\alpha \in A$. Indeed, assuming this special case, we find that if θ is an arbitrary $(n-1)$-form with compact support,

$$\int_{\partial M} \iota^*\theta = \int_{\partial M} \iota^*\left(\sum_{\alpha\in A} \psi_\alpha\theta\right) = \sum_{\alpha\in A}\int_{\partial M} \iota^*(\psi_\alpha\theta)$$
$$= \sum_{\alpha\in A}\int_{\partial M} d(\psi_\alpha\theta) = \int_{\partial M} d\left(\sum_{\alpha\in A}\psi_\alpha\theta\right) = \int_M d\theta.$$

Thus it suffices to prove Stokes' Theorem in the special case where the support of θ is contained in U, where U is the domain of a chart (U,ϕ) of type (2.7) or (2.8). Since the case of type (2.7) is simpler, we consider only the case of type (2.8), and suppose that

$$\phi(U) = (a^1, 0] \times \cdots \times (a^n, b^n).$$

Indeed, we can assume that ϕ is the identity and that U itself is a rectangular set in $\mathbb{R}^n_-$.

Thus we let $(x^1, \ldots, x^n)$ be the usual rectangular cartesian coordinates on U and that the inclusion $\iota : \partial\mathbb{R}^n_- \to \mathbb{R}^n_-$ is defined by

$$\iota(x^2, \ldots, x^n) = (0, x^2, \ldots, x^n).$$

If θ is a smooth $(n-1)$-form on U with compact support in U, we can write

$$\theta = \sum_{i=1}^{n} (-1)^{i-1} f_i dx^1 \wedge \cdots dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^n.$$

Then

$$\iota^*\theta = (f_1 \circ \iota) dx^2 \wedge \cdots \wedge dx^n,$$

while

$$d\theta = \sum_{i=1}^{n} \frac{\partial f^i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n.$$

To verify Stokes' Theorem, we need to calculate the two integrals appearing in (2.6) with $M = U$. For the right-hand integral, we obtain

$$\begin{aligned}\int_U d\theta = \sum_{i=1}^{n} \int_U \frac{\partial f^i}{\partial x^i}(x^1, \ldots, x^n) dx^1 \cdots dx^n \\ = \sum_{i=1}^{n} \int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} \int_{a^1}^{0} \frac{\partial f^i}{\partial x^i}(x^1, \ldots, x^n) dx^1 \cdots dx^n.\end{aligned}$$

Now we note that for $2 \le i \le n$,

$$\begin{aligned}&\int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} \int_{a^1}^{0} \frac{\partial f^i}{\partial x^i}(x^1, \ldots, x^n) dx^1 \cdots dx^n \\ &\quad = \int_{a^n}^{b^n} \cdots \int_{a^{i+1}}^{b^{i+1}} \int_{a^{i-1}}^{b^{i-1}} \cdots \int_{a^1}^{0} \\ &\qquad [f(x^1, \ldots b^i, \cdots, x^n) - f(x^1, \ldots a^i, \cdots, x^n)] \\ &\qquad\qquad dx^1 \cdots dx^{i-1} dx^{i+1} \cdots dx^n = 0,\end{aligned}$$

because f^i has compact support in U, while in the remaining case, we get

$$\begin{aligned}&\int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} \int_{a^1}^{0} \frac{\partial f^1}{\partial x^1}(x^1, \ldots, x^n) dx^1 \cdots dx^n \\ &\quad = \int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} [f(0, \ldots, x^n) - f(a^1, \ldots, x^n)] dx^2 \cdots dx^n \\ &\qquad = \int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} f(0, \ldots, x^n) dx^2 \cdots dx^n.\end{aligned}$$

Thus

$$\int_U d\theta = \int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} f(0, \ldots, x^n) dx^2 \cdots dx^n. \tag{2.9}$$

On the other hand,

$$\begin{aligned} \int_{\partial U} \iota^* \theta = \int_{\partial U} (f_1 \circ \iota)(x^2, \ldots, x^n) dx^2 \cdots dx^n \\ = \int_{a^n}^{b^n} \cdots \int_{a^2}^{b^2} f(0, \ldots, x^n) dx^2 \cdots dx^n. \end{aligned} \tag{2.10}$$

Stokes' Theorem now follows from (2.9) and (2.10).

2.5 de Rham Cohomology

The basic idea of algebraic topology is that of constructing functors from the category of smooth manifolds and smooth maps, or more generally the category of topological spaces and continuous maps to some algebraic category, such as the category of $\mathbb{R}$-algebras and $\mathbb{R}$-algebra homomorphisms. Often these functors can be utilized to translate topological problems into algebraic problems which may be easier to solve. We refer the reader to [7] for a systematic treatment of algebraic topology emphasizing differential forms and the de Rham theory.

The de Rham cohomology is a candidate for the simplest algebraic topology functor, and has the advantage of being well-adapted to applications in differential geometry. We now give a brief introduction to de Rham theory.

We say that an element $\omega \in \Omega^k(M)$ is *closed* if $d\omega = 0$ and *exact* if $\omega = d\theta$ for some $\theta \in \Omega^{k-1}(M)$. We let

$$Z^k(M) = (\text{closed elements of } \Omega^k(M)) = \{\omega \in \Omega^k(M) : d\omega = 0\},$$

$$B^k(M) = (\text{exact elements of } \Omega^k(M)) = \{\omega \in \Omega^k(M) : \omega \in d(\Omega^{k-1}(M)\}.$$

Since $d \circ d = 0$, $B^k(M) \subseteq Z^k(M)$ and we can form the quotient space.

Definition. The *de Rham cohomology* of M of dimension k is the quotient space

$$H^k_{dR}(M; \mathbb{R}) = \frac{Z^k(M)}{B^k(M)}.$$

If $\omega \in Z^k(M)$, we let $[\omega]$ denote its cohomology class in $H^k_{dR}(M; \mathbb{R})$.

Note that by construction, de Rham cohomology is an invariant of the smooth manifold M.

Example 1. If M is a smooth manifold with finitely many connected components, then $Z^0(M)$ is just the space of functions which are constant on each component, while $B^0(M) = 0$, so

$$H^0_{dR}(M; \mathbb{R}) \cong \overbrace{\mathbb{R} \oplus \cdots \oplus \mathbb{R}}^{k},$$

where k is the number of components of M.

Example 2. The reader may recall the method of exact differentials for solving differential equations. This method states that if

$$\omega = Mdx + Ndy \in \Omega^1(\mathbb{R}^2) \text{ satisfies the integrability condition } \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 0,$$

then $\omega = df$ for some smooth real-valued function f on $\mathbb{R}^2$. In that case,

$$f = c \quad \text{is a solution to the DE} \quad Mdx + Ndy = 0.$$

The integrability condition is simply the statement that ω is closed, and the fact that this implies that ω is exact is simply the assertion that $H^1_{dR}(\mathbb{R}^2; \mathbb{R}) = 0$, a special case of the Poincaré Lemma to be presented in the next section.

Suppose on the other hand, that $M = \mathbb{R}^2 - \{(0,0)\}$ and

$$\omega = \frac{xdy - ydx}{x^2 + y^2} \in \Omega^1(M).$$

Then a straightforward calculation shows that ω is closed, but if we let S^1 be the unit circle $x^2 + y^2 = 1$ with counterclockwise parametrization, we find that

$$\int_{S^1} \omega = 2\pi.$$

Thus it follows from Stokes' Theorem that ω cannot be exact, and $H^1_{dR}(M; \mathbb{R})$ is nonzero. The one-dimensional de Rham cohomology detects the hole that is missing at the origin.

Example 3. Let $(M, \langle \cdot, \cdot \rangle)$ be an n-dimensional oriented Riemannian manifold, $(x^1, \ldots, x^n)$ a positively oriented coordinate system defined on an open subset U of M. If

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j \quad \text{and} \quad g = \det(g_{ij}),$$

we define the *volume form* on U to be

$$\omega_U = \sqrt{g} dx^1 \wedge \cdots dx^n.$$

A straightforward calculation shows that the volume forms for two different positively oriented coordinate systems agree on overlaps, and hence the locally defined volume forms fit together to yield a global volume form $\omega \in \Omega^n(M)$. If M is compact and has empty boundary, one of its basic invariants is

$$\text{Volume of } M = \int_M \omega$$

Since the volume form ω has degree n, it must be closed, but it cannot be exact by Stokes's Theorem. Hence $H^n_{dR}(M; \mathbb{R}) \neq 0$.

We have seen that if M is a smooth manifold, the exterior derivative yields a sequence of vector spaces and linear maps

$$\xrightarrow{d} \Omega^k(M) \xrightarrow{d} \Omega^{k+1}(M) \xrightarrow{d} \Omega^{k+2}(M) \xrightarrow{d} ,$$

the linear maps satisfying the identity $d \circ d = 0$. This is called the de Rham *cochain complex*, and is denoted by $\Omega^*(M)$. A smooth map $F : M \to N$ induces a commutative ladder

$$\begin{array}{ccccccc} \xrightarrow{d} & \Omega^k(N) & \xrightarrow{d} & \Omega^{k+1}(N) & \xrightarrow{d} & \Omega^{k+2}(N) & \xrightarrow{d} \\ & {\scriptstyle F^*}\big\downarrow & & {\scriptstyle F^*}\big\downarrow & & {\scriptstyle F^*}\big\downarrow & \\ \xrightarrow{d} & \Omega^k(M) & \xrightarrow{d} & \Omega^{k+1}(M) & \xrightarrow{d} & \Omega^{k+2}(M) & \xrightarrow{d} \end{array} \tag{2.11}$$

which can be regarded as a homomorphism of cochain complexes, and denoted by $F^* : \Omega^*(N) \to \Omega^*(M)$. It follows from the commutativity of (2.26) that the smooth map F induces a vector space homomorphism

$$F^* : H^k_{dR}(N; \mathbb{R}) \to H^k_{dR}(M; \mathbb{R}), \quad \text{for each } k.$$

The direct sum

$$H^*_{dR}(M; \mathbb{R}) = \sum_{k=0}^{\infty} H^k_{dR}(M; \mathbb{R})$$

can be made into a graded commutative algebra over $\mathbb{R}$, the product being the so-called *cup product*, which is defined by

$$[\omega] \cup [\phi] = [\omega \wedge \phi].$$

If $F : M \to N$ is a smooth map, the linear map on cohomology

$$F^* : H^*_{dR}(N; \mathbb{R}) \longrightarrow H^*_{dR}(M; \mathbb{R}) \quad \text{respects the cup product:}$$
$$F^*([\omega] \cup [\phi]) = F^*[\omega] \cup F^*[\phi].$$

Moreover, the identity map on M induces the identity on de Rham cohomology and if $F : M \to N$ and $G : N \to P$ are smooth maps, then $(G \circ F)^* = F^* \circ G^*$, so we can say that

$$M \mapsto H^*_{dR}(M; \mathbb{R}), \qquad (F : M \to N) \mapsto (F^* : H^*_{dR}(N; \mathbb{R}) \to H^*_{dR}(M; \mathbb{R}))$$

is a contravariant functor from the category of smooth manifolds and smooth maps to the category of $\mathbb{R}$-algebras and $\mathbb{R}$-homomorphisms.

2.6 Poincaré Lemma

In order for de Rham cohomology to be useful in solving topological problems, we need to be able to compute it in important cases. As a first step in this direction, we might try to prove the so-called Poincaré Lemma:

Poincaré Lemma. *If U is a convex open subset of $\mathbb{R}^n$, then the de Rham cohomology of U is trivial:*

$$H^k_{dR}(U;\mathbb{R}) \cong \begin{cases} \mathbb{R} \text{ if } k = 0, \\ 0 \text{ if } k \neq 0. \end{cases}$$

However, it turns out to be not much more difficult to prove a very powerful fact about de Rham cohomology: de Rham cohomology is invariant under smooth homotopy.

If $F, G : M \to N$ are smooth maps, we say that they are *smoothly homotopic* if there is a smooth map $H : [0,1] \times M \to N$ such that

$$H(0,p) = F(p) \quad \text{and} \quad H(1,p) = G(p).$$

Homotopy Theorem. *Smoothly homotopic maps $F, G : M \to N$ induce the same map on cohomology,*

$$F^* = G^* : H^k_{dR}(N;\mathbb{R}) \longrightarrow H^k_{dR}(M;\mathbb{R}).$$

To see how the Homotopy Theorem implies the Poincaré Lemma, we first note that if $\{p_0\}$ is a single point, regarded as a zero-dimensional manifold, then

$$H^k_{dR}(\{p_0\};\mathbb{R}) \cong \begin{cases} \mathbb{R} \text{ if } k = 0, \\ 0 \text{ if } k \neq 0. \end{cases}$$

Next observe that if p lies in U, then we have an inclusion map $\iota : \{p_0\} \to U$ and a map $r : U \to \{p_0\}$ such that $r \circ \iota = \mathrm{id}_{\{p_0\}}$. On the other hand, $\iota \circ r$ is homotopic to the identity on U: The map $H : [0,1] \times U \to U$ defined by

$$H(t,p) = tp_0 + (1-t)p \quad \text{satisfies} \quad H(0,p) = p, \quad H(1,p) = p_0.$$

Thus functoriality implies that

$$i^* : H^*_{dR}(U;\mathbb{R}) \to H^*_{dR}(\{p_0\};\mathbb{R}) \quad \text{and} \quad r^* : H^*_{dR}(\{p_0\};\mathbb{R}) \to H^*_{dR}(U;\mathbb{R})$$

are both isomorphisms.

Proof of Homotopy Theorem: We only sketch the key ideas; the reader can refer to § 7 of Chapter VI of [5] for additional details.

The Homotopy Theorem follows from the special case for the inclusion maps

$$i_0, i_1 : M \longrightarrow [0,1] \times M, \qquad i_0(p) = (0,p), \quad i_1(p) = (1,p).$$

Indeed, if $H : [0,1] \times M \to N$ is a smooth homotopy from F to G, then by definition of homotopy, $F = H \circ i_0$ and $G = H \circ i_1$, so

$$i_0^* = i_1^* \Rightarrow F^* = i_0^* \circ H^* = i_1^* \circ H^* = G^*.$$

This special case, however, can be established by integrating over the fiber of the projection on the second factor $[0,1] \times M \to M$. More precisely, let t be the standard coordinate on $[0,1]$, T the vector field tangent to the fiber of $[0,1] \times M$ such that $dt(T) = 1$. We then define integration over the fiber

$$\pi_* : \Omega^k([0,1] \times M) \to \Omega^{k-1}(M) \quad \text{by} \quad \pi_*(\omega)(p) = \int_0^1 (\iota_T \omega)(t,p)dt,$$

where we define the interior product $(\iota_T\omega)(t,p)$ as an element of $\Lambda^k T^*_{(t,p)}([0,1] \times M)$ by the formula

$$(\iota_T\omega)(t,p)(v_1,\ldots,v_{k-1}) = \omega(t,p)(T(t,p),v_1,\ldots v_{k-1}), \qquad \text{for} \quad v_1,\ldots,v_{k-1} \in T_{(t,p)}([0,1]\times M).$$

The integration is possible because the exterior power at (t,p) is canonically isomorphic to $\Lambda^k T^*_{(0,p)}([0,1] \times M)$. The key to proving that $i_0^* = i_1^*$ in cohomology is the "cochain homotopy" formula

$$i_1^*\omega - i_0^*\omega = d(\pi_*(\omega)) + \pi_*(d\omega). \tag{2.12}$$

It follows from this formula that if $d\omega = 0$,

$$i_1^*\omega - i_0^*\omega = d(\pi_*(\omega)) \Rightarrow [i_1^*\omega] = [i_0^*\omega],$$

and hence on the cohomology level $i_0^* = i_1^*$.

Thus to finish the proof of the Homotopy Lemma, it remains only to establish (2.12). Let $\{U_\alpha, \phi_\alpha) : \alpha \in A\}$ be an open cover of M by coordinate neighborhoods. It suffices to show that the coordinate representatives $(\pi_*)_\alpha : \Omega^k(U_\alpha \times [0,1]) \to \Omega^k(U_\alpha)$ satisfy

$$i_1^*\omega - i_0^*\omega = d((\pi_*)_\alpha(\omega)) + (\pi_*)_\alpha(d\omega).$$

Thus let $(U, (x^1,\ldots,x^n)$ be one of the smooth maps in the atlas, so that $(t, x^1, \ldots, x^n)$ are smooth coordinates on $U \times [0,1]$. Letting x stand for the n-tuple, we see that $\pi_* : \Omega^k(U \times [0,1]) \to \Omega^k(U)$ satisfies

$$\pi_*\left(f(t,x)dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) = 0, \tag{2.13}$$

$$\pi_*\left(f(t,x)dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}\right) = \left[\int_0^1 f(t,u)du\right] dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}. \tag{2.14}$$

Since any k-from is a superposition of differential forms treated by (2.13) and (2.14), we need only verify the identity (2.12) in each of these two cases.

In the first case (2.13), $\pi_*(\omega) = 0$ so $d(\pi_*(\omega)) = 0$. On the other hand,

$$d\omega = \frac{\partial f}{\partial t} dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} + (\text{terms not involving } dt),$$

so

$$\pi_*(d\omega) = \left[\int_0^1 \frac{\partial f}{\partial u}(u,x)du\right] dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$
$$= f(1,x)dx^{i_1} \wedge \cdots \wedge dx^{i_k} - f(0,x)dx^{i_1} \wedge \cdots \wedge dx^{i_k} = i_1^*(\omega) - i_0^*(\omega),$$

so the formula is established in this case.

In the other case (2.14),

$$i_1^*(dt) = i_0^*(dt) = 0 \quad \Rightarrow \quad i_1^*(\omega) = i_0^*(\omega) = 0. \tag{2.15}$$

On the one hand,

$$d\omega = \sum_{j=1}^n \frac{\partial f}{\partial x^j} dx^j \wedge dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}$$
$$= -\sum_{j=1}^n \frac{\partial f}{\partial x^j} dt \wedge dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}$$

and hence

$$\pi_*(d\omega) = -\sum_{j=1}^n \left[\int_0^1 \frac{\partial f}{\partial x^j}(u,x)du\right] dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}. \tag{2.16}$$

On the other hand,

$$\pi_*(\omega) = \left[\int_0^1 f(u,x)du\right] dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}},$$

so

$$d\pi_*(\omega) = \sum_{j=1}^n \frac{\partial}{\partial x^j} \left[\int_0^1 f(u,x)du\right] dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}$$
$$= \sum_{j=1}^n \left[\int_0^1 \frac{\partial f}{\partial x^j}(u,x)du\right] dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}. \tag{2.17}$$

In view of (2.15), the desired identity now follows by adding (2.16) and (2.17).

Definition. We say that two smooth manifolds M and N are *smoothly homotopic equivalent* if there exist smooth maps $F : M \to N$ and $G : N \to M$ such that $G \circ F$ and $F \circ G$ are both homotopic to the identity.

For example, the cylinder is smoothly homotopic to a circle. It follows from the Homotopy Theorem that if two manifolds are smoothly homotopic equivalent, they have the same de Rham cohomology.

To provide an application of the Homotopy Theorem, we consider

$$D^n = \{(x^1, \dots, x^n) \in \mathbb{R}^n : (x^1)^2 + \cdots + (x^n)^2 \leq 1\},$$
$$S^{n-1} = \{(x^1, \dots, x^n) \in \mathbb{R}^n : (x^1)^2 + \cdots + (x^n)^2 = 1\}.$$

Lemma. *If $n \geq 2$, there does not exist a smooth map $F : D^n \to S^{n-1}$ which leaves $S^{n=1}$ pointwise fixed.*

Proof: Suppose that there were such a map F, and let $p_0 = F(0)$. Define

$$H : [0,1] \times S^{n-1} \to S^{n-1} \quad \text{by} \quad H(t,p) = F(tp).$$

Then H is a smooth homotopy from c to id where c is the constant map which takes S^{n-1} to p_0 and id is the identity map of S^{n-1}. By the Homotopy Theorem,

$$c^* = \mathrm{id}^* : H^{n-1}_{dR}(S^{n-1}; \mathbb{R}) \to H^{n-1}_{dR}(S^{n-1}; \mathbb{R}).$$

If ω is a smooth n-form on D^n, $c^*(\omega) = 0$, and hence c^* is the zero map. On the other hand, it follows from Stokes's Theorem that $H^{n-1}_{dR}(S^{n-1}; \mathbb{R}) \neq 0$, so $\mathrm{Id}^* = \mathrm{id}$ is not the zero map, a contradiction.

Proposition. *A smooth map $G : D^n \to D^n$ must have a fixed point.*

Proof: If $G : D^n \to D^n$ is a smooth map with no fixed point, define $FD^n \to S^{n-1}$ as follows: For $p \in D^n$, let $L(p)$ denote the line trhough p and $G(p)$ and let $F(p)$ be the point on $S^{n-1} \cap L(p)$ closer to p than $G(p)$. Since G is smooth, F is also smooth. Moreover, F leaves S^{n-1} pointwise fixed, contradicting the previous lemma.

Brouwer Fixed Point Theorem. *A continuous map $f : D^n \to D^n$ must have a fixed point.*

Proof: Suppose that $f : D^n \to D^n$ is a continuous map and $\epsilon > 0$ is given. By the Weierstrass approximation theorem, there is a smooth map $PD^n \to \mathbb{R}^n$ such that $|P(p) - f(p)| < \epsilon$ for $p \in D^n$. let

$$G = \frac{1}{1+\epsilon} P.$$

Then $G : D^n \to D^n$ is a smooth map which satisfies $f - G| < 2\epsilon$.

Suppose now that $f : D^n \to D^n$ is a continuous map without fixed points and let

$$\mu = \inf\{|f(p) - p| : p \in D^n\}.$$

By the argument in the preceding paragraph, we can choose a smooth map $G : D^n \to D^n$ such that $|f - G| < \mu$. Then G is a smooth map without fixed points, contradicting the preceding proposition.

Exercise VII. a. Suppose that X is a smooth vector field on M. Define the *interior product* $\iota_X : \Omega^k(M) \to \Omega^{k-1}(M)$ by

$$\iota_X(\omega)(Y_1, \dots, Y_{k-1}) = \omega(X, Y_1, \dots, Y_{k-1}).$$

Show that if $\omega \in \Omega^k(M)$ and $\theta \in \Omega^l(M)$, then

$$\iota_X(\omega \wedge \theta) = (\iota_X \omega) \wedge \theta + (-1)^k \omega \wedge (\iota_X \theta).$$

b. Define a real linear operator $L_X : \Omega^k(M) \to \Omega^k(M)$ by

$$L_X = d \circ \iota_X + \iota_X \circ d.$$

Show that if $\omega \in \Omega^k(M)$ and $\theta \in \Omega^l(M)$, then

$$L_X(\omega \wedge \theta) = (L_X \omega) \wedge \theta + \omega \wedge (L_X \theta).$$

We call L_X the *Lie derivative* in the direction of X.

2.7 Mayer-Vietoris Sequence

In order to be able to calculate de Rham cohomology effectively, we need one further property of de Rham cohomology, the exactness of the Mayer-Vietoris sequence. This enables us to calculate the de Rham cohomology of a smooth manifold by dividing it up into simpler pieces.

Suppose that U and V are open subsets of a smooth manifold M such that $M = U \cup V$. We then have a diagram of inclusion maps:

$$\begin{tikzcd} U \cap V \arrow[r, "j_U"] \arrow[d, "j_V"'] & U \arrow[d, "i_U"] \\ V \arrow[r, "i_V"] & M \end{tikzcd}$$

We can therefore construct a sequence of vector spaces and linear maps

$$0 \to \Omega^k(M) \xrightarrow{i^*} \Omega^k(U) \oplus \Omega^k(V) \xrightarrow{j^*} \Omega^k(U \cap V) \to 0, \tag{2.18}$$

where

$$i^*(\omega) = (i_U^*(\omega), i_V^*(\omega)), \qquad j^*(\phi_U, \phi_V) = j_U^*(\phi_U) - j_V^*(\phi_V).$$

Since i^* and j^* commute with d, (2.18) yields a sequence of cochain complexes

$$0 \to \Omega^*(M) \to \Omega^*(U) \oplus \Omega^*(V) \to \Omega^*(U \cap V) \to 0. \tag{2.19}$$

Lemma. *The sequence (2.19) is exact; in other words for each k, the sequence (2.18) is exact.*

One easily checks that i^* is injective and $j^* \circ i^* = 0$. If $j^*(\phi_U, \phi_V) = 0$, then $j_U^*(\phi_U) = j_V^*(\phi_V)$ so $j_U^*(\phi_U)$ and $j_V^*(\phi_V)$ fit together to form a smooth form on M such that $i^*(\omega) = (\phi_U, \phi_V)$.

The only difficult step in the proof is to show that j^* is surjective. To do this, we choose a partition of unity $\{\psi_U, \psi_V\}$ subordinate to the open cover $\{U, V\}$ of M. If $\theta \in \Omega^k(U \cap V)$, we define $\phi_U \in \Omega^k(U)$ by

$$\phi_U(p) = \begin{cases} \psi_V(p)\theta(p), & \text{for } p \in U \cap V, \\ 0, & \text{for } p \in U - (U \cap V), \end{cases}$$

and $\phi_V \in \Omega^k(V)$ by

$$\phi_V(p) = \begin{cases} -\psi_U(p)\theta(p), & \text{for } p \in U \cap V, \\ 0, & \text{for } p \in V - (U \cap V). \end{cases}$$

then

$$j^*(\phi_U, \phi_V) = \phi_U|(U \cap V) - \phi_V|(U \cap V) = \psi_V\theta + \psi_U\theta = \theta,$$

so j^* is surjective and the lemma is proven.

The lemma yields a large commutative diagram in which the rows are exact:

$$\begin{array}{ccccc} \downarrow d & & \downarrow d & & \downarrow d \\ 0 \to \Omega^k(M) & \xrightarrow{i^*} & \Omega^k(U) \oplus \Omega^k(V) & \xrightarrow{j^*} & \Omega^k(U \cap V) \to 0 \\ \downarrow d & & \downarrow d & & \downarrow d \\ 0 \to \Omega^{k+1}(M) & \xrightarrow{i^*} & \Omega^{k+1}(U) \oplus \Omega^{k+1}(V) & \xrightarrow{j^*} & \Omega^{k+1}(U \cap V) \to 0 \\ \downarrow d & & \downarrow d & & \downarrow d \end{array} \tag{2.20}$$

Note that i^* and j^* induce homomorphisms on cohomology

$$\begin{aligned} i^* &: H^k_{dR}(M; \mathbb{R}) \to H^k_{dR}(U; \mathbb{R}) \oplus H^k_{dR}(V; \mathbb{R}), \\ j^* &: H^k_{dR}(U; \mathbb{R}) \oplus H^k_{dR}(V; \mathbb{R}) \to H^k_{dR}(U \cap V; \mathbb{R}). \end{aligned} \tag{2.21}$$

The commuting diagram (2.20) allows us to construct a "connecting homomorphism"

$$\Delta : H^k_{dR}(U \cap V; \mathbb{R}) \to H^{k+1}_{dR}(M; \mathbb{R}) \tag{2.22}$$

as follows: If $[\theta] \in H^k_{dR}(U \cap V; \mathbb{R})$, choose a representative $\theta \in \Omega^k(U \cap V)$. Since j^* is surjectve, we can choose $\phi \in \Omega^k(U) \oplus \Omega^k(V)$ so that $j^*(\phi) = \theta$. Then $j^*(d\phi) = dj^*(\phi) = d\theta = 0$, so there is a unique $\omega \in \Omega^{k+1}(M)$ such that $i^*\omega = d\phi$. Finally, $i^*(d\omega) = d(i^*\omega) = d(d\phi) = 0$, and since i^* is injective, $d\omega = 0$. Let $[\omega]$ be the de Rham cohomology class of ω in $H^{k+1}_{dR}(M; \mathbb{R})$ and set $\Delta([\theta]) = [\omega]$. Roughly speaking

$$\Delta = (i^*)^{-1} \circ d \circ (j^*)^{-1}.$$

By the technique of "diagram chasing", one checks that $\Delta([\theta])$ is independent of the choice of ϕ, or of θ representing $[\theta]$.

We can describe Δ explicitly as follows: If θ is a representative of $[\theta] \in H^k_{dR}(U \cap V; \mathbb{R})$, then $\Delta([\theta])$ is represented by the form

$$d(\psi_V \theta) = d\psi_V \wedge \theta \quad \text{or} \quad - d(\psi_U \theta) = -d\psi_U \wedge \theta,$$

two expressions for the same $(k+1)$-form on M (since $\psi_U + \psi_V = 1$) which actually has its support in $U \cap V$.

Mayer-Vietoris Theorem. *The homomorphisms (2.21) and (2.22) fit together to form a long exact sequence*

$$\begin{aligned} \cdots \to H^k_{dR}(M;\mathbb{R}) \to H^k_{dR}(U;\mathbb{R}) \oplus H^k_{dR}(V;\mathbb{R}) \to H^k_{dR}(U \cap V;\mathbb{R}) \\ \to H^{k+1}_{dR}(M;\mathbb{R}) \to H^{k+1}_{dR}(U;\mathbb{R}) \oplus H^{k+1}_{dR}(V;\mathbb{R}) \to \cdots . \end{aligned} \tag{2.23}$$

This exact sequence is called the *Mayer-Vietoris sequence*. Together with the Homotopy Lemma, the Mayer-Vietoris sequence is very helpful in computing the de Rham cohomology.

The proof of exactness of the Mayer-Vietoris sequence follows from the so-called "snake lemma" from algebraic topology: A short exact sequence of cochain complexes such as (2.19) gives rise to a long exact sequence in cohomology such as (2.23).

To prove the snake lemma one must establish three assertions:

1. $\mathrm{Ker}(j^*) = \mathrm{Im}(i^*)$,
2. $\mathrm{Ker}(\Delta) = \mathrm{Im}(j^*)$, and
3. $\mathrm{Ker}(i^*) = \mathrm{Im}(\Delta)$.

Each of these assertions is proven by a diagram chase using the diagram (2.20).

For example, suppose we want to check the second of these assertions, $\mathrm{Ker}(\Delta) = \mathrm{Im}(j^*)$. To see that $\mathrm{Im}(j^*) \subseteq \mathrm{Ker}(\Delta)$, we suppose $[\theta] \in \mathrm{Im}(j^*)$, so $[\theta] = j^*[\phi]$ for some $[\phi] \in H^k \oplus H^k(V)$. Then there are representatives θ and ϕ such that $j^*\phi = \theta$. Note that $d\phi = 0$. Hence

$$\Delta([\theta]) = [(i^*)^{-1} \circ d \circ (j^*)^{-1}(\theta)] = 0.$$

Conversely, suppose that $[\theta] \in \mathrm{Ker}(\Delta)$, so $(i^*)^{-1} \circ d \circ (j^*)^{-1}(\theta)$ is exact, so if $\phi \in (j^*)^{-1}(\theta)$, then $(i^*)^{-1} \circ d\phi = d\omega$, for some $\omega \in \Omega^k(M)$. Then $d(\phi - i^*\omega) = 0$ and $j^*(\phi - i^*\omega) = \theta$. Thus $[\theta] \in \mathrm{Im}(j^*)$.

The other two assertions are proven by similar arguments.

Example 1. We first calculate the cohomology of $S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, noting that $H^0(S^1)$ is isomorphic to the space $\mathbb{R}$ of constant functions on S^1. We decompose S^1 into a union of two open sets

$$U = S^1 \cap \left\{ y > -\frac{1}{2} \right\}, \qquad V = S^1 \cap \left\{ y < \frac{1}{2} \right\},$$

and note that U and V are both diffeomorphic to open intervals, while $U \cap V$ is diffeomorphic to the union of two open intervals, so the cohomologies of these spaces can be calculated from the Poincaré Lemma. It follows from the Mayer-Vietoris sequence that

$$\begin{aligned} 0 \to H^0(S^1) \to H^0(U) \oplus H^0(V) \to H^0(U \cap V) \\ \to H^1(S^1) \to H^1(U) \oplus H^1(V) \to H^1(U \cap V) \to \cdots, \end{aligned}$$

which yields

$$0 \to \mathbb{R} \to \mathbb{R} \oplus \mathbb{R} \to \mathbb{R} \to H^1(S^1) \to 0 \to 0 \to \cdots.$$

It follows that

$$H^k_{dR}(S^1; \mathbb{R}) \cong \begin{cases} \mathbb{R}, & \text{if } k = 0 \text{ or } k = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 2. We next consider $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Once again $H^0(S^2) \cong \mathbb{R}$. This time we decompose S^2 into a union of

$$U = S^2 \cap \left\{ z > -\frac{1}{2} \right\} \quad \text{and} \quad V = S^2 \cap \left\{ z < \frac{1}{2} \right\}.$$

In this case, U and V are both diffeomorphic to open disks, while $U \cap V$ is homotopy equivalent to S^1. It follows from the Mayer-Vietoris sequence that

$$\begin{aligned} 0 \to H^0(S^2) \to H^0(U) \oplus H^0(V) \to H^0(U \cap V) \\ \to H^1(S^2) \to H^1(U) \oplus H^1(V) \to H^1(U \cap V) \\ \to H^2(S^2) \to H^2(U) \oplus H^2(V) \to H^2(U \cap V) \to \cdots, \end{aligned}$$

which yields

$$0 \to \mathbb{R} \to \mathbb{R} \oplus \mathbb{R} \to \mathbb{R} \to H^1(S^2) \to 0 \to \mathbb{R} \to H^2(S^2) \to 0 \to \cdots.$$

It follows that

$$H^k_{dR}(S^2; \mathbb{R}) \cong \begin{cases} \mathbb{R}, & \text{if } k = 0 \text{ or } k = 2, \\ 0, & \text{otherwise.} \end{cases}$$

Example 3. By induction, one can calculate the cohomology of S^n:

$$H^k_{dR}(S^n; \mathbb{R}) \cong \begin{cases} \mathbb{R}, & \text{if } k = 0 \text{ or } k = n, \\ 0, & \text{otherwise.} \end{cases}$$

Exercise VIII. Use the Mayer-Vietoris sequence to determine the de Rham cohomology of the two-sphere Σ_g with g handles, the compact oriented surface of genus g. Hint: Use the fact that Σ_g is orientable and therefore has a volume form which makes the top cohomology nonzero.

2.8 Singular homology*

For manifolds, perhaps the easiest entry point for homology and cohomology is the de Rham theory, which we have just discussed. However, it does not apply to general topological spaces. Moreover, the construction of singular homology and cohomology is in many ways even simpler, and gives more refined invariants. For a compact manifold, the singular homology groups are finitely generated abelian groups, which may contain torsion summands, such as $\mathbb{Z}_2$; this torsion is invisible in de Rham theory. An excellent extended treatment of singular homology is given in [14], and we will describe only the very basic ideas here.

2.8.1 Definition of singular homology*

Let $\mathcal{T}$ be the category whose objects are topological spaces and whose morphisms are continuous maps, and let $\mathcal{A}$ be the category whose objects are abelian groups and whose morphisms are group homomorphisms. Singular homology of degree n is a *covariant functor* from $\mathcal{T}$ to $\mathcal{G}$. This means that singular homology assigns to each topological space X an abelian group $H_n(X)$ and to each continuous map $F : X \to Y$ a group homomorphism

$$F_* : H_n(X) \to H_n(Y),$$

such that the identity map id_X on X induces the identity homomorphism on $H_k(X)$, and whenever $F : X \to Y$ and $G : Y \to Z$ are two continuous maps, then

$$(G \circ F)_* = G_* \circ F_* : H_n(X) \to H_n(Z).$$

We begin the construction of singular homology by defining the *standard n-simplex* to be

$$\Delta^n = \{(t^0, \dots t^n) \in \mathbb{R}^{n+1} : \sum t^i = 1, t^i \geq 0\}.$$

It possesses standard *face maps*

$$\delta_i : \Delta^{n-1} \to \Delta^n, \quad \delta_i(t^0, \dots t^{n-1}) = (t_0, \dots t^{i-1}, 0, t_i, \dots, t^{n-1}) \tag{2.24}$$

for $0 \leq i \leq n$. If X is a topological space, a *singular n-simplex in X* is a continuous map $f : \Delta^n \to X$, and we let

$$S_n(X) = \{ \text{ singular } n\text{-simplices in } X \ \}.$$

We can then define *face operators*

$$\partial_i : S_n(X) \to S_{n-1}(X) \quad \text{by} \quad \partial_i(f) = f \circ \delta_i.$$

It is straightforward to check that the face operators satisfy the identities

$$\partial_i \circ \partial_j = \partial_{j-1} \circ \partial_i \quad \text{if } i < j. \tag{2.25}$$

We now let $C_n(X)$ be the free abelian group generated by the space $S_n(X)$ of singular simplices. Thus elements of the abelian group $C_n(X)$ are simply finite sums

$$a_1 f_1 + \cdots + a_k f_k,$$

where $a_1, \ldots a_k$ are elements of $\mathbb{Z}$. We can define a *boundary operator*

$$\partial : C_n(X) \to C_{n-1}(X) \quad \text{by} \quad \partial(f) = \sum_{i=0}^{n} (-1)^i \partial_i(f).$$

It then follows from (2.25) that $\partial \circ \partial = 0$. We thus obtain a sequence of groups and group homomorphisms

$$\xrightarrow{\partial} C_n(X) \xrightarrow{\partial} C_{n-1}(X) \xrightarrow{\partial} C_{n-2}(X) \xrightarrow{\partial} \tag{2.26}$$

which is called a *chain complex* over $\mathbb{Z}$, and is denoted by $C_*(X)$.

A continuous map $F : X \to Y$ between topological spaces induces a map

$$F_* : S_n(X) \to S_n(Y) \quad \text{by} \quad F_*(f) = F \circ f : \Delta^n \to Y,$$

a map which commutes with the face operators. This map defines a group homomorphism $F_* : C_n(X) \to C_n(Y)$ which commutes with the boundary operator, and hence induces a commutative ladder

$$\begin{array}{ccccccc}
\xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & C_{n-1}(X) & \xrightarrow{\partial} & C_{n-2}(X) & \xrightarrow{\partial} \\
 & \downarrow{\scriptstyle F_*} & & \downarrow{\scriptstyle F_*} & & \downarrow{\scriptstyle F_*} & \\
\xrightarrow{\partial} & C_n(Y) & \xrightarrow{\partial} & C_{n-1}(Y) & \xrightarrow{\partial} & C_{n-2}(Y) & \xrightarrow{\partial}
\end{array} \tag{2.27}$$

which is called a *chain map* and regarded as a morphism in the category of chain complexes. We write $F_* : C_*(X) \to C_*(Y)$. The correspondence

$$X \mapsto C_*(X), \quad (F : X \to Y) \mapsto (F_* : C_*(X) \to C_*(Y))$$

satisfies the two identities

$$(\mathrm{id})_* = \mathrm{id}, \qquad (G \circ F)_* = G_* \circ F_*,$$

so we say it defines a *covariant functor* from the the category of topological spaces and continuous maps to the category of chain complexes and chain maps.

Given any chain complex $C_*(X)$, we define the corresponding space of n-*cycles* to be

$$Z_n(X) = \{c \in C_n(X) : \partial c = 0\},$$

and the space of n-*boundaries* to be

$$B_n(X) = \{c \in C_n(X) : c = \partial d, \text{ for some } d \in C_{n+1}(X)\ \}.$$

Since $\partial \circ \partial = 0$, $B_n(X) \subseteq Z_n(X)$ so we can form the quotient group

$$H_n(X;\mathbb{Z}) = \frac{Z_n(X)}{B_n(X)},$$

which we call the n-th homology group of the chain complex. When this construction is applied to the chain complex generated by singular simplices, the resulting homology is called the n-th *singular homology group* of X with integer coefficients. If $F : X \to Y$ is a continuous map, the commutative ladder (2.27) shows that F induces a group homomorphism

$$F_* : H_n(X;\mathbb{Z}) \to H_n(Y;\mathbb{Z}).$$

For each $n \in \mathbb{Z}$; we thus obtain a covariant functor from the category of topological spaces and continuous maps to the category of groups and group homomorphisms

$$X \mapsto H_n(X;\mathbb{Z}), \qquad (F : X \to Y) \quad \mapsto \quad (F_* : H_n(X;\mathbb{Z}) \to H_n(Y;\mathbb{Z})).$$

Simple examples. First note that if X is any topological space, $H_0(X;\mathbb{Z})$ is the free abelian group generated by the path components of X. Indeed, the zero simplices of X are just points, and if $p, q \in X$, then $q - p$ is a boundary exactly when p and q are joined by a continuous path. Moreover, we easily check that if $\{p\}$ is a topological space consisting of a single point, then

$$H_k(\{p\};\mathbb{Z}) \cong \begin{cases} \mathbb{Z} & \text{if } k = 0, \\ 0, & \text{otherwise.} \end{cases}$$

More generally, we have an analog of the Poincaré Lemma; the first step towards proving it requires the notion of chain homotopy.

If $F_*, G_* : C_*(X) \to C_*(Y)$ are chain maps, we say that they are *chain homotopic* if for each $n \in \mathbb{Z}$, there are group homomorphisms $D_n : C_n(X) \to C_{n+1}(Y)$ such that $\partial D_n + D_{n-1}\partial = F_* - G_*$. If F_* and G_* are chain homotopic and $z \in Z_n(X)$, then

$$F_*(z) - G_*(z) = (\partial D_n + D_{n-1}\partial)(z) = \partial D(z),$$

so $F_*([z]) = G_*([z])$ on the homology level. In other words, chain homotopic maps induce the same map on homology. The usefulness of chain homotopies is illustrated by the proof of the following proposition:

Proposition. *If X is a convex subset of $\mathbb{R}^n$ with the induced topology, then*

$$H_k(X;\mathbb{Z}) \cong \begin{cases} \mathbb{Z}, & \textit{if } k = 0, \\ 0, & \textit{otherwise.} \end{cases}$$

To prove this, let $s_* : C_*(X) \to C_*(\{p\})$ be the chain map defined by

$$c_n = 0 \quad \text{if } n \neq 0, \qquad c_0 \left(\sum n_x x \right) = \sum n_x p, \quad \text{if } n = 0.$$

Let p_0 be a fixed point of X. If $f : \Delta^n \to X$ is a singular simplex in X, then set

$$Df(t^0, \dots t^{n+1}) = \begin{cases} t^0 p_0 + (1-t^0) f\left(\frac{t^1}{1-t^0}, \dots, \frac{t^{n+1}}{1-t^0} \right) & \text{if } t_0 \neq 1, \\ p_0, & \text{if } t^0 = 1. \end{cases}$$

If $f \in C_0(X)$, then

$$\partial D(f) = f - p_0 = \mathrm{id}_*(f) - s_*(f).$$

More generally, if $f \in C_n(X)$, then

$$\partial D f = f + \sum_{j=1}^{n+1} (-1)^j D(f \circ \delta_{j-1}) = f - \sum_{j=0}^{n} (-1)^j D(f \circ \delta_j),$$

$$D \partial f = f + D \left(\sum_{j=0}^{n} (-1)^j (f \circ \delta_j) \right) = \sum_{j=0}^{n} (-1)^j D(f \circ \delta_j).$$

Hence if $n \geq 1$,

$$\partial D f + D \partial f = f = \mathrm{id}_*(f) - s_*(f).$$

Thus D is a chain homotopy from the identity to the constant map to a point. The statement of the proposition follows from this fact.

Singular homology can be computed for more complicated examples once one has the Homotopy and Mayer-Vietoris Theorems for singular theory, which we now describe.

If $F : X \to Y$ and $G : X \to Y$ are continuous maps, we say that F and G are *homotopic* if there is a continuous map

$$H : [0,1] \times X \to Y \qquad \text{such that} \qquad H(p,0) = F, \quad H(p,1) = G(p).$$

Then just like in de Rham theory, we have:

Homotopy Theorem. *If $F, G : X \to Y$ are homotopic continuous maps, then*

$$F_* = G_* : H_k(X; \mathbb{Z}) \to H_k(Y; \mathbb{Z}).$$

Mayer-Vietoris Theorem. *If U and V are open subsets of a topological space X with $X = U \cup V$, then we have a long exact sequence*

$$\begin{aligned} \cdots \to H_k(U \cap V; \mathbb{Z}) \to H_k(U; \mathbb{Z}) \oplus H_k(V; \mathbb{Z}) \to H_k(M; \mathbb{Z}) \\ \to H_{k-1}(U \cap V; \mathbb{Z}) \to H_{k-1}(U; \mathbb{Z}) \oplus H_{k-1}(V; \mathbb{Z}) \to \cdots . \end{aligned} \tag{2.28}$$

Proofs of these theorems, as well as the calculations for many examples, can be found in [14], as well as other texts on algebraic topology. The important thing for the reader to recognize at this point is the formal similarity between de Rham cohomology and singular homology, except that one is a contravariant functor, the other covariant.

2.8.2 Singular cohomology*

Suppose that G is an abelian group, or more generally, an R-module, where R is a commutative ring with identity. Given a chain complex

$$\longrightarrow C_{n+1} \xrightarrow{\partial} C_n \xrightarrow{\partial} C_{n-1} \longrightarrow,$$

we can construct a corresponding *cochain complex*

$$\longrightarrow \mathrm{Hom}(C_{n-1}, G) \xrightarrow{\delta} \mathrm{Hom}(C_n, G) \xrightarrow{\delta} \mathrm{Hom}(C_{n+1}, G) \longrightarrow,$$

where $\mathrm{Hom}(C_n, G)$ denotes the space of group homomorphisms from C_n to G and

$$\delta(\phi)(c) = \phi(\partial c), \quad \text{for } \phi \in \mathrm{Hom}(C_n, G) \text{ and } c \in C_{n+1}.$$

Note that when G is an R-module, so is $\mathrm{Hom}(C_{n+1}, G)$, and the maps δ are G-module morphisms.

For example, if we apply this construction to the singular chain complex

$$C_*(X) : \longrightarrow C_{n+1}(X) \xrightarrow{\partial} C_n(X) \xrightarrow{\partial} C_{n-1}(X) \longrightarrow$$

of a topological space X, we obtain the complex of *singular cochains* with coefficients in G,

$$C^*(X; G) : \longrightarrow C^{n-1}(X; G) \xrightarrow{\delta} C^n(X; G) \xrightarrow{\delta} C^{n+1}(X; G) \longrightarrow.$$

We call

$$Z^n(X; G) = \mathrm{Ker}(\delta : C^n(X; G) \to C^{n+1}(X; G))$$

the space of *cocycles* and

$$B^n(X; G) = \mathrm{Im}(\delta : C^{n-1}(X; G) \to C^n(X; G))$$

the space of *coboundaries*, and define the *singular cohomology* of degree n with coefficients in G:

$$H^n(X; \mathbb{R}) = \frac{Z^n(X; G)}{B^n(X; G)}.$$

A continuous map $F : X \to Y$ induces a chain homomorphism $F^* : C^*(Y; G) \to C^*(X; G)$, which in turn induces a homomorphism $F^* : H^n(Y; G) \to H^n(X; G)$ for each $n \in \mathbb{Z}$. This gives a contravariant functor

$$X \mapsto H^n(X : G), \qquad (F : X \to Y) \mapsto (F^* : H^n(Y; G) \to H^n(X; G)).$$

In particular, we can consider singular cohomology with coefficients iin the integers $\mathbb{Z}$, or with coefficients in the field $\mathbb{R}$ of real numbers. Moreover, the inclusion $\mathbb{Z} \subseteq \mathbb{R}$ induces cochain homomorphisms $C^*(x;\mathbb{Z}) \to C^*(X;\mathbb{R})$ which induces a "coefficient homomorphism"

$$H^n(X;\mathbb{Z}) \to H^n(X;\mathbb{R}), \quad \text{for each } n \in \mathbb{Z}.$$

Remark. If one develops the skill for calculating cohomology with integer coefficients, one finds that it is slightly more refined, but also a little more difficult to compute. For example, if $M = \mathbb{R}P^2$, one finds that

$$H^k(\mathbb{R}P^2;\mathbb{Z}) \cong \begin{cases} \mathbb{Z} & \text{if } k = 0, \\ \mathbb{Z}_2 & \text{if } k = 2, \\ 0, & \text{otherwise,} \end{cases}, \quad \text{while} \quad H^k_{dR}(\mathbb{R}P^2;\mathbb{R}) \cong \begin{cases} \mathbb{R} & \text{if } k = 0, \\ 0, & \text{otherwise.} \end{cases}$$

With real coefficients, one loses torsion such as $\mathbb{Z}_2$, but that fact often makes it easier to compute.

The Isomorphism Theorem of de Rham. *If M is a smooth manifold, its singular cohomology with coefficients in $\mathbb{R}$ is isomorphic to its de Rham cohomology:*

$$H^n(M;\mathbb{R}) \cong H^n_{dR}(M;\mathbb{R}).$$

The proof of this theorem is somewhat lengthy. It is actually relatively easy to construct a chain map from the de Rham complex $\Omega^*(M)$ to the complex $C^*(M;\mathbb{R})$ of real singular cochains; the difficulty is to show that the map induces an isomorphism on cohomology.

To construct that chain map from $\Omega^*(M)$ to $C^*(M;\mathbb{R})$, we start by supposing that $\omega \in \Omega^n(M)$. We can then define the integral of ω over f to be the real number

$$\langle f, \omega \rangle = \int_{\Delta^n} f^*\omega,$$

where the standard n-simplex Δ^n is given the orientation determined by leaving out the coordinate t^0 and taking the remaining coordinates in the order $(t^1, \ldots, t^n)$. More generally, if $c = a_1 f_1 + \cdots + a_k f_k$ is a singular n-chain, we define the integral of ω over c by

$$\langle c, \omega \rangle = a_1 \langle f_1, \omega \rangle + \cdots + a_k \langle f_k, \omega \rangle.$$

This yields a homomorphism from $C_n(M)$ to $\mathbb{R}$, so any $\omega \in \Omega^n(M)$ defines a corresponding singular cochain $\omega : C_n(M) \to \mathbb{R}$. Thus we have an inclusion $i^n : \Omega^n(M) \subseteq C^n(M;\mathbb{R})$ for each $n \in \mathbb{Z}$ and the question arises as to whether these inclusions fit into a cochain map. This will in fact be the case if and only if the following diagram commutes:

$$\begin{array}{ccc}
\Omega^{n+1}(M) & \xrightarrow{i^{n+1}} & C^{n+1}(M;\mathbb{R}) \\
{\scriptstyle d}\big\uparrow & & {\scriptstyle \delta}\big\uparrow \\
\Omega^{n+1}(M) & \xrightarrow{i^{n}} & C^{n}(M;\mathbb{R}).
\end{array}$$

the commutativity of this diagram is exactly the content of the following:

Stokes' Theorem for Chains. *Suppose that M is a smooth manifold. If $c \in S_n(M)$ and $\theta \in \Omega^{n-1}(M)$, then*

$$\langle c, d\theta\rangle = \langle \partial c, \theta\rangle. \tag{2.29}$$

Once we have Stokes' Theorem for chains, the chain map $i^* : \omega^*(M) \subseteq C^*(M;\mathbb{R})$ defines a homomorphism

$$i^* : H^n_{dR}(M;\mathbb{R}) \longrightarrow H^n(M;\mathbb{R}).$$

One can prove the de Rham Theorem by showing that i^* is an isomorphism on the cohomology level.

As preparation for the proof of (2.29), we first note that if $\sigma \in S_{n+1}$, the symmetric group of bijections from $\{0, 1, \ldots, n\}$ to itself, we can define a linear map $T_\sigma : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ by

$$T_\sigma(x^0, \ldots, x^n) = (x^{\sigma(0)}, \ldots, x^{\sigma(n)}).$$

Then one checks directly that

$$\langle f \circ T_\sigma, \omega\rangle = \operatorname{sgn}(\sigma)\langle f, \omega\rangle. \tag{2.30}$$

Proof of Stokes' Theorem for chains: It suffices to show that

$$\langle f, d\theta\rangle = \langle \partial f, \theta\rangle,$$

when f is an n-simplex; in other words, that

$$\int_{\Delta^n} f^*(d\theta) = \sum_{i=0}^{n}(-1)^i \int_{\Delta^{n-1}} (f \circ \delta_i)^*\theta,$$

or equivalently,

$$\int_{\Delta^n} d(f^*\theta) = \sum_{i=0}^{n}(-1)^i \int_{\Delta^{n-1}} \delta_i^*(f^*\theta).$$

Thus we need only show that if ϕ is a smooth $(n-1)$-form on Δ^n, then

$$\int_{\Delta^n} d\phi = \sum_{i=0}^{n}(-1)^i \int_{\Delta^{n-1}} \delta_i^*\phi.$$

For positively oriented coordinates on Δ^n we take $(t^1, \ldots, t^n)$ on $\mathbb{R}^{n+1}$ and note that on Δ^n,

$$t^0 = 1 - \sum_{i=1}^{n} t^i.$$

Then the $(n-1)$-form ϕ will be a sum of monomials,

$$\phi = \sum_{i=0}^{n} (-1)^{i-1} f^i(t^1, \ldots, t^n) dt^1 \wedge \cdots \wedge dt^{i-1} \wedge dt^{i+1} \wedge \cdots \wedge dt^n,$$

and it would suffice to prove Stokes' Theorem for each of the monomials in the sum. But by applying (2.30), we can reduce any such monomial to the form

$$\phi = f(t^1, \ldots, t^n) dt^2 \wedge \cdots \wedge dt^n,$$

so that

$$d\phi = \frac{\partial f}{\partial t^1}(t^1, \ldots, t^n) dt^1 \wedge \cdots \wedge dt^n.$$

Note that $\delta_j^* \phi = 0$ unless $j = 0$ or $j = 1$.

Denote the restricted coordinates on Δ^{n-1} by $(s^1, \ldots, s^{n-1})$ with

$$s^0 = 1 - \sum_{j=1}^{n-1} s^j.$$

Then

$$t^0 \circ \delta_0 = 0, \quad t^i \circ \delta_0 = s^{i-1} \quad \text{for } 1 \leq i \leq n,$$

so

$$\begin{aligned} \delta_0^* \phi = f(s^0, \ldots s^{n-1}) ds^1 \wedge \cdots \wedge ds^{n-1} \\ = f\left(1 - \sum_{j=1}^{n-1} s^j, \ldots s^{n-1}\right) ds^1 \wedge \cdots \wedge ds^{n-1}. \end{aligned}$$

On the other hand,

$$t^0 \circ \delta_1 = s^0, t^1 \circ \delta_1 = 0, t^2 \circ \delta_1 = s^1 \ldots, \ldots, t^n \circ \delta_i = s^{n-1},$$

so

$$\delta_1^* \phi = f(0, s^2, \ldots, s^{n-1}) ds^1 \wedge \cdots \wedge ds^{n-1}.$$

Thus we can conclude that

$$\int_{\Delta^n} d\phi = \int_0^1 \left[\int_0^{1-t^n} \cdots \left[\int_0^{1-\sum_{i\geq 2} t^i} \frac{df}{dt^1} dt^1 \right] \cdots dt^{p-1} \right] dt^p$$

$$= \int_0^1 \left[\int_0^{1-t^n} \cdots \left[f\left(1 - \sum_{i=2}^{p} t^i, t^2, \ldots t^p\right) - f(0, t^2, \ldots, t^p) \right] \cdots dt^{p-1} \right] dt^p$$

$$= \int_0^1 \left[\int_0^{1-s^{n-1}} \cdots \left[f\left(1 - \sum_{i=1}^{p-1} s^i, s^1, \ldots s^{p-1}\right) \right.\right.$$
$$\left.\left. - f(0, s^1, \ldots, s^{p-1}) \right] \cdots ds^{p-2} \right] ds^{p-1}$$

$$= \int_{\Delta^{p-1}} \delta_0^*(\phi) - \int_{\Delta^{p-1}} \delta_1^*(\phi) = \sum_{i=0}^{n} (-1)^i \int_{\Delta^{p-1}} \delta_i^*(\phi),$$

finishing the proof of the theorem.

2.8.3 Proof of the de Rham Theorem*

As cohomology theory developed, many different definitions of cohomology groups were proposed. In addition to the de Rham and singular cohomologies, yet a third cohomology due Čech was based upon open covers of a topological space. It turns out to be easiest to prove the de Rham isomorphism theorem by showing that both de Rham and singular cohomologies are isomorphic to Čech.

For the simplest definition of Čech cohomology, we need a *good cover* of X, an open cover $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ of a topological space X such that any nonempty intersection $U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq 0$ of open sets in the cover is contractible. It is actually rather difficult to construct good covers for an arbitrary topological space, so the general defintion of Čech cohomology requires taking direct limits over open covers. However, for smooth manifolds there are no such problems. There are two ways of constructing good covers for smooth manifolds.

Method I. We let $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ be a locally finite open cover of $\mathcal{M}$ by sets which are geodesically convex with respect to some Riemannian metric on $\mathcal{M}$. Geodesically convex subsets of a Riemannian manifold are automatically contractible. Since the intersection of geodesically convex sets is geodesically convex, the intersection of any collection of elements from $\mathcal{U}$ is contractible.

Method II. For this method, we need to assume as known the fact that any smooth manifold can be triangulated. Given a vertex v in the triangulation, the *open star* U_v of the vertex is all of the open simplices in the triangulation which contain v in their closure. We start by taking the first barycentric subdivision of a given triangulation. The collection

$$\mathcal{U} = \{U_v : \ v \text{ is a vertex of the barycentric subdivision triangulation } \}$$

is then a good open cover of M. Then any nonempty intersection of $(p+1)$

elements of $\mathcal{U}$ is actually an open p-simplex in the triangulation, hence contractible.

The second method motivates the following definition. By a *p-simplex* of $\mathcal{U}$ we mean an ordered $(p+1)$-tuple $(\alpha_0, \ldots, \alpha_p)$ of indices such that

$$U_{\alpha_0} \cap U_{\alpha_1} \cap \cdots \cap U_{\alpha_p} \neq \emptyset,$$

and we let $\check{\mathrm{S}}_p(\mathcal{U})$ denote the collection of all p-simplices of $\mathcal{U}$. If $\sigma \in \check{\mathrm{S}}_p(\mathcal{U})$, we call

$$|\ \sigma| = U_{\alpha_0} \cap U_{\alpha_1} \cap \cdots \cap U_{\alpha_p}$$

the *support* of σ. If $\sigma = (\alpha_0, \ldots, \alpha_p)$ we let

$$\partial_i \sigma = (\alpha_0, \alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_p),$$

where α_i has been left out.

Finally, if G is an abelian group or an R-module, where R is a commutative ring with identity, we let $\check{\mathrm{C}}^p(\mathcal{U}, G)$ be the set of functions $c : \check{\mathrm{S}}_p(\mathcal{U}) \to G$ such that whenever $\sigma \in S_{p+1}$, the symmetric group on $p+1$ letters,

$$c(\alpha_{\sigma(0)}, \alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(p)}) = (\mathrm{sgn})\sigma c(\alpha_0, \alpha_1, \ldots, \alpha_p) \in G.$$

For simplicity we will often write

$$c_{\alpha_0, \alpha_1, \ldots, \alpha_p} \quad \text{for} \quad c(\alpha_0, \alpha_1, \ldots, \alpha_p).$$

Of course, for the proof of the de Rham Theorem, the important case is that where $G = \mathbb{R}$. We will define the Čech coboundary

$$\delta : \check{\mathrm{C}}^p(\mathcal{U}, G) \to \check{\mathrm{C}}^{p+1}(\mathcal{U}, G) \quad \text{by} \quad \delta(f) = \sum_{i=1}^{p+1} (-1)^i (f \circ \partial_i).$$

One readily verifies that $\delta \circ \delta = 0$, so we obtain the Čech cochain complex

$$C^*(\mathcal{U}, G) : \longrightarrow C^{n-1}(\mathcal{U}, G) \xrightarrow{\ \delta\ } C^n(\mathcal{U}, G) \xrightarrow{\ \delta\ } C^{n+1}(\mathcal{U}, G) \longrightarrow \tag{2.31}$$

The reader who is familiar with simplicial homology and cohomology will recognize that when the open cover is obtained by method II, the Čech cochain complex is isomorphic to the cochain complex of simplicial cohomology with coefficients in G.

The cohomology of the Čech cochain complex (2.31),

$$\check{\mathrm{H}}^p(\mathcal{U}, \mathbb{R}) = \frac{\mathrm{Ker}(\delta : \check{\mathrm{C}}^p(\mathcal{U}, \mathbb{R}) \to \check{\mathrm{C}}^{p+1}(\mathcal{U}, \mathbb{R}))}{\mathrm{Im}(\delta : \check{\mathrm{C}}^{p-1}(\mathcal{U}, \mathbb{R}) \to \check{\mathrm{C}}^p(\mathcal{U}, \mathbb{R})}$$

is called the Čech cohomology of the good covering $\mathcal{U}$, with coefficients in $\mathbb{R}$. For a manifold M, we will see that the Čech cohomology does not depend upon the choice of good cover. When the cover is constructed by method II, it is isomorphic to the simplicial cohomology of M which is isomorphic to the singular cohomology of M by techniques described in algebraic topology texts, such as [14].

We emphasize that for more general topological spaces, it may not be possible to find good covers, and it is more difficult to define Čech cohomology; it must be defined by a taking a direct limit over finer and finer open covers.

Our goal in this section is to show that the de Rham cohomology of a smooth manifold is isomorphic to its Čech cohomology of any good cover. In view of the results from algebraic topology cited above, this gives a proof of de Rham's Isomorphism Theorem.

The idea is to use a generalized Mayer-Vietoris argument as in [7]. We thus construct a double complex $K^{*,*}$ in which the (p,q)-element is

$$K^{p,q} = \check{\mathrm{C}}^p(\mathcal{U}, \Omega^q),$$

which is defined to be the space of functions ω which assign to p-simplex $(\alpha_0, \ldots, \alpha_p)$ a smooth q-form

$$\omega_{\alpha_0 \cdots \alpha_p} \in \Omega^q(U_{\alpha_0} \cap \cdots \cap U_{\alpha_p}).$$

As before, we require that if the order of elements in a sequence is permuted, $\omega_{\alpha_0,\ldots,\alpha_p}$ changes by the sign of the permutation; thus, for example,

$$\omega_{\alpha_0\alpha_1} = -\omega_{\alpha_1\alpha_0}, \quad \omega_{\alpha\alpha} = 0, \qquad \text{and so forth.}$$

We have two differentials on the double complex, the exterior derivative

$$d : \check{\mathrm{C}}^p(\mathcal{U}, \Omega^q) \to \check{\mathrm{C}}^p(\mathcal{U}, \Omega^{q+1}) \quad \text{defined by} \quad (d\omega)_{\alpha_0\cdots\alpha_p} = d\omega_{\alpha_0\cdots\alpha_p},$$

and the *Čech differential*

$$\delta : \check{\mathrm{C}}^p(\mathcal{U}, \Omega^q) \to \check{\mathrm{C}}^{p+1}(\mathcal{U}, \Omega^q)$$

defined by

$$(\delta\omega)_{\alpha_0\cdots\alpha_{p+1}} = \sum_{i=0}^{p+1} (-1)^i \omega_{\alpha_0\cdots\hat{\alpha}_i\cdots\alpha_{p+1}},$$

the forms on the right being restricted to the intersection and the hat indicating that an index is omitted. It is immediately verified that $\delta \circ \delta = 0$.

Remark. The cohomology of the cochain complex

$$\ldots \to \check{\mathrm{C}}^{p-1}(\mathcal{U}, \Omega^q) \to \check{\mathrm{C}}^p(\mathcal{U}, \Omega^q) \to \check{\mathrm{C}}^{p+1}(\mathcal{U}, \Omega^q) \to \cdots$$

is often called *sheaf cohomology* with coefficients in the sheaf of smooth q-forms on M.

The first differential d in the double complex is exact except when $q = 0$ by the Poincaré Lemma, while in the case $q = 0$ we find that

$$[\text{Kernel of } d : \check{\mathrm{C}}^p(\mathcal{U}, \Omega^0) \to \check{\mathrm{C}}^p(\mathcal{U}, \Omega^1)] = \check{\mathrm{C}}^p(\mathcal{U}; \mathbb{R}),$$

the space of Čech cochains for the covering $\mathcal{U}$ on $\mathcal{M}$. The Čech cohomology of the cover $\mathcal{U}$ with coefficients in $\mathbb{R}$ is by definition the cohomology of the cochain complex

$$\cdots \to \check{\mathrm{C}}^{p-1}(\mathcal{U}; \mathbb{R}) \to \check{\mathrm{C}}^p(\mathcal{U}; \mathbb{R}) \to \check{\mathrm{C}}^{p+1}(\mathcal{U}; \mathbb{R}) \to \cdots \tag{2.32}$$

and is denoted by $\check{\mathrm{H}}^p(\mathcal{M}; \mathbb{R})$.

We claim that the second differential is exact except when $p = 0$. It is at this point that we need a partition of unity $\{\psi_\alpha : \alpha \in A\}$ subordinate to $\mathcal{U}$. Given a δ-cocycle $\omega \in \check{\mathrm{C}}^p(\mathcal{U}, \Omega^q)$, we set

$$\tau_{\alpha_1 \cdots \alpha_{p-1}} = \sum_\alpha \psi_\alpha \omega_{\alpha\alpha_1 \cdots \alpha_{p-1}} \in \check{\mathrm{C}}^{p-1}(\mathcal{U}, \Omega^q).$$

Then

$$(\delta\tau)_{\alpha_0 \cdots \alpha_p} = \sum_{i,\alpha} (-1)^i \psi_\alpha \omega_{\alpha\alpha_0 \cdots \hat{\alpha}_i \cdots \alpha_p}$$

and since ω is a cocycle,

$$\delta\omega_{\alpha\alpha_0 \cdots \alpha_p} = \omega_{\alpha_0 \cdots \alpha_p} - \sum (-1)^i \omega_{\alpha\alpha_0 \cdots \hat{\alpha}_i \cdots \alpha_p} = 0.$$

Thus it follows from the fact that $\sum \psi_\alpha = 1$ that

$$(\delta\tau)_{\alpha_0 \cdots \alpha_p} = \sum \psi_\alpha \omega_{\alpha_0 \cdots \alpha_p} = \omega_{\alpha_0 \cdots \alpha_p}.$$

establishing exactness. When $p = 0$, we find that

$$\text{Kernel of } \delta : \check{\mathrm{C}}^0(\mathcal{U}, \Omega^q) \to \check{\mathrm{C}}^1(\mathcal{U}, \Omega^q) = \Omega^q(M),$$

the space of smooth q-forms on $\mathcal{M}$.

We can summarize the previous discussion by stating that the rows and columns in the following commutative diagram are exact:

$$\begin{array}{ccccccccccc}
 & & & & \cdot & & \cdot & & \cdot & & \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & \Omega^2(M) & \to & \check{\mathrm{C}}^0(\mathcal{U}, \Omega^2) & \to & \check{\mathrm{C}}^1(\mathcal{U}, \Omega^2) & \to & \check{\mathrm{C}}^2(\mathcal{U}, \Omega^2) & \to \cdots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & \Omega^1(M) & \to & \check{\mathrm{C}}^0(\mathcal{U}, \Omega^1) & \to & \check{\mathrm{C}}^1(\mathcal{U}, \Omega^1) & \to & \check{\mathrm{C}}^2(\mathcal{U}, \Omega^1) & \to \cdots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & \Omega^0(M) & \to & \check{\mathrm{C}}^0(\mathcal{U}, \Omega^0) & \to & \check{\mathrm{C}}^1(\mathcal{U}, \Omega^0) & \to & \check{\mathrm{C}}^2(\mathcal{U}, \Omega^0) & \to \cdots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & \check{\mathrm{C}}^0(\mathcal{U}, \mathbb{R}) & & \check{\mathrm{C}}^1(\mathcal{U}, \mathbb{R}) & & \check{\mathrm{C}}^2(\mathcal{U}, \mathbb{R}) & & \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & 0 & & 0 & & 0 & &
\end{array}$$

The remainder of the proof uses this diagram. Given a de Rham cohomology class $[\omega] \in H^p_{dR}(\mathcal{M};\mathbb{R})$ with p-form representative ω we construct a corresponding cohomology class $s([\omega])$ in the Čech cohomology $\check{\mathrm{H}}^p(\mathcal{M};\mathbb{R})$ as follows: The differential form defines an element $\omega^{0p} \in \check{\mathrm{C}}^0(\mathcal{U},\Omega^p)$ by simply restricting ω to the sets in the cover. It is readily checked that ω^{0p} is closed with respect to the *total differential* $D = \delta + (-1)^p d$ on the double complex $K^{*,*} = \check{\mathrm{C}}^*(\mathcal{U},\Omega^*)$. Using the Poincaré Lemma, we construct an element $\omega^{0,p-1} \in \check{\mathrm{C}}^0(\mathcal{U},\Omega^{p-1})$ such that $d\omega^{0,p-1} = \omega^{0p}$. Let $\omega^{1,p-1} = \delta\omega^{0,p-1}$ and observe that $d\omega^{1,p-1} = 0$ and $\omega^{1,p-1}$ is cohomologous to ω^{0p} with respect to D. Using the Poincaré Lemma again, we construct an element $\omega^{1,p-2} \in \check{\mathrm{C}}^1(\mathcal{U},\Omega^{p-2})$ such that $d\omega^{1,p-2} = \omega^{1,p-1}$. Let $\omega^{2,p-2} = \delta\omega^{1,p-2}$ and note that $\omega^{2,p-2}$ is cohomologous to ω^{0p} with respect to D. Continue in this fashion until we reach a D-cocycle $\omega^{p0} \in \check{\mathrm{C}}^p(\mathcal{U},\Omega^0)$ which is cohomologous to ω^{0p}. Since $d\omega^{p0} = 0$, each function $\omega^{p0}_{\alpha_0\cdots\alpha_p}$ is constant, and thus ω^{p0} determines a Čech cocycle $s(\omega)$ whose cohomology class is $s([\omega])$.

By the usual diagram chasing, the cohomology class obtained is independent of choices made. Moreover, reversing the zig-zag construction described in the preceding paragraph yields an inverse to s. This finishes our sketch of the proof of the following:

Theorem. *If M is a smooth manifold with a Riemannian metric and $\mathcal{U}$ is a good open cover of M, then the de Rham cohomology of M is isomorphic to the Čech cohomology:*

$$H^n_{dR}(M;\mathbb{R}) \cong \check{H}^n(M;\mathbb{R}).$$

Remark. Establishing exactness of the rows in the double complex is just a generalization of the argument used to prove exactness of the Mayer-Vietoris sequence. Thus Bott and Tu [7] call the above argument a generalized Mayer-Vietoris argument. The argument is useful in many other contexts.

2.9 The Hodge star

In the last few sections, we have described the foundations for the beautiful edifice of algebraic topology. Of course, topology can be justified as a beautiful subject in its own right, but it also has numerous applications to the partial differential equations which arise in geometry. Our next goal is to describe one such application, Hodge's theorem that any de Rham cohomology class contains a unique solution to Laplace's equation.

We begin with an oriented Riemannian or pseudo-Riemannian manifold $(M,\langle,\cdot,\cdot\rangle)$ of dimension n. For each $p \in M$, the nondegenerate symmetric billinear form $\langle\cdot,\cdot\rangle$ defines an isomorphism

$$\flat : T_pM \to T_p^*M, \quad \flat(v) = \langle v,\cdot\rangle, \quad \text{with inverse} \quad \sharp : T_p^*M \to T_pM.$$

These isomorphism allow us to transport the nondegenerate symmetric bilinear form

$$\langle \cdot, \cdot \rangle : T_pM \times T_pM \to \mathbb{R} \quad \text{to} \quad \langle \cdot, \cdot \rangle : T_p^*M \times T_p^*M \to \mathbb{R}.$$

If $(x^1, \dots, x^n)$ is a smooth coordinate system on $U \subseteq M$, and

$$\langle \cdot, \cdot \rangle | U = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j,$$

then on the cotangent space we have

$$\langle dx^i|_p, dx^j|_p \rangle = g^{ij}(p),$$

where $(g^{ij}(p))$ is the matrix inverse to $g_{ij}(p)$.

We can extend $\langle \cdot, \cdot \rangle$ from T_p^*M to the entire exterior algebra $\Lambda^* T_p^*M$. First, we define a multilinear map

$$\mu : \overbrace{(T_p^*M \times \cdots \times T_p^*M)}^{k} \times \overbrace{(T_p^*M \times \cdots \times T_p^*M)}^{k} \to \mathbb{R}$$

by

$$\mu((\alpha_1, \dots, \alpha_k), (\beta_1, \dots, \beta_k)) = \det \begin{pmatrix} \langle \alpha_1, \beta_1 \rangle & \cdots & \langle \alpha_1, \beta_k \rangle \\ \cdot & & \cdot \\ \langle \alpha_k, \beta_1 \rangle & \cdots & \langle \alpha_k, \beta_k \rangle \end{pmatrix}.$$

This map is skew-symmetric in each set of k variables, when the other is kept fixed, so it defines a symmetric bilinear form

$$\langle \cdot, \cdot \rangle : \Lambda^k T_p^*M \times \Lambda^k T_p^*M \to \mathbb{R} \tag{2.33}$$

such that

$$\langle \alpha_1 \wedge \cdots \wedge \alpha_k, \beta_1 \wedge \cdots \wedge \beta_k \rangle = \det \begin{pmatrix} \langle \alpha_1, \beta_1 \rangle & \cdots & \langle \alpha_1, \beta_k \rangle \\ \cdot & & \cdot \\ \langle \alpha_k, \beta_1 \rangle & \cdots & \langle \alpha_k, \beta_k \rangle \end{pmatrix}.$$

In the special case where $(M, \langle, \cdot, \cdot \rangle)$ is a Riemannian manifold, the symmetric bilinear form $\langle, \cdot, \cdot \rangle$ on $\Lambda^k T_p^*M$ is positive definite. One sees this most easily by noting that if $(\theta^1, \dots, \theta^n)$ is an orthonormal basis for T_p^*M, then

$$\{\theta^{i_1} \wedge \cdots \wedge \theta^{i_k} : i_1 < \cdots < i_k\}$$

is an orthonormal basis for $\Lambda^k T_p^*M$. hence

$$\left\langle \sum a_{i_1 \cdots i_k} \theta^{i_1} \wedge \cdots \wedge \theta^{i_k}, \sum a_{i_1 \cdots i_k} \theta^{i_1} \wedge \cdots \wedge \theta^{i_k} \right\rangle = \sum (a_{i_1 \cdots i_k})^2 \geq 0,$$

with equality if and only if all the $a_{i_1 \cdots i_k}$'s are zero.

By a similar argument, which we leave to the reader, one can show that if $(M, \langle, \cdot, \cdot \rangle)$ is only a pseudo-Riemannian manifold, then the symmetric bilinear form $\langle, \cdot, \cdot \rangle$ on $\Lambda^k T_p^*M$ is nondegenerate.

If $(x^1, \ldots, x^n)$ is a positively oriented coordinate system on $U \subseteq M$, we can define the volume form on U by

$$\Theta|U = \sqrt{|\det(g_{ij})|} dx^1 \wedge \cdots \wedge dx^n.$$

As we have remarked before, the locally defined volume forms agree on coordinate overlaps, so they fit together to yield a globally defined volume form Θ on M. Thus if $D \subseteq M$ is a region with a smooth boundary, the volume of D is given by the formula

$$\text{Volume of } D = \int_D \Theta.$$

In the Riemannian case, if $(\theta^1, \ldots, \theta^n)$ is a basis of smooth one-forms on U such that $\langle \theta^i, \theta^j \rangle = \delta^{ij}$, then $\Theta_U = \theta^1 \wedge \cdots \wedge \theta^n$. In the more general pseudo-Riemannian case, if

$$(\langle \theta^i, \theta^j \rangle) = \begin{pmatrix} -I_{p\times p} & 0 \\ 0 & I_{q\times q} \end{pmatrix},$$

where $p + q = n$, then $\Theta_U = \pm\theta^1 \wedge \cdots \wedge \theta^n$.

Using the volume form, we will now define a linear map

$$\star : \Lambda^k T_p^* M \longrightarrow \Lambda^{n-k} T_p^* M$$

to be called the *Hodge star*. The Hodge star is crucial for a full understanding of Riemannian geometry, but unlike the exterior derivative d it does not pull back under smooth maps (unless they are orientation-preserving isometries).

Proposition. *For each integer k, $0 \leq k \leq n$, there is a unique linear map $\star : \Lambda^k T_p^* M \to \Lambda^{n-k} T_p^* M$ such that whenever α and β are elements of $\Lambda^k T_p^* M$, then*

$$\alpha \wedge \star\beta = \langle \alpha, \beta \rangle \ \Theta(p), \tag{2.34}$$

where $\Theta(p)$ is the evaluation of the volume form at p.

The proof is simpler in the Riemannian case, because we do not have to worry so much about signs. The idea is to take a fixed positively oriented orthonormal basis $(\theta^1, \ldots, \theta^n)$ for $T_p^* M$ and derive an explicit formula in terms of this basis. (By positively oriented we mean that $\Theta = \theta^1 \wedge \cdots \wedge \theta^n$.)

Given a fixed positively oriented orthonormal basis $(\theta^1, \ldots, \theta^n)$, we claim that (2.34) implies the explicit formula

$$\star\left(\theta^{\sigma(1)} \wedge \cdots \wedge \theta^{\sigma(k)}\right) = (\text{sgn } \sigma)\theta^{\sigma(k+1)} \wedge \cdots \wedge \theta^{\sigma(n)}, \tag{2.35}$$

whenever σ is a permutation of $\{1, \ldots, n\}$. Note that this formula is invariant under a transformation which replaces σ by $\sigma \circ \tau$, where τ is a permutation of $(1, \ldots, k)$ or $(k + 1, \ldots, n)$.

So we can assume without loss of generality that $\sigma(1) < \cdots < \sigma(k)$ and $\sigma(k + 1) < \cdots < \sigma(n)$. Under these hypotheses, suppose that

$$\star\left(\theta^{\sigma(1)} \wedge \cdots \wedge \theta^{\sigma(k)}\right) = \sum_{j_{k+1} < \cdots < j_n} c_{j_{k+1} \cdots j_n} \theta^{j_{k+1}} \wedge \cdots \wedge \theta^{j_n}.$$

For a fixed term in this sum, choose $j_1 < \cdots < j_k$ so that $(j_1, \ldots, j_k, j_{k+1}, \ldots, j_n)$ is a permutation of $(1, \ldots, n)$. Then

$$\left(\theta^{j_1} \wedge \cdots \wedge \theta^{j_k}\right) \wedge \star\left(\theta^{\sigma(1)} \wedge \cdots \wedge \theta^{\sigma(k)}\right)$$
$$= c_{j_{k+1}\cdots j_n} \theta^{j_1} \wedge \cdots \wedge \theta^{j_k} \wedge \theta^{j_{k+1}} \wedge \cdots \wedge \theta^{j_n},$$

all the other terms wedging to zero. On the other hand, it follows from (2.34) that

$$\left(\theta^{j_1} \wedge \cdots \wedge \theta^{j_k}\right) \wedge \star\left(\theta^{\sigma(1)} \wedge \cdots \wedge \theta^{\sigma(k)}\right)$$
$$= \begin{cases} \pm\theta^1 \wedge \cdots \wedge \theta^n, & \text{if } (j_1, \ldots, j_k) = (\sigma(1), \ldots, \sigma(k)), \\ 0, & \text{otherwise.} \end{cases}$$

hence

$$c_{j_{k+1}\cdots j_n} = \begin{cases} \pm 1, & \text{if } (j_1, \ldots, j_k) = (\sigma(1), \ldots, \sigma(k)), \\ 0, & \text{otherwise.} \end{cases}$$

It is now easy to verify that

$$c_{\sigma(k+1)\cdots\sigma(n)} = \operatorname{sgn} \sigma,$$

thereby establishing (2.35). This proves uniqueness, because the Hodge star is uniquely determined by its effect on a basis.

To prove existence, one uses (2.35) to define the Hodge star on a fixed orthonormal basis, and checks that it satisfies (2.34).

Note that it follows from the proof that if $(\theta^1, \ldots, \theta^n)$ is *any* orthonormal basis for T_p^*M, then (2.35) holds for that basis. This fact is extremely useful in calculating the Hodge star.

The proof in the general pseudo-Riemannian case is similar, except that the notion of orthonormal basis must be modified in the obvious manner.

Remark: One can give a very useful geometric interpretation of the Hodge star in the Riemannian case. Suppose that W is a k-dimensional subspace of T_p^*M with orthonormal basis $(\theta^1, \ldots, \theta^k)$. Complete $(\theta^1, \ldots, \theta^k)$ to a positively oriented orthonormal basis $(\theta^1, \ldots, \theta^n)$ for T_p^*M. Then $(\theta^{k+1}, \ldots, \theta^n)$ is a positively oriented orthonormal basis for the oriented orthogonal complement $W^\perp$ to W in T_p^*M.

Of course, the Hodge star extends immediately to a linear map

$$\star : \Omega^k(M) \to \Omega^{n-k}(M)$$

which is also called the Hodge star. In the Riemannian case, it is easy to verify that

$$\star(\star\alpha) = (-1)^{k(n-k)}\alpha, \qquad \text{for } \alpha \in \Omega^k(M). \tag{2.36}$$

Example 1. We first consider $\mathbb{E}^2$ with its usual Euclidean coordinates (x, y). In this case, $\Theta = \star 1 = dx \wedge dy$, and

$$\star(dx) = dy, \quad \star(dy) = -dx, \quad \text{so} \quad \star(M dx + N dy) = -N dx + M dy.$$

We can think of the Hodge star in this case as a counterclockwise rotation through 90 degrees. More generally, if $(M, \langle , \cdot, \cdot \rangle)$ is an oriented two-dimensional Riemannian manifold, we can define a *counterclockwise rotation throught* 90 *degrees*

$$J : \Omega^1(M) \to \Omega^1(M) \quad \text{by} \quad J(\omega) = \star(\omega).$$

Example 2. We next consider $\mathbb{E}^3$ with its usual Euclidean coordinates (x, y, z). In this case, $\Theta = \star 1 = dx \wedge dy \wedge dz$. Moreover,

$$\star(dx) = dy \wedge dz, \quad \star(dy) = dz \wedge dx, \quad \star(dz) = dx \wedge dy,$$

$$\star(dy \wedge dz) = dx, \quad \star(dz \wedge dx) = dy, \quad \star(dx \wedge dy) = dz.$$

Note that if α and β are elements of $\Omega^1(\mathbb{E}^3)$, then so is $\star(\alpha \wedge \beta)$, and one can check that its components are the same as those of the cross product $\alpha \times \beta$. More generally, if $(M, \langle , \cdot, \cdot \rangle)$ is an oriented three-dimensional Riemannian manifold we can define a *cross product*

$$\times : \Omega^1(M) \times \Omega^1(M) \to \Omega^1(M) \quad \text{by} \quad \alpha \times \beta = \star(\alpha \wedge \beta).$$

Example 3. Finally, we consider Minkowski space-tme $\mathbb{L}^4$ with its standard coordinates (t, x, y, z), and we take the speed of light to be $c = 1$ so that the Lorentz metric is

$$\langle \cdot, \cdot \rangle = -dt \otimes dt + dx \otimes dx + y \otimes dy + dz \otimes dz. \tag{2.37}$$

In this case, the volume form is

$$\Theta = \star 1 = dt \wedge dx \wedge dy \wedge dz.$$

Clearly, $\star(dt \wedge dx) = \pm dy \wedge dz$. To determine the sign, we note that

$$\langle dt \wedge dx, dt \wedge dx \rangle = -1, \quad \text{so} \quad (dt \wedge dx)(\star dt \wedge dx) = -\Theta.$$

It follows that

$$\star(dt \wedge dx) = -dy \wedge dz, \quad \star(dt \wedge dy) = -dz \wedge dy, \quad \star(dt \wedge dz) = -dx \wedge dy.$$

By similar arguments, one verifies that

$$\star(dy \wedge dz) = dt \wedge dx, \quad \star(dz \wedge dx) = dt \wedge dy, \quad \star(dx \wedge dy) = dt \wedge dz.$$

This Hodge star is invariant under orientation-preserving Lorentz transformation, those orientation-preserving linear transformations of $\mathbb{L}^4$ which leave invariant the flat Lorentz metric (2.37).

Maxwell's equations. Using the Hodge star on Minkowski space-time with the standard Lorentz metric, we can give a particularly elegant formulation of Maxwell's equations from electricity and magnetism, equations formulated by James Clerk Maxwell in 1873. This formulation has the advantage that it extends to the curved space-times of general relativity.

In Maxwell's theory of electricity and magnetism, one imagines that one is given the charge density $\rho(t,x,y,z)$ and the current density

$$\mathbf{J}(t,x,y,z) = J_x \frac{\partial}{\partial x} + J_y \frac{\partial}{\partial y} + J_z \frac{\partial}{\partial z}.$$

The charge and current density should determine the electric and magnetic fields

$$\mathbf{E}(t,x,y,z) = E_x \frac{\partial}{\partial x} + E_y \frac{\partial}{\partial y} + E_z \frac{\partial}{\partial z} = (\text{electric field})$$
$$\text{and} \quad \mathbf{B}(t,x,y,z) = B_x \frac{\partial}{\partial x} + B_y \frac{\partial}{\partial y} + B_z \frac{\partial}{\partial z} = (\text{magnetic field})$$

by means of Maxwell's equations, which are expressed in terms of the divergence and curl operations studied in second year calculus as

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \tag{2.38}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \qquad \nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = 4\pi \mathbf{J}. \tag{2.39}$$

To express these equations in space-time formalism, it is convenient to replace the electric and magnetic fields by a single covariant tensor field of rank two, the so called *Faraday tensor*:

$$\mathcal{F} = -E_x dt \wedge dx - E_y dt \wedge dy - E_z dt \wedge dz$$
$$+ B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy.$$

Then the Hodge star interchanges the $\mathbf{E}$ and the $\mathbf{B}$ fields:

$$\star \mathcal{F} = B_x dt \wedge dx + B_y dt \wedge dy + B_z dt \wedge dz$$
$$+ E_x dy \wedge dz + E_y dz \wedge dx + E_z dx \wedge dy.$$

Exercise IX. a. Show that in Minkowski space-time $\mathbb{L}^4$,

$$\star\star = (-1)^{k+1} : \Omega^k(\mathbb{L}^4) \longrightarrow \Omega^k(\mathbb{L}^4).$$

b. Determine $\star dt$, $\star dx$, $\star dy$ and $\star dz$.

c. Show that Maxwell's equations can be expressed in terms of the Faraday tensor as

$$d\mathcal{F} = 0, \qquad d(\star\mathcal{F}) = \star(4\pi\mathcal{J}), \quad \text{where} \quad \mathcal{J} = -\rho dt + J_x dx + J_y dy + J_z dz.$$

This formulation of Maxwell's equations is described in more detail in Chapter 4 of [27]. From this viewpoint, it is apparent that Maxwell's equations are invariant under the action of the Lorentz group, the linear transformations from $\mathbb{L}^4$ to itself which preserve the Lorentz metric of Minkowski space-time, since the Hodge star is completely defined by the Lorentz metric. This invariance was one of the major reasons for the discovery of special relativity; Maxwell's equations were not invariant under the same group of transformations as Newtonian mechanics. The formulation of Maxwell's equations in terms of the Faraday tensor can be extended immediately to the curved space-times of general relativity. Maxwell's equations thus motivate the study of the operator $\star d \star$ on differential forms.

One approach to solving Maxwell's equations on $\mathbb{L}^4$ is to write $\mathcal{F} = d\mathcal{A}$, where $\mathcal{A}$ is a one-form called the *vector potential*. Appropriately chosen, the vector potential will solve an equation similar to the wave equation, for which an elegant theory has been developed. This approach runs into a snag in a general space-time because the de Rham cohomology class $[\mathcal{F}]$ may not be zero. Resolving this question leads to the theory of connections in vector bundles as we will see later.

2.10 The Hodge Laplacian

In addition to the exterior derivative, an n-dimensional Riemannian manifolds haa a *codifferential* δ which goes in the opposite direction,

$$\delta = (-1)^{nk+1} \star d \star : \Omega^{k+1}(M) \longrightarrow \Omega^k(M).$$

The only thing that is difficult to remember about the codifferential is the sign. For now, note that if M is even-dimensional, $\delta = - \star d \star$. No matter what the dimension of M, it follows from (2.36) that $\delta \circ \delta = 0$.

To explain where the sign comes from, we suppose that $(M, \langle \cdot, \cdot \rangle)$ is a compact oriented Riemannian manifold, possibly with boundary ∂M. We can define a positive definite L^2 inner product

$$(\cdot, \cdot) : \Omega^k(M) \times \Omega^k(M) \longrightarrow \mathbb{R}$$

by setting

$$(\phi, \psi) = \int_M \phi \wedge \star\psi = \int_M \langle \phi, \psi \rangle \Theta.$$

This inner product make $\Omega^k(M)$ into a pre-Hibert space. The funny sign is introducted to make the following proposition valid:

Proposition. *If $\phi \in \Omega^k(M)$ and $\psi \in \Omega^{k+1}(M)$, then*

$$(d\phi, \psi) - (\phi, \delta\psi) = \int_{\partial M} \phi \wedge \star\psi, \tag{2.40}$$

where $\delta = (-1)^{nk+1} \star d \star$.

The proposition is a consequence of Stokes' Theorem:

$$\begin{aligned}
\int_{\partial M} \phi \wedge \star\psi = \int_M d(\phi \wedge \star\psi) &= \int_M d\phi \wedge \star\psi + (-1)^k \int_M \phi \wedge (d \star \psi) \\
&= \int_M d\phi \wedge \star\psi + (-1)^k (-1)^{(n-k)k} \int_M \phi \wedge \star(\star d \star \psi) \\
&= \int_M d\phi \wedge \star\psi - \int_M \phi \wedge \star(\delta\psi) = (d\phi, \psi) - (\phi, \delta\psi).
\end{aligned}$$

Suppose $p \in \partial M$ and that ν is a unit-length element of $T_p^* M$ which points out of ∂M. The volume forms Θ_M and Θ_M are then related by the formula

$$\Theta_M = \nu \wedge \Theta_{\partial M}.$$

If $\phi \in \Omega^k(M)$ and $\psi \in \Omega^{k+1}(M)$, then

$$\nu \wedge \phi \wedge \star\psi = \langle \nu \wedge \phi, \psi \rangle \Theta_M \quad \Rightarrow \quad \phi \wedge \star\psi = \langle \nu \wedge \phi, \psi \rangle \Theta_{\partial M},$$

so we can rewrite (2.40) as

$$(d\phi, \psi) - (\phi, \delta\psi) = \int_{\partial M} \langle \nu \wedge \phi, \psi \rangle \Theta_{\partial M}. \tag{2.41}$$

If we also use the Riemannian metric to identify ν with a unit length tangent vector, it follows from the identity $\langle \nu \wedge \phi, \psi \rangle = \langle \phi, \iota_\nu \psi \rangle$, where ι_ν is the interior product discussed in Exercise VIII, that we can write this formula as

$$(d\phi, \psi) - (\phi, \delta\psi) = \int_{\partial M} \langle \phi, \iota_\nu \psi \rangle \Theta_{\partial M}. \tag{2.42}$$

Note that if $\partial M = \emptyset$, equation (2.40) becomes

$$(d\phi, \psi) = (\phi, \delta\psi). \tag{2.43}$$

The sign in the definition of δ was chosen to make this identity hold. Because of the identity, the codifferential δ is also called the *formal adjoint* to d.

Given an oriented Riemannian manifold, possibly with boundary, we now have two first order differential operators d and δ which satisfy the identities $d^2 = 0$ and $\delta^2 = 0$. Thus

$$\Delta = -(d + \delta)^2 = -d\delta - \delta d : \Omega^k(M) \longrightarrow \Omega^k(M).$$

Definition. The *Hodge Laplacian* is the second order differential operator

$$\Delta : \Omega^k(M) \longrightarrow \Omega^k(M) \quad \text{defined by} \quad \Delta = -(d\delta + \delta d).$$

We say that an element $\phi \in \Omega^k(M)$ is *harmonic* if $\Delta\phi = 0$.

Dangerous curve. Most geometry books use the opposite sign in the definition of the Hodge Laplacian. We have chosen the sign so that the Hodge Laplacian agrees with the Laplace operator used by engineers and physicists when M is Euclidean space.

Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a compact Riemannian manifold without boundary. It then follows from the identity (2.43) that

$$(-\Delta\phi, \psi) = ((d\delta + \delta d)\phi, \psi) = (\delta\phi, \delta\psi) + (d\phi, d\psi) = \cdots = (\phi, -\Delta\psi), \quad (2.44)$$

for all $\phi, \psi \in \Omega^k(M)$. Equation (2.44) states that the Laplace operator Δ is *formally self-adjoint*. Moreover, $\Delta\phi = 0$ if and only if ϕ satisfies the *weak form* of Laplace's equation on k-forms:

$$(\delta\phi, \delta\psi) + (d\phi, d\psi) = 0, \quad \text{for all } \psi \in \Omega^k(M). \quad (2.45)$$

We can take $\psi = \phi$, so that

$$\Delta\phi = 0 \quad \Rightarrow \quad (\delta\phi, \delta\phi) + (d\phi, d\phi) = 0.$$

Since the inner product $(\cdot, \cdot)$ is positive definite, it follows that $d\phi = 0 = \delta\phi$, and we conclude:

Proposition. *If $(M, \langle \cdot, \cdot \rangle)$ is a compact oriented Riemannian manifold without boundary, harmonic k-forms on M are exactly those forms which are both closed and coclosed:*

$$\Delta\phi = 0 \quad \Leftrightarrow \quad d\phi = 0 = \delta\phi. \quad (2.46)$$

The Laplace operator on functions. To gain some intuition, we focus on the simplest case, the Laplace operator on functions. In this case,

$$\Delta(f) = -\delta d(f), \quad \text{since} \quad \delta(f) = 0$$

and $\delta = - \star d \star$.

We imagine that $(x^1, \ldots, x^n)$ is a smooth positively oriented coordinate system on M, so that

$$\Theta = \sqrt{g} dx^1 \wedge \cdots \wedge dx^n.$$

Clearly,

$$\star(dx^i) = \sum_{j=1}^{n} (-1)^{j-1} h^{ij} dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n,$$

for certain functions h^{ij}. Hence

$$dx^i \wedge \star dx^j = \cdots = h^{ij} dx^1 \wedge \cdots \wedge dx^n.$$

On the other hand, it follows from (2.34) that

$$dx^i \wedge \star dx^j = \langle dx^i, dx^j \rangle \sqrt{g} dx^1 \wedge \cdots \wedge dx^n = g^{ij} \sqrt{g} dx^1 \wedge \cdots \wedge dx^n.$$

It follows that $h^{jk} = g^{jk}\sqrt{g}$, and hence

$$\star(dx^i) = \sum_{j=1}^{n}(-1)^{k-1} g^{ij}\sqrt{g} dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n.$$

It follows that

$$\begin{aligned}\star\,(df) &= \star\left(\sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i\right) \\ &= \sum_{i,j=1}^{n} (-1)^{j-1} g^{ij}\sqrt{g}\frac{\partial f}{\partial x^i} dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n,\end{aligned}$$

and

$$d \star d(f) = \sum_{i,j=1}^{n} \frac{\partial}{\partial x^j}\left(g^{ij}\sqrt{g}\frac{\partial f}{\partial x^i}\right) dx^1 \wedge \ldots \wedge dx^n.$$

Thus we finally conclude that

$$\Delta(f) = \star d \star df = \frac{1}{\sqrt{g}} \sum_{i,j=1}^{n} \frac{\partial}{\partial x^j}\left(g^{ij}\sqrt{g}\frac{\partial f}{\partial x^i}\right). \tag{2.47}$$

This formula for the Laplace operator is incredibly useful. Consider, for example, the flow of heat on a smooth surface $M \subseteq \mathbb{E}^3$. From the elementary theory of PDE's we expect the temperature on a smooth homogeneous surface to be described by a function $u : M \times [0, \infty) \to \mathbb{R}$ which satisfies an initial value problem

$$\frac{\partial u}{\partial t} = c^2 \Delta u, \qquad u(p, 0) = h(p),$$

where $h : M \to \mathbb{R}$ is the initial temperature distribution. However, in order to make sense of this equation, we need to define a Laplace operator acting on scalar functions on a smooth surface. The Laplace operator to use is the one given by (2.47).

Exercise X. Suppose that $M = \mathbb{S}^2$, the standard unit two-sphere in $\mathbb{E}^3$, with Riemannian metric expressed in spherical coordinates as

$$\langle \cdot, \cdot \rangle = (\sin^2 \phi) d\theta \otimes d\theta + d\phi \otimes d\phi.$$

Determine the Hodge Laplacian on functions in this case.

Remark. We could use this expression for the Laplacian to solve the initial-value problem for the flow of heat over the unit sphere. The technique of separation of variables and Legendre polynomials gives a very explicit representation of the solution.

2.11 The Hodge Theorem

Suppose now that $(M, \langle \cdot, \cdot \rangle)$ is a compact oriented Riemannian manifold without boundary, and let

$$\mathcal{H}^k(M) = \{ \text{ harmonic } k\text{-forms on } M \ \} = \{\phi \in \Omega^k(M) : \Delta\phi = 0\}.$$

Hodge Theorem. *Every de Rham cohomology class has a unique harmonic representative; thus*

$$H^k_{dR}(M; \mathbb{R}) \cong \mathcal{H}^k(M).$$

This gives an important relationship between topology and solutions to linear elliptic systems of partial differential equations on smooth manifolds.

This theorem has important topological consequences. Thus for example, one easily checks that $\star\Delta = \Delta\star$, so if ϕ is a harmonic k-form, so is $\star\phi$. Thus we obtain:

Poincaré Duality Theorem. *If M is a compact oriented manifold of dimension n,*

$$H^k_{dR}(M; \mathbb{R}) \cong H^{n-k}_{dR}(M; \mathbb{R}).$$

We already know that if M is a compact oriented Riemannian manifold, the volume form represents a nontrivial element of $H^n_{dR}(M; \mathbb{R})$. Poincaré duality enables us to make a finer statement:

Corollary. *If M is a compact connected oriented manifold of dimension n,*

$$H^n_{dR}(M; \mathbb{R}) \cong \mathbb{R}.$$

Of course, Poincaré duality simplifies the calculation of de Rham cohomology of compact oriented manifolds via the Homotopy and Mayer-Vietoris Theorems.

Example. Let us suppose that

$$M = T^n = \overbrace{S^1 \times \cdots \times S^1}^{n},$$

with the flat metric defined by requiring that the covering

$$\pi : \mathbb{E}^n \to T^n, \qquad \pi(x^1, \ldots, x^n) = (e^{2\pi i x^1}, \ldots, e^{2\pi i x^n})$$

be a local isometry. We define one-forms $(\theta^1, \ldots, \theta^n)$ on T^n by $\pi^*\theta^i = dx^i$. Then $(\theta^1, \ldots, \theta^n)$ form a positively oriented orthonormal basis for the one-forms on M and it follows from (2.35) that

$$d(\theta^{i_1} \wedge \cdots \wedge \theta^{i_k}) = 0, \qquad \delta(\theta^{i_1} \wedge \cdots \wedge \theta^{i_k}) = 0.$$

If $\omega \in \Omega^k(T^n)$, then

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \ldots i_k} \theta^{i_1} \wedge \cdots \wedge \theta^{i_k},$$

for some smooth real-valued functions $f_{i_1 \ldots i_k}$ on T^n and if these functions are constant, $d\omega = 0 = \delta\omega$ and hence ω is harmonic. Conversely, a direct calculation shows that

$$\Delta\omega = \sum_{i_1 < \cdots < i_k} (\Delta f_{i_1 \ldots i_k}) \theta^{i_1} \wedge \cdots \wedge \theta^{i_k},$$

and hence if ω is harmonic, so is each function $f_{i_1 \ldots i_k}$. But it follows from (2.46) that the only harmonic functions on a compact oriented manifolds are constant on each connected component, so the only harmonic forms on T^n are

$$\omega = \sum_{i_1 < \cdots < i_k} c_{i_1 \ldots i_k} \theta^{i_1} \wedge \cdots \wedge \theta^{i_k},$$

where the $c_{i_1 \ldots i_k}$'s are constants. In other words,

$$\{\theta^{i_1} \wedge \cdots \wedge \theta^{i_k} : i_1 < \cdots < i_k\}$$

is a basis for $\mathcal{H}^k(T^n)$ and hence the dimension of $H^k_{dR}(T^n; \mathbb{R})$ is $\binom{n}{k}$. Thus Hodge theory yields a very explicit representation for the cohomology of the torus in terms of harmonic forms.

There are two steps to the proof of the Hodge Theorem. Uniqueness of harmonic representatives is easy. If ω_1 and ω_2 were two harmonic k-forms on M representing the same cohomology class, say $\omega_1 - \omega_2 = d\phi$, then

$$(\omega_1 - \omega_2, \omega_1 - \omega_2) = (\omega_1 - \omega_2, d\phi) = (\delta\omega_1 - \delta\omega_2, \phi) = 0,$$

and positive-definiteness of $(\cdot, \cdot)$ implies that $\omega_1 - \omega_2 = 0$, so $\omega_1 = \omega_2$.

Thus the difficult step is establishing existence of a harmonic representative of a given cohomology class. A proof of this step is beyond the scope of the course, but we will briefly sketch the idea behind the proof, which is similar to the argument given in [16], Chapter 2. (Some readers may want to skip this sketch on a first reading.) Let ω_0 be a smooth k-form representing a given de Rham cohomology class $[\omega_0]$. Then any k-form in the same cohomology class must be of the form $\omega = \omega_0 + d\alpha$, where α is a smooth $(k-1)$-form. The idea is to find a minimum for the function

$$F : \Omega^{k-1}(M) \to \mathbb{R} \quad \text{defined by} \quad F(\alpha) = (\omega_0 + d\alpha, \omega_0 + d\alpha).$$

If we knew that the function F possessed a smooth minimum α_0, we could apply first variation to obtain

$$\begin{aligned} 0 = \frac{d}{dt}(\omega_0 + d\alpha_0 + td\beta, \omega_0 + d\alpha_0 + td\beta)\bigg|_{t=0} \\ = 2(\omega_0 + d\alpha_0, d\beta), \quad \text{for all } \beta \in \Omega^{k-1}(M), \end{aligned} \quad (2.48)$$

which would imply $\delta(\omega_0 + d\alpha) = 0$. Then $\omega_0 + d\alpha_0$ would be both closed and coclosed, and hence the desired harmonic representative for the given de Rham cohomology class.

To construct the minimum, we might take a minimizing sequence for F, a sequence $\{\alpha_i\}$ of $(k-1)$-forms such that

$$F(\alpha_i) \to \mu = \inf\{F(\alpha) : \alpha \in \Omega^{k-1}(M)\}.$$

The difficulty is that such a sequence might not converge, because we could replace α_i by $\alpha_i + d\phi_i$ for some $(k-2)$-form ϕ_i without changing the value of F, and it might be that $\phi_i \to \infty$. Thus we focus instead on the sequence $\{\omega_i = \omega_0 + d\alpha_i\}$ in $\Omega^k(M)$.

To show that a subsequence of $\{\omega_i\}$ converges, we first complete the space $\Omega^k(M)$ of smooth k-forms on M with respect to the inner product $(\cdot, \cdot)$, thereby obtaining a Hilbert space $L^2(\Omega^k(M))$. The closure of $\omega_0 + d(\Omega^{k-1}(M))$ is a translate $\omega_0 + H$ of a closed Hilbert subspace $H \subseteq L^2(\Omega^k(M))$. Instead of minimizing F directly we construct an element ω_∞ of $\omega_0 + H$ which is closest to the origin. The sequence $\{\omega_i\}$ must then converge to ω_∞ in $\omega_0 + H$.

We can conclude that

$$(\omega_\infty, \delta\phi) = 0 \quad \text{and} \quad (\omega_\infty, \delta\phi) = 0, \quad \text{for all } \phi \in \Omega^k(M), \tag{2.49}$$

the first since ω_∞ is a limit of closed forms, the second by the first variation argument (2.48), since ω_∞ is a minimum for the map

$$F : (\text{closure of } d\Omega^{k-1}(M)) \to \mathbb{R} \quad \text{defined by} \quad F(\gamma) = (\omega_0 + \gamma, \omega_0 + \gamma).$$

In the terminology of PDE theory (2.49) states that ω_∞ is *weakly* closed and coclosed.

We have explained the part of the proof that can be done without "elliptic regularity theory," which implies that a k-form which is weakly closed and coclosed is actually smooth, and hence a bona fide harmonic form. To apply the regularity theory, one uses the fact that

$$d + \delta : \Omega^k(M) \to \Omega^{k+1}(M) \oplus \Omega^{k-1}(M)$$

is an elliptic operator. Unfortunately, the regularity theory requires the development of considerable analysis: construction of Sobolev spaces, the Sobolev Lemma and the Rellich Lemma. We refer to [33] and [10] for presentation of this theory.

Once we know that ω_∞ is a smooth harmonic form, it is not difficult to show that it lies in the same de Rham cohomology class as ω_0 and all of the ω_i's. One way of establishing this would be to use the de Rham Isomorphism Theorem, since the fact that $\omega_i \to \omega_\infty$ in L^2 implies that the integral of ω_i over any singular cycle approaches the integral of ω_∞ over that cycle.

2.12 d and δ in terms of moving frames

The Hodge Laplacian is only one of many Laplace operators that arise in differential geometry. In the next few sections, we will describe another, the so-called "rough Laplacian." The relationship between the Hodge and rough Laplacians gives an important relationships between curvature and topology of Riemannian manifolds.

Definition. A linear operator $D : \Omega^*(M) \to \Omega^*(M)$ is said to be a *derivation* if

$$D(\phi \wedge \psi) = D(\phi) \wedge \psi + \phi \wedge D(\psi), \quad \text{for all } \phi, \psi \in \Omega^*(M).$$

It is said to be *skew-derivation* if instead

$$D(\phi \wedge \psi) = D(\phi) \wedge \psi + (-1)^k \phi \wedge D(\psi), \quad \text{for all } \phi \in \Omega^k(M) \text{ and } \psi \in \Omega^*(M).$$

Thus the Lie derivative L_X in the direction of a vector field X is a derivation, while the exterior derivative d is a skew-derivation. Another skew-derivation is the interior product ι_X in the direction of a vector field X on $\omega^*(M)$.

If $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold and X is a vector field on M, then the Levi-Civita connection defines a derivation ∇_X of $\Omega^*(M)$. Indeed, if Y is a smooth vector field on M, the Levi-Civita connection defines a covariant derivative $\nabla_X Y$ of Y in the direction of X. If $\omega \in \Omega^k(M)$, we define the *covariant derivative* of ω in the direction of X in such a way that the Leibniz rule will hold. This forces

$$\begin{aligned} &X(\omega(Y_1, \ldots, Y_k)) \\ &\qquad = (\nabla_X \omega)(Y_1, \ldots, Y_k) + \omega(\nabla_X Y_1, \ldots, Y_k) + \cdots + \omega(Y_1, \ldots, \nabla_X Y_k), \end{aligned}$$

or

$$\begin{aligned} &(\nabla_X \omega)(Y_1, \ldots, Y_k) \\ &\qquad = X(\omega(Y_1, \ldots, Y_k)) - \omega(\nabla_X Y_1, \ldots, Y_k) - \cdots - \omega(Y_1, \ldots, \nabla_X Y_k). \end{aligned} \tag{2.50}$$

It is readily verified that

$$\nabla_X (f\omega + \phi) = X(f)\omega + f\nabla_X \omega + \nabla_X \phi, \qquad \nabla_{fX+Y}\omega = f\nabla_X \omega + \nabla_Y \omega.$$

Moreover, ∇_X satisfies the derivation property,

$$\nabla_X(\phi \wedge \psi) = \nabla_X(\phi) \wedge \psi + \phi \wedge \nabla_X(\psi).$$

Suppose now that $(M, \langle \cdot, \cdot \rangle)$ is an n-dimensional oriented Riemannian manifold and U is an open subset of M.

Definition. A *moving orthonormal frame* on U is an ordered collection $(e_1, \ldots e_n)$ of vector fields on U such that

$$\langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

The *dual orthonormal coframe* is the ordered collection $(\theta^1, \ldots, \theta^n)$ defined so that

$$\theta^i(e_j) = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Of course, in terms of the induced inner product on T^*M,

$$\langle \theta^i, \theta^j \rangle = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

We define a linear operator $\varepsilon_{\theta^i} : \Omega^*(M) \to \Omega^*(M)$ by $\varepsilon_{\theta^i}(\phi) = \theta^i \wedge \phi$. The exterior derivative can then be expressed in terms of the covariant derivative and ε_{θ^i}:

Proposition 1. *$d = \sum_{i=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i}$ on U.*

To prove this proposition, we first check it on functions. If f is a smooth real-valued function on U and $p \in U$, we can choose coordinates $(x^1, \ldots, x^n)$ on a neighborhood of p such that

$$x^1(p) = \cdots x^n(p) = 0, \quad e_i(p) = \left.\frac{\partial}{\partial x^i}\right|_p, \quad \theta^i(p) = dx^i|_p.$$

Then

$$\sum_{i=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i}(f)(p) = \sum_{i=1}^n dx^i|_p \left.\frac{\partial}{\partial x^i}\right|_p (f) = df(p).$$

Next we check the formula on one-forms. If ω is a smooth one-form on U, then

$$\begin{aligned}
\left(\sum_{i=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i}\omega \right)(X,Y) &= \sum_{i=1}^n \left(\theta^i \wedge \nabla_{e_i}\omega \right)(X,Y) \\
= \sum_{i=1}^n \theta^i(X)\nabla_{e_i}\omega(Y) - \sum_{i=1}^n \theta^i(Y)\nabla_{e_i}\omega(X) &= (\nabla_X\omega)(Y) - (\nabla_Y\omega)(X) \\
&= X(\omega(Y)) - Y(\omega(X)) - \omega(\nabla_X Y) + \omega(\nabla_Y X) \\
&= X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) = d\omega(X,Y),
\end{aligned}$$

which gives the formula for one-forms. In the calculation we used the fact that

$$\sum_{i=1}^n \theta^i(X)\nabla_{e_i} = \nabla_X.$$

Thus to finish the proof of the proposition, we need only cite the following lemma:

Lemma. *If D and D' are both derivations, or both skew-derivations, of $\Omega^*(M)$ such that*

$$D(f) = D'(f), \quad \text{for } f \in \Omega^0(M) \quad \text{and} \quad D(\omega) = D'(\omega), \quad \text{for } \omega \in \Omega^1(M),$$

then $D = D'$.

We sketch the proof of the lemma for derivations, the case of skew-derivations being the same except for a few signs. Any element $\omega \in \Omega^k(M)$ can be divided into a sum of k-forms with supports in local coordinate systems, so it suffices to prove a local coordinate version of the lemma. If $(U, (x^1, \dots, x^n)$ is a smooth coordinate system and the support of ω is contained in U, it can be written as

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

for some choice of smooth functions $f_{i_1 \dots i_k}$. Thus if D and D' are both derivations, we have

$$D(\omega) = \sum_{i_1 < \cdots < i_k} D(f_{i_1 \dots i_k}) dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$
$$+ \sum_{i_1 < \cdots < i_k} f_{i_1 \dots i_k} D(dx^{i_1}) \wedge \cdots \wedge dx^{i_k} + \cdots + \sum_{i_1 < \cdots < i_k} f_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge D(dx^{i_k}),$$

$$D'(\omega) = \sum_{i_1 < \cdots < i_k} D'(f_{i_1 \dots i_k}) dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$
$$+ \sum_{i_1 < \cdots < i_k} f_{i_1 \dots i_k} D'(dx^{i_1}) \wedge \cdots \wedge dx^{i_k} + \cdots + \sum_{i_1 < \cdots < i_k} f_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge D'(dx^{i_k}).$$

By hypothesis, the right-hand sides of the two expressions are the same, so $D = D'$ as required.

Of course, we can ask whether the codifferential

$$\delta = (-1)^{(n(k+1)+1} \star d \star : \Omega^k(M) \longrightarrow \Omega^{k-1}(M)$$

has a similar expression

Proposition 2. *$\delta = -\sum_{i=1}^n \iota_{e_i} \circ \nabla_{e_i}$ on U.*

But this actually follows from Proposition 1, together with the straightforward calculation that

$$\iota_{e_i} = (-1)^{nk+n} \star \varepsilon_{\theta^i} \star : \Omega^k(M) \longrightarrow \Omega^{k-1}(M).$$

Indeed, from this formula, and the fact that $\star$ commutes with ∇_X, it follows that

$$\delta = (-1)^{nk+n+1} \star \circ \sum_{i=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i} \circ \star$$
$$= (-1)^{nk+n+1} \sum_{i=1}^n \star \circ \varepsilon_{\theta^i} \circ \star \circ \nabla_{e_i} = -\sum_{i=1}^n \iota_{e_i} \circ \nabla_{e_i}.$$

2.13 The rough Laplacian

There is a second Laplace operator that is defined on $\Omega^*(M)$. As preparation for defining it, we define an operator

$$\nabla_{X,Y} : \Omega^*(M) \to \Omega^*(M) \quad \text{by} \quad \nabla_{X,Y}(\omega) = \nabla_X \nabla_Y \omega - \nabla_{\nabla_X Y}\omega,$$

for vector fields X and Y on M.

Lemma. $\nabla_{fX,Y}\omega = f\nabla_{X,Y}\omega = \nabla_{X,fY}\omega$.

The proof is a straightforward calculation; for example, consider the first equality:

$$\begin{aligned} \nabla_{fX,Y}\omega &= \nabla_{fX}\nabla Y\omega - \nabla_Y \nabla_{fX}\omega - \nabla[fX,Y]\omega \\ &= f\nabla_X\nabla_Y\omega - Y(f)\nabla_X\omega - f\nabla_Y\nabla X\omega - \nabla_{fXY-Y(f)X-fYX} \\ &= f\nabla_X\nabla_Y\omega - f\nabla_Y\nabla X\omega - f\nabla[X,Y] = f\nabla_{X,Y}\omega. \end{aligned}$$

This lemma implies that we have a well-defined linear map

$$\nabla_{x,y} : \Lambda^* T_p^* M \to \Lambda^* T_p^* M, \quad \text{for each choice of } x, y \in T_p M.$$

Definition. The *rough Laplacian* on $\Omega^*(M)$ is the linear operator $\Delta_R : \Omega^*(M) \to \Omega^*(M)$ defined by

$$\Delta_R(\omega) = \sum_{i=1}^n \nabla_{e_i,e_i}(\omega) \quad \text{on} \quad U,$$

whenever $(e_1, \dots e_n)$ is a moving orthonormal frame on U. (It is immediately verified that the expression on the right is independent of the choice of moving orthonormal frame.)

If $\omega \in \Omega^k(M)$ we define $\|\nabla\omega\|$ by

$$\|\nabla\omega\|^2 = \sum_{i=1}^n \|\nabla_{e_i}\omega\|^2,$$

whenever $(e_1, \dots e_n)$ is a moving orthonormal frame.

Proposition. *If $(M, \langle\cdot,\cdot\rangle)$ is a compact oriented Riemannian manifold with empty boundary and volume form Θ_M, then for any $\omega \in \Omega^k(M)$,*

$$\int_M \langle -\Delta_R(\omega), \omega\rangle \Theta_M = \int_M \|\nabla\omega\|^2 \Theta_M.$$

To prove the proposition, we define a vector field X on M by

$$X = \sum_{i=1}^n \langle \nabla_{e_i}\omega, \omega\rangle e_i,$$

whenever $(e_1, \ldots e_n)$ is a positively oriented moving orthonormal frame on $U \subseteq M$. Clearly, X is independent of choice of moving frame, hence globally defined on M, and in fact,

$$X = \frac{1}{2}\mathrm{grad}\langle \omega, \omega \rangle.$$

By positively oriented, we mean that if $(\theta^1, \ldots, \theta^n)$ is the orthonormal coframe on U dual to $(e_1, \ldots e_n)$, then

$$\Theta_M = \theta^1 \wedge \cdots \wedge \theta^n.$$

Now set

$$\psi = \iota_X \Theta_M = \sum_{i=1}^{n} (-1)^{i-1} \langle \nabla_{e_i} \omega, \omega \rangle \theta^1 \wedge \cdots \wedge \theta^{i-1} \wedge \theta^{i+1} \wedge \cdots \wedge \theta^n.$$

To finish the proof of the proposition, it will suffice to show that

$$d\psi = \langle \Delta_R(\omega), \omega \rangle \Theta_M + \|\nabla \omega\|^2 \Theta_M. \tag{2.51}$$

We need only prove (2.51) at a given point $p \in M$, and we can that the point lies within an open neighborhood on which we have a moving orthonormal frame $(e_1, \ldots, e_n)$ with corresponding orthonormal coframe $(\theta^1, \ldots, \theta^n)$. Moreover, we can assume without loss of generality that

$$\nabla_{e_i} e_j(p) = 0, \quad \text{and hence} \quad \nabla_{e_i} \theta^j(p) = 0.$$

Since $\nabla_X : \Omega^*(M) \to \Omega^*(M)$ is a derivation,

$$\nabla_{e_i}(\theta^{j_1} \wedge \cdots \wedge \theta^{j_k})(p) = 0.$$

Hence it follows from Proposition 1 of the previous section that

$$\begin{aligned} &d\psi(p) \\ &\quad = \sum_{i=1}^{n} (-1)^{i-1} \varepsilon_{\theta^i} \circ \nabla_{e_i} \left(\langle \nabla_{e_i} \omega, \omega \rangle \right) \theta^1 \wedge \cdots \wedge \theta^{i-1} \wedge \theta^{i+1} \wedge \cdots \wedge \theta^n(p) \\ &\qquad\qquad = \sum_{i=1}^{n} \left[\langle \nabla_{e_i} \nabla_{e_i} \omega, \omega \rangle + \langle \nabla_{e_i} \omega, \nabla_{e_i} \omega \rangle \right] \theta^1 \wedge \cdots \wedge \theta^n(p) \\ &\qquad\qquad\qquad\qquad = [\langle \Delta_R(\omega) \omega \rangle + \|\nabla \omega\|^2] \Theta_M(p), \end{aligned}$$

which establishes (2.51) and finishes the proof of the proposition.

2.14 The Weitzenböck formula

We now have two operators, the Hodge Laplacian Δ and the rough Laplacian Δ_R on $\Omega^*(M)$. It turns out that the relationship between the two is given by the Riemann-Christoffel curvature of the Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$.

Here the curvature of the Levi-Civita connection on $\Omega^*(M)$ is defined by the expected formula,

$$R(X,Y)\omega = \nabla_X\nabla_Y\omega - \nabla_Y\nabla_X\omega - \nabla_{[X,Y]}\omega,$$

whenever X and Y are smooth vector fields on M. Since the connection is symmetric, $R(X,Y)f = 0$, for $f \in \Omega^*(M)$ and, since the Riemannian metric provides an isomorphism between TM and T^*M preserving inner products, $R(X,Y)$ on T^*M is essentially the same as $R(X,Y)$ on TM. Moreover, one can check that $R(X,Y)$ is a derivation, so it is completely determined by its effect on $\Omega^0(M)$ and $\Omega^1(M)$ by the lemma of § 2.12.

To derive the relationship between the two Laplace operators, we use the identities,

$$d = \sum_{i=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i}, \qquad \delta = -\sum_{i=1}^n \iota_{e_i} \circ \nabla_{e_i}.$$

We will need the following easily verified identities for the interior and exterior products:

$$\varepsilon_{\theta^i} \circ \iota_{e_j} = -\iota_{e_j} \circ \varepsilon_{\theta^i}, \quad \text{if } i \neq j,$$
$$\varepsilon_{\theta^i} \circ \iota_{e_i} + \iota_{e_i} \circ \varepsilon_{\theta^i} = \mathrm{id}.$$

We now calculate the Hodge Laplacian at a point $p \in M$, and we assume without loss of generality that

$$\nabla_{e_i} e_j(p) = 0, \quad \text{and} \quad \nabla_{e_i}\theta^j(p) = 0,$$

as before. This implies that $[e_i, e_j](p) = \nabla_{e_i} e_j(p) - \nabla_{e_j} e_i(p) = 0$ as well. Thus

$$\begin{aligned}
\Delta(\omega) &= -d\delta(\omega) - \delta d(\omega) \\
&= \sum_{i,j=1}^n \varepsilon_{\theta^i} \circ \nabla_{e_i} \circ \iota_{e_j} \circ \nabla_{e_j}(\omega) + \sum_{i,j=1}^n \iota_{e_j} \circ \nabla_{e_j} \circ \varepsilon_{\theta^i} \circ \nabla_{e_i}(\omega) \\
&= \sum_{i,j=1}^n \varepsilon_{\theta^i} \circ \iota_{e_j} \circ \nabla_{e_i} \circ \nabla_{e_j}(\omega) + \sum_{i,j=1}^n \iota_{e_j} \circ \varepsilon_{\theta^i} \circ \nabla_{e_j} \circ \nabla_{e_i}(\omega) \\
&= \sum_{i=1}^n \nabla_{e_i}\nabla_{e_i}(\omega) + \sum_{i<j} \varepsilon_{\theta^i} \circ \iota_{e_j} (\nabla_{e_i}\nabla_{e_j} - \nabla_{e_j}\nabla_{e_i})(\omega).
\end{aligned}$$

We conclude that

$$\Delta(\omega) = \Delta_R(\omega) + \frac{1}{2}\sum_{i\neq j} (\varepsilon_{\theta^i} \circ \iota_{e_j} + \iota_{e_i} \circ \varepsilon_{\theta^j}) R(e_i, e_j)(\omega). \tag{2.52}$$

This formula, relating the two Laplacians is called the *Weitzenböck formula*.

2.15 Ricci curvature and Hodge theory

One of the most studied problems in differential geometry concerns relationships between curvature and topology of Riemannian manifolds. In addition to the Riemann-Christoffel curvature tensor studied in §1.8, we have various "contractions" of the curvature which are also important.

Definition. The *Ricci curvature* of a pseudo-Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ is the bilinear form

$$\mathrm{Ric} : T_pM \times T_pM \to \mathbb{R} \quad \text{defined by} \quad \mathrm{Ric}(x,y) = (\text{Trace of } v \mapsto R(v,x)y).$$

It follows from the curvature symmetries that Ric is symmetric, that is,

$$\mathrm{Ric}(x,y) = \mathrm{Ric}(y,x).$$

Indeed, if $(e_1, \ldots, e_n)$ is a basis for T_pM such that $\langle e_i, e_j\rangle = 0$ if $i \neq j$ and

$$\langle e_i, e_i\rangle = 1, \quad \text{for } 1 \leq i \leq p, \qquad \langle e_i, e_i\rangle = -1, \quad \text{for } p+1 \leq i \leq n, \tag{2.53}$$

then

$$\mathrm{Ric}(x,y) = \sum_{i=1}^{p} \langle R(e_i,x)y, e_i\rangle - \sum_{i=p+1}^{n} \langle R(e_i,x)y, e_i\rangle = \mathrm{Ric}(y,x),$$

the last equality following from the curvature symmetries presented in §1.9. We say that a Riemannian manifold has positive Ricci curvature if for each $p \in M$, the symmetric bilinear form $\mathrm{Ric} : T_pM \times T_pM \to \mathbb{R}$ is positive definite. Examples of manifolds of positive Ricci curvature include the spheres, the compact Lie groups $O(n)$ and $U(n)$ with biinvariant Riemannian metrics, and the complex projective space.

In this section, we will describe how Hodge's theory of harmonic forms allows us to draw conclusions about the topology of manifolds with positive Ricci curvature. Our approach is to use the Weitzenböck formula (2.52):

$$-\Delta\omega = -\Delta_R(\omega) + \mathcal{K}(\omega), \quad \text{where} \quad \mathcal{K}(\omega) = -\frac{1}{2}\sum_{i \neq j} (\varepsilon_{\theta^i} \circ \iota_{e_j} + \iota_{e_i} \circ \varepsilon_{\theta^j}) R(e_i, e_j)(\omega).$$

It follows from this formula and the proposition from §2.13 that

$$\begin{aligned} \int_M \langle -\Delta\omega, \omega\rangle \Theta_M &= \int_M \langle -\Delta_R(\omega) + \mathcal{K}(\omega), \omega\rangle \Theta_M \\ &= \int_M \|\nabla\omega\|^2 \Theta_M + \int_M \langle \mathcal{K}(\omega), \omega\rangle \Theta_M. \end{aligned} \tag{2.54}$$

If we can show that whenever $\omega \in \Omega^k(M)$, the curvature term $\langle \mathcal{K}(\omega), \omega\rangle > 0$, it will follow that

$$\int_M \langle -\Delta\omega, \omega\rangle \Theta_M > 0, \quad \text{for all } \omega \in \Omega^k(M).$$

From this we can conclude that there cannot be any nonzero harmonic k-forms, and hence it follows from Hodge's Theorem that $H^k_{dR}(M;\mathbb{R}) = 0$.

This approach to determining relationships between curvature and topology is called *Bochner's technique*, and has many applications to finding relationships between curvature and topology. We will apply this technique to the case where $\omega \in \Omega^1(M)$ to obtain:

Theorem. *Let (M, ω) be a compact connected oriented Riemannian manifold with positive Ricci curvature. Then $H^1_{dR}(M;\mathbb{R}) = 0$.*

To carry out the calculation of $\langle \mathcal{K}(\omega), \omega\rangle$ when $\omega \in \Omega^1(M)$, we use the Riemannian metric $\langle \cdot, \cdot \rangle$ on M to identify T_pM with T_p^*M via the isomorphism

$$v \in T_pM \quad \mapsto \quad \langle v, \cdot \rangle \in T_p^*M.$$

Using this isomorphism, we could identify the elements e_i of a moving orthonormal frame with the corresponding elements θ^i of the dual orthonormal coframe. It then suffices to show that

$$\langle \mathcal{K}(\theta^i), \theta^i \rangle = \langle \mathcal{K}(e_i), e_i \rangle > 0, \quad \text{for } 1 \le i \le n.$$

But

$$\begin{aligned}
\langle \mathcal{K}(e_k), e_k \rangle &= -\frac{1}{2} \sum_{i \neq j} \left\langle (\varepsilon_{\theta^i} \circ \iota_{e_j} + \iota_{e_i} \circ \varepsilon_{\theta^j}) R(e_i, e_j) e_k, e_k \right\rangle \\
&= -\frac{1}{2} \sum_{i \neq j} \left\langle R(e_i, e_j) e_k, (\varepsilon_{\theta^j} \circ \iota_{e_i} + \iota_{e_j} \circ \varepsilon_{\theta^i}) e_k \right\rangle \\
&= -\frac{1}{2} \sum_{i \neq j} \left\langle R(e_i, e_j) e_k, (\varepsilon_{\theta^j} \circ \iota_{e_i} - \varepsilon_{\theta^i} \circ \iota_{e_j}) e_k \right\rangle \\
&= -\frac{1}{2} \sum_{j=1}^{n} \langle R(e_k, e_j) e_k, e_j \rangle + \frac{1}{2} \sum_{j=1}^{n} \langle R(e_i, e_k) e_k, e_i \rangle \\
&= \operatorname{Ric}(e_k, e_k).
\end{aligned}$$

Thus if M has positive Ricci curvature, it does indeed follow from (2.54) that there are no nonzero harmonic one-forms on M and $H^1_{dR}(M;\mathbb{R}) = 0$.

2.16 The curvature operator and Hodge theory

Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold. Just as we have defined $\Lambda^2 T_p^*M$, we could define $\Lambda^2 T_pM$ and use the Riemannian metric to define a positive definite inner product

$$\langle \cdot, \cdot \rangle : \Lambda^2 T_p^*M \times \Lambda^2 T_p^*M \longrightarrow \mathbb{R}.$$

We can then organize the information contained in the Riemann-Christoffel curvature tensor into a symmetric linear map on $\Lambda^2 T_pM$.

Definition. The *curvature operator* of a Riemmannian manifold $(M, \langle \cdot, \cdot \rangle)$ at a point $p \in M$ is the linear map

$$\mathcal{R} : \Lambda^2 T_p M \to \Lambda^2 T_p M \quad \text{defined by} \quad \langle \mathcal{R}(x \wedge y), z \wedge w \rangle = \langle R(x, y)w, z \rangle.$$

The curvature symmetries imply that $\mathcal{R}$ is well defined and symmetric:

$$\langle \mathcal{R}(x \wedge y), z \wedge w \rangle = \langle x \wedge y, \mathcal{R}(z \wedge w) \rangle .$$

We say that $(M, \langle \cdot, \cdot \rangle)$ has *positive curvature operators* if for every $p \in M$,

$$\langle \mathcal{R}(\xi), \xi \rangle > 0, \quad \text{for all nonzero } \xi \in \Lambda^2 T_p^* M.$$

We can ask the question of whether having positive curvature operators puts constraints on the topology of a Riemannian manifold. A first answer to this question was given by the following theorem:

Theorem of Gallot and Meyer (1975). *Let $(M, \langle \cdot, \cdot \rangle)$ be a compact connected oriented n-dimensional Riemannian manifold with positive curvature operators. Then*

$$H^k_{dR}(M; \mathbb{R}) \cong \begin{cases} \mathbb{R}, & \text{if } k = 0 \text{ or } k = n, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, M has the same de Rham cohomology as S^n.

A complete proof of this theorem is beyond the scope of the course, but we will give a sketch. The idea is relatively simple.

Since M is connected $H^0_{dR}(M; \mathbb{R}) \cong \mathbb{R}$ while Poincaré duality shows that $H^n_{dR}(M; \mathbb{R}) \cong \mathbb{R}$. Thus we need only verify that

$$0 < k < n \quad \Rightarrow \quad H^0_{dR}(M; \mathbb{R}) \cong \mathcal{H}^k(M) = 0.$$

In other words, we need only verify that M has no harmonic k-forms when $0 < k < n$. Applying (2.54) to the case where $\Delta\omega = 0$, we see that it would suffice to show that if ω is a nonzero k-form, where $0 < k < n$, then

$$\begin{aligned} &\langle \mathcal{R}(\xi), \xi \rangle > 0, \quad \text{for all nonzero } \xi \in \Lambda^2 T_p M \\ &\qquad\qquad \Rightarrow \quad \langle \mathcal{K}(\omega), \omega \rangle > 0, \text{for all nonzero } \omega \in \Lambda^2 T_p^* M. \end{aligned}$$

This is a purely algebraic problem that was solved by Gallot and Meyer. We will sketch further details in the next section.

The Hodge Laplacian is a *linear* elliptic operator. It might be hoped that stronger results could be obtained using *nonlinear* PDE's. This was realized by Böhm and Wilking [6] who showed in 2006 that compact simply connected Riemannian manifolds with positive curvature operators are in fact diffeomorphic to spheres by using the nonlinear Ricci curvature evolutions equations introduced by Hamilton. (We say that a connected manifold M is simply connected if any path $\gamma : [0, 1] \to M$ such that $\gamma(0) = \gamma(1)$ can be continuously deformed to a point; see Chapter 1 of [14].)

2.17 Proof of Gallot-Meyer Theorem*

For those who are curious about how the proof goes, we provide a sketch of the argument, using the notion of Clifford algebra, which is important for other applications as well, and will be discussed later in the course (see §5.6). Our argument follows [19].

Using the Riemannian metric we can identify T_p^*M with T_pM and $\Lambda^*T_p^*M$ with Λ^*T_pM. Thus if we have an orthonormal basis $(e_1, \ldots, e_n)$ at a given point $p \in M$, we identify θ^i with e_i.

The idea behind Clifford algebras is to define a new product (denoted by a dot) on Λ^*T_pM by requiring that it be associative and

$$e_i \cdot e_j + e_j \cdot e_i = -2\delta_{ij}.$$

With this new product, we call Λ^*T_pM the Clifford algebra of T_pM and denote it by $\mathrm{Cl}(T_pM)$. Thus the *Clifford algebra* of T_pM is the associative algebra with identity generated by the elements of $v \in T_pM$ subject to the relations

$$v \cdot w + w \cdot v = -2\langle v, w\rangle, \quad \text{for all } v, w \in T_pM.$$

The Clifford algebras $\mathrm{Cl}(T_pM)$ at the various points $p \in M$ fit together to form a vector bundle $\mathrm{Cl}(TM)$ of rank 2^n over M. We identify

$$\Omega^*(M) = \{ \text{ sections of } \Lambda^*(TM) \ \} \quad \text{with} \quad \{ \text{ sections of } \mathrm{Cl}(TM) \ \}.$$

With this new Clifford multiplication, we can write

$$d + \delta = \sum_{i=1}^{n} e_i \cdot \nabla_{e_i}.$$

Note that $(d + \delta)^2 = -\Delta$, and we call $d + \delta$ a Dirac operator. The Clifford product simplifies the form of the operator $\mathcal{K}$:

$$\langle \mathcal{K}(\omega), \omega\rangle = \frac{1}{2} \sum_{i \neq j} \langle e_i \cdot e_j \cdot R(e_i, e_j)(\omega), \omega\rangle .$$

If U is an open subset of M on which we have defined a moving orthonormal frame $(e_1, \ldots, e_n)$, it is readily checked that

$$\langle e_i \cdot \omega, \phi\rangle = -\langle \omega, e_i \cdot \phi\rangle, \quad \text{for } \omega, \phi \in \Omega^*(U). \tag{2.55}$$

We can define a map

$$\mathrm{ad}(e_i \cdot e_j) : \Omega^*(U) \to \Omega^*(U) \quad \text{by} \quad \mathrm{ad}(e_i \cdot e_j)(\omega) = e_i \cdot e_j \cdot \omega - \omega \cdot e_i \cdot e_j.$$

It is easy to check that $\mathrm{ad}(e_i \cdot e_j)$ is a derivation with respect to Clifford multiplication.

If $I = (i_1, \ldots, i_k)$, we let

$$e_I = e_{i_1} \cdot e_{i_2} \cdots e_{i_k}.$$

It is easy to check that $\mathrm{ad}(e_i \cdot e_j)$ preserves degree that is, that

$$\mathrm{ad}(e_i \cdot e_j) : \Omega^k(U) \to \Omega^k(U).$$

Moreover,

$$\mathrm{ad}(e_i \cdot e_j)(e_I) = \begin{cases} 0, & \text{if } i, j \in I, \\ 0, & \text{if } i \notin I \text{ and } j \notin I, \\ 2e_i \cdot e_j \cdot e_I, & \text{otherwise,} \end{cases}$$

while if $i \in I$ and $j \notin I$,

$$\mathrm{ad}(e_i \cdot e_j)(e_I) = \pm e_{I \cup \{j\} - \{i\}}.$$

Finally, we check that

$$\mathrm{ad}(e_i \cdot e_j)(\omega) = 0 \quad \text{for all } i, j \quad \Rightarrow \omega \in \Omega^0(U) \oplus \Omega^n(U).$$

Note also the important fact that

$$e_i \cdot e_j \cdot e_I - \frac{1}{2}\mathrm{ad}(e_i \cdot e_j)(e_I) \in \Omega^{k-2}(U) \oplus \Omega^{k+2}(U). \tag{2.56}$$

It follows from (2.56) and (2.55) that if $\omega \in \Omega^k(U)$, where $0 < k < n$,

$$\begin{aligned} \langle \mathcal{K}(\omega), \omega \rangle = \frac{1}{2} \sum_{i \neq j} \langle e_i \cdot e_j \cdot R(e_i, e_j)(\omega), \omega \rangle \\ = -\frac{1}{4} \sum_{i \neq j} \langle R(e_i, e_j)(\omega), \mathrm{ad}(e_i \cdot e_j)(\omega) \rangle . \end{aligned} \tag{2.57}$$

Recall that is terms of components,

$$R(e_i, e_j)e_k = \sum R_{ijlk} e_l = -\sum R_{ijlk} e_l.$$

Lemma. *If $\omega \in \Omega^k(U)$, then*

$$R(e_i, e_j)\omega = -\frac{1}{4} \sum_{k,l=1}^{n} R_{ijkl} \mathrm{ad}(e_k \cdot e_l)(\omega).$$

The proof of this lemma is left as an exercise. The idea is to show that both sides are derivations that agree on $\Omega^0(U)$ and $\Omega^1(U)$. Each of these steps is a straightforward computation.

This lemma together with (2.57) yields

$$\begin{aligned}\langle \mathcal{K}(\omega), \omega\rangle &= \frac{1}{16}\sum_{i,j,k,l} R_{ijkl}\left\langle \mathrm{ad}(e_i\cdot e_j)(\omega), \mathrm{ad}(e_k\cdot e_l)(\omega)\right\rangle \\ &= \frac{1}{16}\sum_{i,j,k,l}\left\langle \mathcal{R}(e_i\wedge e_j)e_k\wedge e_l\right\rangle\left\langle \mathrm{ad}(e_i\cdot e_j)(\omega), \mathrm{ad}(e_k\cdot e_l)(\omega)\right\rangle.\end{aligned}$$

Thus we find that evaluated at any given point $p \in M$,

$$\begin{aligned}\langle \mathcal{K}(\omega), \omega\rangle &= \frac{1}{4}\sum_{i<j,k<l}\left\langle \mathcal{R}(e_i\wedge e_j), e_k\wedge e_l\right\rangle\left\langle \mathrm{ad}(e_i\cdot e_j)(\omega), \mathrm{ad}(e_k\cdot e_l)(\omega)\right\rangle \\ &= \frac{1}{4}\sum_{\alpha,\beta}\left\langle \mathcal{R}(\xi_\alpha), \xi_\beta\right\rangle\left\langle \mathrm{ad}(\xi_\alpha)(\omega), \mathrm{ad}(\xi_\beta)(\omega)\right\rangle,\end{aligned}$$

where the ξ_α's form an orthonormal basis for $\Lambda^2 T_pM$. By a well-known theorem from linear algebra, we can choose such an orthonormal basis so that

$$\langle \mathcal{R}(\xi_\alpha), \xi_\beta\rangle = \lambda_\alpha \delta_{\alpha\beta},$$

and the hypothesis "positive curvature opertors" implies that all of the λ_α's are positive. Then

$$\langle \mathcal{K}(\omega), \omega\rangle = \frac{1}{4}\sum_{\alpha}\lambda_\alpha\left\langle \mathrm{ad}(\xi_\alpha)(\omega), \mathrm{ad}(\xi_\alpha)(\omega)\right\rangle \geq 0,$$

and equals zero only if $\omega = 0$. Thus under the assumption that M has positive curvature operators, there are no harmonic k-forms for $0 < k < n$ and the Gallot-Meyer Theorem is proven.

Chapter 3

Curvature and topology

3.1 The Hadamard-Cartan Theorem

Recall that the curvature is the most important local invariant of a Riemannian or pseudo-Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$. We could imagine defining the curvature at a given point $p \in M$ as follows. First we construct "Riemannian normal coordinates" $(U, x^1, \ldots x^n)$ centered at p as in §1.14. In terms of these coordinates, the Riemannian metric assumes the form

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j,$$

where each component g_{ij} has a Taylor series expansion

$$g_{ij} = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^{n} R_{ikjl}(p) x^k x^l + (\text{higher order terms}),$$

for certain constants $R_{ikjl}(p)$ which satisfy the curvature symmetries:

$$R_{ikjl}(p) = -R_{kijl}(p) = -R_{iklj}(p) = R_{jlik}(p),$$
$$R_{ikjl}(p) + R_{kjil}(p) + R_{jikl}(p) = 0.$$

We can then define a quadrilinear map

$$R : T_pM \times T_pM \times T_pM \times T_pM \to \mathbb{R}$$
$$\text{by} \quad R\left(\left.\frac{\partial}{\partial x^i}\right|_p, \left.\frac{\partial}{\partial x^k}\right|_p, \left.\frac{\partial}{\partial x^j}\right|_p, \left.\frac{\partial}{\partial x^l}\right|_p \right) = R_{ikjl}(p),$$

thereby obtaining an element $R \in \otimes^4 T_p^* M$. As we allow p to vary over M, we thereby obtain a covariant tensor field of rank four,

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{F}(M),$$
$$R(X, Y, Z, W)(p) = R(X(p), Y(p), Z(p), W(p)),$$

which we call the Riemann-Christoffel curvature tensor.

Thus curvature measures the deviation from flatness in the coordinates which are as flat as possible near a given point p.

As we saw in §2.16, we can organize the curvature into a curvature operator

$$\mathcal{R} : \Lambda^2 T_p M \to \Lambda^2 T_p M \quad \text{defined by} \quad \langle \mathcal{R}(x \wedge y), z \wedge w \rangle = \langle R(x,y)w, z \rangle.$$

The curvature symmetries imply that $\mathcal{R}$ is well defined and symmetric:

$$\langle \mathcal{R}(x \wedge y), z \wedge w \rangle = \langle x \wedge y, \mathcal{R}(z \wedge w) \rangle .$$

Hence by theorems from linear algebra, $\Lambda^2 T_p M$ has a basis consisting of eigenvectors for $\mathcal{R}$ and all of the eigenvalues of $\mathcal{R}$ are real. We say that $(M, \langle \cdot, \cdot \rangle)$ has *positive curvature operators* if for every $p \in M$, the eigenvalues of $\mathcal{R}$ are positive and that it has *nonpositive curvature operators* if for every $p \in M$, the eigenvalues of $\mathcal{R}$ are nonpositive.

The Hadamard-Cartan Theorem implies that if $(M, \langle \cdot, \cdot \rangle)$ is a complete Riemannian manifold with nonpositive curvature operators, then for each $p \in M$ the exponential map $\exp_p : T_p M \to M$ is a smooth covering in the following sense. We say that a smooth map $\pi : \tilde{M} \to M$ is a *smooth covering* if π is onto, and each $q \in M$ possesses an open neighborhood U such that $\pi^{-1}(U)$ is a disjoint union of open sets each of which is mapped diffeomorphically by π onto U. Such an open set $U \subset M$ is said to be *evenly covered*.

We say that a connected manifold M is *simply connected* if any path $\gamma : [0,1] \to M$ such that $\gamma(0) = \gamma(1)$ can be continuously deformed to a point. It is a theorem from basic topology as we will see later in §3.5 (or see Chapter 1 of [14]) that a smooth covering of a simply connected space must be a diffeomorphism. Thus the Hadamard-Cartan Theorem implies that a simply connected complete Riemannian manifold with nonpositive curvature operators must be diffeomorphic to Euclidean space. This contrasts with a recent theorem of Böhm and Wilking [6] that compact simply connected Riemannian manifolds with positive curvature operators are diffeomorphic to spheres.

Actually, however, the Hadamard-Cartan Theorem is somewhat stronger. Nonpositive curvature operators can be replaced by the weaker hypothesis that for every $p \in M$,

$$\langle \mathcal{R}(x \wedge y), x \wedge y \rangle \le 0, \quad \text{for all decomposable } x \wedge y \in \Lambda^2 T_p^* M.$$

This is just the assumption that $(M, \langle \cdot, \cdot \rangle)$ has *nonpositive sectional curvatures*, where sectional curvatures are defined as follows: If σ is a two-dimensional subspace of $T_p M$ such that the restriction of $\langle \cdot, \cdot \rangle$ to σ is nondegenerate, the *sectional curvature* of σ is

$$K(\sigma) = \frac{\langle R(x,y)y, x \rangle}{\langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle^2},$$

whenever (x, y) is a basis for σ.

Hadamard-Cartan Theorem. *Let $(M, \langle \cdot, \cdot \rangle)$ be a complete connected n-dimensional Riemannian manifold with nonpositive sectional curvatures. Then the exponential map*

$$exp_p : T_pM \longrightarrow M$$

is a smooth covering.

The reason we can get by with sectional curvatures in the hypothesis is that it is sectional curvatures which governs the behavior of geodesics. In fact, we claim that positive sectional curvatures cause geodesics emanating from a point $p \in M$ to converge, while nonpositive sectional curvatures cause them to diverge. It is this fact which underlies the proof of the Hadamard-Cartan Theorem that we present in the next several sections.

3.2 Parallel transport along curves

Let $(M, \langle \cdot, \cdot \rangle)$ be a pseudo-Riemannian manifold with Levi-Civita connection ∇. Suppose that $\gamma : [a, b] \to M$ is a smooth curve. A *smooth vector field in M along γ* is a smooth function

$$X : [a, b] \to TM \quad \text{such that} \quad X(t) \in T_{\gamma(t)}M \quad \text{for all } t \in [a, b].$$

Note that we can define the *covariant derivative* of such a vector field along γ:

$$\nabla_{\gamma'} X : [a, b] \to TM, \quad (\nabla_{\gamma'} X)(t) \in T_{\gamma(t)}M.$$

If $(x^1, \ldots, x^n)$ are local coordinates in terms of which

$$X(t) = \sum_{i=1}^{n} f^i(t) \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)}, \qquad \gamma'(t) = \sum_{i=1}^{n} \frac{d(x^i \circ \gamma)}{dt}(t) \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)},$$

then a short calculation shows that

$$\nabla_{\gamma'} X(t) = \sum_{i=1}^{n} \left[\frac{df^i}{dt}(t) + \sum_{j,k=1}^{n} \Gamma^i_{jk}(\gamma(t)) \frac{d(x^j \circ \gamma)}{dt}(t) f^k(t) \right] \frac{\partial}{\partial x^i}\bigg|_{\gamma(t)}.$$

Definition. We say that a vector field X along γ is *parallel* if $\nabla_{\gamma'} X \equiv 0$.

Proposition. *If $\gamma : [a, b] \to M$ is a smooth curve, $t_0 \in [a, b]$ and $v \in T_{\gamma(t_0)}M$, then there is a unique vector field X along γ which is parallel along γ and takes the value v at t_0:*

$$\nabla_{\gamma'} X \equiv 0 \quad \text{and} \quad X(t_0) = v. \tag{3.1}$$

Proof: Suppose that in terms of local coordinates,

$$v = \sum_{i=1}^{n} a^i \frac{\partial}{\partial x^i}\bigg|_{\gamma(t_0)}.$$

Then (3.1) is equivalent to the linear initial value problem

$$\frac{df^i}{dt} + \sum_{j,k=1}^{n} \Gamma^i_{jk} \frac{d(x^j \circ \gamma)}{dt} f^k = 0, \quad f^i(t_0) = a^i.$$

It follows from the theory of ordinary differential equations that this initial value problem has a unique solution defined on the interval $[a, b]$.

If $\gamma : [a, b] \to M$ is a smooth path we can define a vector space isomorphism

$$\tau_\gamma : T_{\gamma(a)}M \to T_{\gamma(b)}M \quad \text{by} \quad \tau_\gamma(v) = X(b),$$

where X is the unique vector field along γ which is parallel and satisfies the initial condition $X(a) = v$. Similarly, we can define such an isomorphism τ_γ if γ is only piecewise smooth. We call τ_γ the *parallel transport* along γ.

Note that if X and Y are parallel along γ, then since the Levi-Civita connection ∇ is metric, $\langle X, Y\rangle$ is constant along γ; indeed,

$$\gamma'\langle X, Y\rangle = \langle \nabla_{\gamma'} X, Y\rangle + \langle X, \nabla_{\gamma'} Y\rangle = 0.$$

It follows that τ_γ is an isometry from $T_{\gamma(a)}M$ to $T_{\gamma(b)}M$.

Parallel transport depends very much on the path γ. For example, we could imagine parallel transport on the unit two-sphere $\mathbb{S}^2 \subseteq \mathbb{E}^3$ along the following piecewise smooth geodesic triangle γ: We start at the north pole $n \in \mathbb{S}^2$ and follow the prime meridian to the equator, then follow the equator through θ radians of longitude, and finally follow a meridian of constant longitude back up to the north pole. The resulting isometry from $T_n\mathbb{S}^2$ to itself is then just a rotation through the angle θ.

3.3 Geodesics and curvature

We now consider the differential equation that is generated when we have a "deformation through geodesics."

Suppose that $\gamma : [a, b] \to M$ is a smooth curve and that $\alpha : (-\epsilon, \epsilon) \times [a, b] \to M$ is a smooth map such that $\alpha(0, t) = \gamma(t)$. We can consider the map α as defining a family of smooth curves

$$\bar\alpha(s) : [a, b] \to M, \quad \text{for } s \in (-\epsilon, \epsilon), \quad \text{such that} \quad \bar\alpha(0) = \gamma,$$

if we set $\bar\alpha(s)(t) = \alpha(s, t)$. A *smooth vector field in M along α* is a smooth function

$$X : (-\epsilon, \epsilon) \times [a, b] \to TM$$
$$\text{such that} \quad X(s, t) \in T_{\alpha(s,t)}M \quad \text{for all } (s, t) \in (-\epsilon, \epsilon) \times [a, b].$$

We can take the covariant derivatives $\nabla_{\partial/\partial s}X$ and $\nabla_{\partial/\partial t}X$ of such a vector field along α just as we did for vector fields along curves. (In fact, we already carried out this construction in a special case in §1.6.) If $(x^1,\dots,x^n)$ are local coordinates in terms of which

$$X(s,t) = \sum_{i=1}^n f^i(s,t) \frac{\partial}{\partial x^i}\bigg|_{\alpha(s,t)}$$

$$\text{and we write} \quad \frac{\partial}{\partial s}(s,t) = \frac{\partial \alpha}{\partial s}(s,t) = \sum_{i=1}^n \frac{\partial (x^i \circ \alpha)}{\partial s}(s,t) \frac{\partial}{\partial x^i}\bigg|_{\alpha(s,t)},$$

then a short calculation yields

$$(\nabla_{\partial/\partial s} X)(s,t) = \sum_{i=1}^n \left[\frac{\partial f^i}{\partial s} + \sum_{j,k=1}^n (\Gamma^i_{jk} \circ \alpha) \frac{\partial (x^j \circ \alpha)}{\partial s} f^k \right] (s,t) \, \frac{\partial}{\partial x^i}\bigg|_{\alpha(s,t)}.$$

A similar local coordinate formula can be given for $\nabla_{\partial/\partial t} X$.

Of course, important examples of vector fields along α include

$$\frac{\partial \alpha}{\partial s} \quad \text{and} \quad \frac{\partial \alpha}{\partial t},$$

and it follows quickly from the local coordinate formulae that

$$\nabla_{\partial/\partial s}\left(\frac{\partial \alpha}{\partial t}\right) = \nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial s}\right).$$

Just as in §1.8, the covariant derivatives do not commute, and this failure is described by the curvature: Thus if X is a smooth vector field along α,

$$\nabla_{\partial/\partial s} \circ \nabla_{\partial/\partial t} X - \nabla_{\partial/\partial t} \circ \nabla_{\partial/\partial s} X = R\left(\frac{\partial \alpha}{\partial s}, \frac{\partial \alpha}{\partial t}\right) X.$$

We say that α a *deformation* of γ and call

$$X(t) = \frac{\partial \alpha}{\partial s}(0,t) \in T_{\gamma(t)} M$$

the corresponding *deformation field*.

Proposition 1. *If α is a deformation such that each $\bar{\alpha}(s)$ is a geodesic, then the deformation field X must satisfy Jacobi's equation:*

$$\nabla_{\gamma'} \nabla_{\gamma'} X + R(X, \gamma')\gamma' = 0. \tag{3.2}$$

Proof: Since $\bar{\alpha}(s)$ is a geodesic for every s,

$$\nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial t}\right) \equiv 0 \quad \text{and hence} \quad \nabla_{\partial/\partial s} \nabla_{\partial/\partial t}\left(\frac{\partial \alpha}{\partial t}\right) = 0.$$

By the definition of curvature (see §1.8)

$$\nabla_{\partial/\partial t}\nabla_{\partial/\partial s}\left(\frac{\partial\alpha}{\partial t}\right) + R\left(\frac{\partial\alpha}{\partial s},\frac{\partial\alpha}{\partial t}\right)\frac{\partial\alpha}{\partial t} = 0$$
$$\text{or}\quad \nabla_{\partial/\partial t}\nabla_{\partial/\partial t}\left(\frac{\partial\alpha}{\partial s}\right) + R\left(\frac{\partial\alpha}{\partial s},\frac{\partial\alpha}{\partial t}\right)\frac{\partial\alpha}{\partial t} = 0.$$

Evaluation at $s = 0$ now yields (3.2), finishing the proof.

Remark. The Jacobi equation can be regarded as the linearization of the geodesic equation near a given geodesic γ.

Definition. A vector field X along a geodesic γ which satisfies the Jacobi equation (3.2) is called a *Jacobi field*.

Suppose that γ is a unit speed geodesic and that $(E_1, \ldots E_n)$ are parallel orthonormal vector fields along γ such that $E_1 = \gamma'$. We can then define the component functions of the curvature with respect to $(E_1, \ldots E_n)$ by

$$R(E_k, E_l)E_j = \sum_{i=1}^{n} R^i_{klj} E_i,$$

where our convention is that the upper index i gets lowered to the third position. If $X = \sum f^i E_i$, then the Jacobi equation becomes

$$\frac{d^2 f^i}{dt^2} + \sum_{j=1}^{n} R^i_{j11} f^j = 0. \tag{3.3}$$

This second order linear system of ordinary differential equations will possess a $2n$-dimensional vector space of solutions along γ. The Jacobi fields which vanish at a given point will form a linear subspace of dimension n.

Example. Suppose that $(M, \langle\cdot,\cdot\rangle)$ is a complete Riemannian manifold of constant sectional curvature k. Thus

$$R(X,Y)W = k[\langle Y, W\rangle X - \langle X, W\rangle Y].$$

In this case, X will be a Jacobi field if and only if

$$\nabla_{\gamma'}\nabla_{\gamma'}X = -R(X,\gamma')\gamma' = k[\langle X\gamma'\rangle\gamma' - \langle\gamma',\gamma'\rangle X].$$

Equivalently, if we assume that γ is unit speed and write $X = \sum f^i E_i$, where $(E_1, \ldots, E_n)$ is a parallel orthonormal frame along γ such that $E_1 = \gamma'$, then

$$\begin{cases} \frac{d^2 f^1}{dt^2} = 0, \\ \frac{d^2 f^i}{dt^2} = -k f^i, \quad \text{for } 2 \le i \le n. \end{cases}$$

The solutions are

$$f^1(t) = a^1 + b^1 t,$$

and for $2 \leq i \leq n$,

$$f^i(t) = \begin{cases} a^i \cos(\sqrt{k}t) + b^i \sin(\sqrt{k}t), & \text{for } k > 0, \\ a^i + b^i t, & \text{for } k = 0, \\ a^i \cosh(\sqrt{-k}t) + b^i \sinh(\sqrt{-k}t), & \text{for } k < 0, \end{cases} \tag{3.4}$$

Here $a^1, b^1, \ldots, a^n, b^n$ are constants of integration to be that are determined by the initial conditions.

Definition. Suppose that $\gamma : [a, b] \to M$ is a geodesic in a pseudo-Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ with $\gamma(a) = p$ and $\gamma(b) = q$. We say that p and q are *conjugate* along γ if $p \neq q$ and there is a nonzero Jacobi field X along γ such that $X(a) = 0 = X(b)$.

For example, antipodal points on $\mathbb{S}^n$ are conjugate along the great circle geodesics which join them, while $\mathbb{E}^n$ and $\mathbb{H}^n$ do not have any conjugate points.

Suppose that p is a point in a geodesically complete pseudo-Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ and $v \in T_pM$. We can then define a geodesic $\gamma_v : [0, 1] \to M$ by $\gamma_v(t) = \exp_p(tv)$. We say that v belongs to the *conjugate locus* in T_pM if $\gamma_v(0)$ and $\gamma_v(1)$ are conjugate along γ_v.

Proposition 2. *A vector $v \in T_pM$ belongs to the conjugate locus if and only if $(\exp_p)_*$ is singular at v, that is, there is a nonzero vector $w \in T_v(T_pM)$ such that $(\exp_p)_{*v}(w) = 0$.*

Proof: We use the following construction: If $w \in T_v(T_pM)$, we define

$$\alpha_w : (\epsilon, \epsilon) \times [0, 1] \to M \quad \text{by} \quad \alpha_w(s, t) = \exp_p(t(v + sw)).$$

We set

$$X_w(t) = \frac{\partial \alpha_w}{\partial s}(0, t),$$

A Jacobi field along γ_v which vanishes at $\gamma_v(0)$. As w ranges throughout T_pM, X_w rnges throughout the n-dimensional space of Jacobi fields along γ_v which vanish at $\gamma_v(0)$.

$\Leftarrow$: If $(\exp_p)_{*v}(w) = 0$, then X_w is a Jacobi field along γ_v which vanishes at $\gamma_v(0)$ and $\gamma_v(1)$, so v belongs to the conjugate locus.

$\Rightarrow$: If v belongs to the conjugate locus, there is a Jacobi field along γ_v which vanishes at $\gamma_v(0)$ and $\gamma_v(1)$, and this vector field must be of the form X_w for some $w \in T_v(T_pM)$. But then $(\exp_p)_{*v}(w) = X_w(1) = 0$, and hence $(\exp_p)_*$ is singular at v.

Example. Let us consider the n-sphere $\mathbb{S}^n$ of constant curvature one. If p is the north pole in $\mathbb{S}^n$, it follows from (3.4) that the conjugate locus in $T_p\mathbb{S}^n$ is a family of concentric spheres of radius $k\pi$, where $k \in \mathbb{N}$.

3.4 Proof of the Hadamard-Cartan Theorem

Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold. A point $p \in M$ is said to be a *pole* if the conjugate locus in T_pM is empty. For example, it follows from the explicit formulae we derived for Jacobi fields that any point in Euclidean space $\mathbb{E}^n$ or hyperbolic space $\mathbb{H}^n$ is a pole.

Pole Theorem. *If $(M, \langle \cdot, \cdot \rangle)$ is a complete connected Riemannian manifold and $p \in M$ is a pole, then $\exp_p : T_pM \to M$ is a smooth covering.*

To prove this, we need to show that π is onto and each $p \in M$ has an open neighborhood U such that $\pi^{-1}(U)$ is the disjoint union of open sets, each of which is mapped diffeomorphically by π onto U.

Since $\exp_p$ is nonsingular at every $v \in T_pM$, we can define a Riemannian metric $\langle\langle \cdot, \cdot \rangle\rangle$ on T_pM by

$$\langle\langle x, y \rangle\rangle = \langle (\exp_p)_*(x), (\exp_p)_*(y) \rangle, \quad \text{for all } x, y \in T_v(T_pM).$$

Locally, $\exp_p$ is an isometry from $(T_pM, \langle\langle \cdot, \cdot \rangle\rangle)$ to $(M, \langle \cdot, \cdot \rangle)$ and it takes lines through the origin in T_pM to geodesics through $p \in M$. Hence lines through the origin must be geodesics in the Riemannian manifold $(T_pM, \langle\langle \cdot, \cdot \rangle\rangle)$. It therefore follows from the Hopf-Rinow Theorem from §1.18 that $(T_pM, \langle\langle \cdot, \cdot \rangle\rangle)$ is complete. Thus the theorem will follow from the following lemma:

Lemma. *If $\pi : \tilde{M} \to M$ is a local isometry of connected Riemannian manifolds with $\tilde{M}$ complete, then π is a smooth covering.*

Proof of lemma: Let $q \in M$. We need to show that q lies in an open set $U \subseteq M$ which is evenly covered, i.e. that $\pi^{-1}(U)$ is a disjoint union of open sets each of which is mapped diffeomorphically onto U.

There exists $\epsilon > 0$ such that $\exp_q$ maps the open ball of radius 2ϵ in T_qM diffeomorphically onto $\{r \in M : d(q, r) < 2\epsilon\}$. Let $\{\tilde{q}_\alpha : \alpha \in A\}$ be the set of points in $\tilde{M}$ which are mapped by π to q, and let

$$U = \{r \in M : d(r, q) < \epsilon\}, \qquad \tilde{U}_\alpha = \{\tilde{r} \in \tilde{M} : d(\tilde{r}, \tilde{q}_\alpha) < \epsilon\}.$$

Choose a point $\tilde{q}_\alpha \in \pi^{-1}(q)$, and let

$$B_\epsilon = \{v \in T_qM : \|v\| < \epsilon\}, \quad \tilde{B}_\epsilon = \{\tilde{v} \in T_{\tilde{q}_\alpha}\tilde{M} : \|\tilde{v}\| < \epsilon\}.$$

We then have a commutative diagram

$$\begin{array}{ccc} \tilde{B}_\epsilon & \xrightarrow{\pi_*} & B_\epsilon \\ {\scriptstyle \exp_{\tilde{q}_\alpha}}\downarrow & & {\scriptstyle \exp_q}\downarrow \\ \tilde{U}_\alpha & \xrightarrow{\pi} & U \end{array}$$

Note that $\exp_{\tilde{q}_\alpha}$ is globally defined and maps onto $\tilde{U}_\alpha$ because $\tilde{M}$ is complete, and π_* and $\exp_q$ are diffeomorphisms. Hence π maps $\tilde{U}_\alpha$ diffeomorphically onto U.

If $\tilde{r} \in \tilde{U}_\alpha \cap \tilde{U}_\beta$, we would have geodesics $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\beta$ of length $< \epsilon$ from $\tilde{q}_\alpha$ and $\tilde{q}_\beta$ to $\tilde{r}$. These would project to geodesics γ_α and γ_β of length $< \epsilon$ from q to $r = \pi(\tilde{r})$. By uniqueness of geodesics in normal coordinate charts, we would have $\gamma_\alpha = \gamma_\beta$. Since π is a local isometry, $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\beta$ would satisfy the same initial conditions at $\tilde{r}$. Thus $\tilde{\gamma}_\alpha = \tilde{\gamma}_\beta$, so $\tilde{q}_\alpha = \tilde{q}_\beta$ and $\alpha = \beta$. We have shown that $\tilde{U}_\alpha \cap \tilde{U}_\beta \neq \emptyset$ only if $\alpha = \beta$.

Suppose now that $\tilde{r} \in \pi^{-1}(U)$, with $r = \pi(\tilde{r}) \in U$. Then there is a unit-speed geodesic γ from r to q of length $< \epsilon$. There is a unit-speed geodesic $\tilde{\gamma}$ in $\tilde{M}$ starting from $\tilde{r}$ whose initial conditions project to those of γ. Then $\pi \circ \tilde{\gamma} = \gamma$ and hence $\tilde{\gamma}$ proceeds from $\tilde{r}$ to $\tilde{q}_\alpha$ in time $< \epsilon$ for some $\alpha \in A$. Thus $\tilde{r} \in \tilde{U}_\alpha$ for some $\alpha \in A$, and

$$\pi^{-1}(U) = \bigcup \{\tilde{U}_\alpha : \alpha \in A\}.$$

Thus every point in M lies in an open set which is evenly covered. One easily checks that $\pi(\tilde{M})$ is both open and closed in M. Since M is connected, π is surjective and the lemma is proven.

The Hadamard-Cartan Theorem now follows from the following:

Theorem. *If $(M, \langle \cdot, \cdot \rangle)$ is a complete connected Riemannian manifold which has the property that its curvature R satisfies the condition*

$$\langle R(x,y)y, x \rangle \leq 0, \quad \text{for all } x, y \in T_qM \text{ and all } q \in M.$$

Then any point $p \in M$ is a pole.

Proof: Suppose that $p \in M$, $v \in T_pM$ and $\gamma(t) = \exp_p(tv)$. We need to show that $p = \gamma(0)$ and $q = \gamma(1)$ are not conjugate along γ.

Suppose, on the contrary, that X is a nonzero Jacobi field along γ which vanishes at $\gamma(0)$ and $\gamma(1)$. Thus

$$\nabla_{\gamma'} \nabla_{\gamma'} X + R(X, \gamma')\gamma' = 0, \quad \langle \nabla_{\gamma'} \nabla_{\gamma'} X, X \rangle = -\langle R(X, \gamma')\gamma', X \rangle \geq 0.$$

Hence

$$\int_0^1 \langle \nabla_{\gamma'} \nabla_{\gamma'} X, X \rangle dt \geq 0,$$

and integrating by parts yields

$$\int_0^1 \left[\frac{d}{dt} \langle \nabla_{\gamma'} X, X \rangle - \langle \nabla_{\gamma'} X, \nabla_{\gamma'} X \rangle \right] dt \geq 0,$$

and since the first term integrates to zero, we obtain

$$\int_0^1 -\langle \nabla_{\gamma'} X, \nabla_{\gamma'} X \rangle dt \geq 0.$$

It follows that $\nabla_{\gamma'} X \equiv 0$ and hence X is identically zero, a contradiction.

3.5 The fundamental group

This section gives a brief treatment of the notion of fundamental group, a topic familiar to many readers; our treatment is brief since this topic is treated in detail in Math 221B. If you have not seen the fundamental group before, focus first on the definitions of fundamental group, simply connected and universal covers, and return to the proofs after you see how these concepts are used. A detailed treatement of the fundamental group can be found in Chapter 1 of [14], which is available on the internet.

3.5.1 Definition of the fundamental group

Suppose that X is a metrizable topological space and that x_0 and x_1 are points of X. We let

$$P(X; x_0, x_1) = \{ \text{ continuous paths } \gamma : [0, 1] \to X : \gamma(0) = x_0, \gamma(1) = x_1\}.$$

If $\gamma, \lambda \in P(X, x_0, x_1)$ we say that γ and λ are *homotopic* relative to the endpoints $\{0, 1\}$ and write $\gamma \simeq \lambda$ if there is a continuous map $\alpha : [0, 1] \times [0, 1] \to X$ such that

$$\alpha(x, 0) = x_0, \quad \alpha(s, 1) = x_1, \quad \alpha(0, t) = \gamma(t), \quad \alpha(1, t) = \lambda(t).$$

We let $\pi_1(X, x_0, x_1)$ denote the quotient space of $P(X, x_0, x_1)$ by the equivalence relation defined by $\simeq$, and if $\gamma \in P(X, x_0, x_1)$, we let $[\gamma] \in \pi_1(X; x_0, x_1)$ denote the corresponding equivalence class. If d is a metric defining the topology on X, we define a metric on $P(X; x_0, x_1)$ by

$$d(\gamma, \lambda) = \sup\{d(\gamma(t), \lambda(t)) : t \in [0, 1]\},$$

then $P(X; x_0, x_1)$ becomes a metric space itself and it has a resulting topology. In this case, $\pi_1(X, x_0, x_1)$ can be regarded as the collection of path components of $P(X; x_0, x_1)$.

Suppose that $\gamma \in P(X; x_0, x_1)$ and $\lambda \in P(X; x_1, x_2)$, and define $\gamma \cdot \lambda \in P(X; x_0, x_2)$ by

$$(\gamma \cdot \lambda)(t) = \begin{cases} \gamma(2t), & \text{for } t \in [0, 1/2], \\ \lambda(2t - 1), & \text{for } t \in [1/2, 1]. \end{cases}$$

Finally, if $[\gamma] \in \pi_1(X; x_0, x_1)$ and $[\lambda] \in \pi_1(X; x_1, x_2)$, we claim that we can define a product $[\gamma][\lambda] = [\gamma \cdot \lambda] \in \pi_1(X; x_0, x_2)$. We need to show that this product

$$\pi_1(X; x_0, x_1) \times \pi_1(X; x_1, x_2) \longrightarrow \pi_1(X; x_0, x_2)$$

is well-defined; in other words, if $\gamma \simeq \tilde{\gamma}$ and $\lambda \simeq \tilde{\lambda}$, then

$$\gamma \cdot \lambda \simeq \tilde{\gamma} \cdot \tilde{\lambda}.$$

We show that $\gamma \simeq \tilde{\gamma}$ implies that $\gamma \cdot \lambda \simeq \tilde{\gamma} \cdot \lambda$. If α is the homotopy from γ to $\tilde{\gamma}$, we define

$$\beta : [0,1] \times [0,1] \to X \quad \text{by} \quad \beta(s.t) = \begin{cases} \alpha(s, 2t), & \text{for } t \in [0, 1/2], \\ \lambda(2t-1), & \text{for } t \in [1/2, 1]. \end{cases}$$

This gives the required homotopy from $\gamma \cdot \lambda$ to $\tilde{\gamma} \cdot \lambda$. The fact that $\lambda \simeq \tilde{\lambda}$ implies that $\gamma \cdot \lambda \simeq \gamma \cdot \tilde{\lambda}$ is quite similar.

The case where $x_0 = x_1$ is particularly important. We denote $\pi_1(X, x_0, x_0)$ by $\pi_1(X, x_0)$, and call it the *fundamental group* of X at x_0.

Theorem. *The multiplication operation defined above makes $\pi_1(X, x_0)$ into a group.*

To prove this we must first show that $\pi_1(X, x_0)$ has an identity. We let ε be the constant path, $\varepsilon(t) = x_0$ for all $t \in [0,1]$, and claim that

$$[\gamma \cdot \varepsilon] = [\gamma] = [\varepsilon \cdot \gamma], \quad \text{for all } [\gamma] \in \pi_1(X, x_0).$$

We prove the first equality, the other being similar; to do this, we need to construct a homotopy from $\gamma \cdot \varepsilon$ to γ. We simply define $\alpha : [0,1] \times [0,1] \to X$ by

$$\alpha(s,t) = \begin{cases} \gamma\left(\frac{2t}{s+1}\right), & \text{for } t \leq (1/2)(s+1), \\ x_0, & \text{for } t \geq (1/2)(s+1). \end{cases}$$

Then $\alpha(0,t) = (\gamma \cdot \varepsilon)(t)$ and $\alpha(1,t) = \gamma(t)$.

To prove associativity of multiplication, we need to show that if γ, λ and μ are elements of $P(X, x_0)$, the

$$(\gamma \cdot \lambda) \cdot \mu \simeq \gamma \cdot (\lambda \cdot \mu).$$

To do this, we define $\alpha : [0,1] \times [0,1] \to X$ by

$$\alpha(s,t) = \begin{cases} \gamma\left(\frac{4t}{s+1}\right), & \text{for } t \leq (1/4)(s+1), \\ \lambda(4t - s - 1), & \text{for } (1/4)(s+1) \leq t \leq (1/4)(s+2), \\ \mu\left(\frac{4t-s-2}{2-t}\right), & \text{for } t \geq (1/4)(s+2). \end{cases}$$

Then $\alpha(0,t) = (\gamma \cdot \lambda) \cdot \mu$ while $\alpha(1,t) = \gamma \cdot (\lambda \cdot \mu)$, so multiplication is indeed associative.

Finally, given $\gamma \in P(X, x_0)$, we define $\gamma^{-1}(t) = \gamma(1-t)$. To prove that $[\gamma^{-1}]$ is the inverse to $[\gamma]$, we must show that

$$\gamma \cdot \gamma^{-1} \simeq \varepsilon \quad \text{and} \quad \gamma^{-1} \cdot \gamma \simeq \varepsilon.$$

For the first of these, we define a homotopy $\alpha : [0,1] \times [0,1] \to X$ by

$$\alpha(s,t) = \begin{cases} \gamma(2t), & \text{for } t \leq (1/2)(1-s), \\ \gamma(1-s), & \text{for } (1/2)(1-s) \leq t \leq (1/2)(1+s), \\ \gamma(2-2t), & \text{for } t \geq (1/2)(1+s). \end{cases}$$

Then $\alpha(0,t) = (\gamma \cdot \gamma^{-1})(t)$ while $\alpha(s,1) = x_0$. The homotopy for $\gamma^{-1} \cdot \gamma \simeq \varepsilon$ is constructed in a similar fashion.

A continuous map $F : X \to Y$ with $F(x_0) = y_0$ induces a map

$$F_\sharp : P(X,x_0) \to P(Y,y_0) \quad \text{by} \quad F_\sharp(\gamma) = F \circ \gamma,$$

and it is easily checked that

$$\gamma \simeq \lambda \quad \Rightarrow \quad F \circ \gamma \simeq F \circ \lambda,$$

so that $F_\sharp$ induces a set-theoretic map

$$F_\sharp : \pi_1(X,x_0) \to \pi_1(Y,y_0).$$

Moreover, it is immediate that $F_\sharp$ is in fact a group homomorphism. We thus obtain a covariant function from the catergory of pointed metrizable topological spaces (X,x_0) and continuous maps $F : (X,x_0) \to (Y,y_0)$ preserving base points to the category of groups and group homomorphisms.

Remark. If X is pathwise connected, the fundamental groups based at different points are isomorphic. This is proven by techniques similar to those utilized in the proof of the preceding theorem. Indeed, if $\gamma \in P(X,x_0,x_1)$, we can define a map $h_\gamma : P(X,x_0) \to P(X,x_1)$ by

$$h_\gamma(\lambda) = \begin{cases} \gamma(1-3t), & \text{for } t \in [0,1/3], \\ \lambda(3t-1), & \text{for } t \in [1/3,2/3], \\ \gamma(3t-2), & \text{for } t \in [2/3,1]. \end{cases}$$

By arguments similar to those used in the proof of the preceding theorem, one checks that this yields a well-defined group homomorphism

$$h_{[\gamma]} : \pi_1(X,x_0) \to \pi_1(X,x_1) \quad \text{by} \quad h_{[\gamma]}([\lambda]) = [h_\gamma \lambda].$$

Finally, if $\gamma^{-1} \in P(X,x_1,x_0)$ is defined by $\gamma^{-1}(t) = \gamma(1-t)$, one checks that $h_{[\gamma^{-1}]}$ is an inverse to $h_{[\gamma]}$.

Definition. We say that a metrizable topological space X is *simply connected* if it is pathwise connected and $\pi_1(X,x_0) = 0$. (The above remark shows that this condition does not depend on the choice of base point x_0.

3.5.2 Homotopy lifting

To calculate the fundamental groups of spaces, one often uses the notion of covering space. A continuous map $\pi\tilde{X} \to X$ is a *covering* if it is onto and every $x \in X$ lies in an open neighborhood U such that $\pi^{-1}(U)$ is a disjoint union of open sets each of which is mapped homeomorphically by π onto U. Such an open set is said to be *evenly covered*. Coverings have two important useful properties:

Homotopy Lifting Theorem. *Suppose that $\pi : \tilde{X} \to X$ is a covering. If*

$$\tilde{\gamma} : [0,1] \to \tilde{X}, \qquad \alpha : [0,1] \times [0,1] \to X$$

are continuous maps such that $\pi(\tilde{\gamma})((t) = \alpha(0,t)$, thene there exists a continuous map

$$\tilde{\alpha} : [0,1] \times [0,1] \to \tilde{X}$$

such that $\tilde{\alpha}(0,t) = \tilde{\gamma}(t)$ and $\pi \circ \tilde{\alpha} = \alpha$.

In words, the homotopy α can be lifted to $\tilde{\alpha}$ taking values in $\tilde{X}$.

To prove this, we let $\mathcal{U}$ be an open cover of $[01,] \times [0,1] \subseteq \mathbb{R}^2$ consisting of open sets of the form $(a,b) \times (c,d)$ where a, b, c, d are rational and

$$\alpha((a,b) \times (c,d) \cap [0,1] \times [0,1])$$

lies in an evenly covered open subset of M. Since $[0,1] \times [0,1]$ is compact, a finite subcollection

$$\{(a_1, b_1 \times (c_1, d_1), \ldots, (a_k, b_k) \times (c_k, d_k)\}$$

of $\mathcal{U}$ covers $[0,1] \times [0,1]$. Choose a positive integer m so that $ma_1, \ldots, md_k$ are all integers and let $n = 2m$. For $1 \leq i, j \leq n$, let

$$D_{ij} = \left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right].$$

Then $\alpha(D_{ij})$ is contained in an evenly covered open subset U_{ij} of X.

The idea now is to define $\tilde{\alpha}$ inductively on D_{11}, D_{12}, …,D_{1n}, D_{21}, …, D_{2n}, …, D_{n1}, …, D_{nn}.

When we get to the (i,j)-stage, $\tilde{\alpha}$ is already defined on a connected part of the bounary of D_{ij} and the image lies in some $\tilde{U}_{ij}$ which is mapped homeomorphically onto an evenly covered open subset U_{ij} of X. We are forced to define $\tilde{\alpha}|D_{ij}$ by

$$\tilde{\alpha}|D_{ij} = (\pi|\tilde{U}_{ij})^{-1} \circ H|\alpha_{ij}.$$

This gives the unique extension of $\tilde{\alpha}$ to D_{ij} and an induction on i and j then finishes the proof of the Unique Path Lifting Theorem.

Remark. In the Homotopy Lifting Theorem, we could consider the case of a degenerate path $\tilde{\gamma}(t) \equiv \tilde{p}$ and a degenerate homotopy $\alpha(s,t) = \lambda(s)$. In this case, the Homotopy Lifting Theorem gives rise to a existence of a path $\tilde{\lambda}$ covering a given path λ in X. The following theorem shows that this lifted path is unique:

Unique Path Lifting Theorem. *Suppose that $\pi : \tilde{X} \to X$ is a covering. If $\gamma, \lambda : [0,1] \to \tilde{X}$ are two continuous maps such that $\gamma(0) = \lambda(0)$ and $\pi \circ \gamma = \pi \circ \lambda$, then $\gamma = \lambda$.*

To prove this, we let $J = \{t \in [0,1] : \gamma(t) = \lambda(t)\}$. We claim that J is both open and closed. Indeed, if $t \in J$, $\gamma(t) = \lambda(t)$ and $\gamma(t) = \lambda(t)$ lies in some open set $\tilde{U}$ which is mapped homeomorphically onto an open set U in X. Clearly $t \subseteq \gamma^{-1}(\tilde{U}) \cap \lambda^{-1}(\tilde{U}) \subseteq J$. Hence J is open.

On the other hand, if $t \in J$, there exist open sets $\tilde{U}_1$ and $\tilde{U}_2$ such that $\gamma(t) \in \tilde{U}_1$ and $\lambda(t) \in \tilde{U}_2$, where the two sets $\tilde{U}_1$ and $\tilde{U}_2$ are disjoint open sets mapped homemorphically by π onto an open subset U of X. Thus

$$t \in \gamma^{-1}(\tilde{U}_1) \cap \lambda^{-1}(\tilde{U}_2) \subseteq [0,1] - J,$$

and J is closed.

Example. Suppose that

$$\pi : \mathbb{R} \to \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\} \quad \text{by} \quad \pi(t) = e^{2\pi i t}. \tag{3.5}$$

One checks that π is a smooth covering. We can use the previous theorems to calculate the fundamental group $\pi_1(\mathbb{S}^2, 1)$.

Indeed, suppose that $\gamma \in P(S^1, 1)$. Then the Unique Path Lifting Theorem implies that there is a unique $\tilde{\gamma} : [0,1] \to \mathbb{R}$ such that $\tilde{\gamma}(0) = 0$ and $\pi \circ \tilde{\gamma} = \gamma$. Since $\gamma(1) = 1$, there exists an element $k \in \mathbb{Z}$ such that $\tilde{\gamma}(1) = k$. If $\gamma \simeq \lambda \in P(\mathbb{S}^2, 1)$ by means of a homotopy $\alpha : [0,1] \times [0,1] \to \mathbb{S}^1$, we can use the Homotopy Lifting Theorem to construct $\tilde{\alpha} : [0,1] \times [0,1] \to \mathbb{R}$ such that $\alpha(0,t) = \tilde{\gamma}(t)$ and $\pi \circ \tilde{\alpha} = \alpha$. Unique path lifting implies that $\tilde{\alpha}(s,0) = 0$, $\tilde{\alpha}(s,1) = k$. and $\tilde{\alpha}(1,t) = \tilde{\lambda}(t)$. Thus $\tilde{\gamma}(1) = \tilde{\lambda}(1)$, and we obtain a well-defined map

$$h : \pi_1(\mathbb{S}^1, 1) \to \mathbb{Z} \quad \text{such that} \quad h([\gamma]) = \tilde{\gamma}(1).$$

It is easily checked that h is a homomorphism and since $h([e^{2\pi k i t}]) = k$ for $k \in \mathbb{Z}$, we see that h is surjection. Finally, if $h([\gamma]) = 0$, then $\tilde{\gamma}(1) = 0$ and hence $\tilde{\gamma}$ is homotopic to a constant, and γ itself must be homotopic to a constant. Thus we conclude that $\pi_1(\mathbb{S}^2, 1) \cong \mathbb{Z}$.

Degree of maps from S^1 to S^1: Suppose that $F : S^1 \to S^1$ is a continuous map. Then $\gamma : S^1 \to S^1$ determines a homomorphism of fundamental groups

$$\gamma_\sharp : \pi_1(S^1) \to \pi_1(S^1),$$

and since $\pi_1(S^1) \cong \mathbb{Z}$, this group homomorphism must be multiplication by some integer $n \in \mathbb{Z}$. We set $\deg(\gamma) = n$ and call it the *degree* of γ.

Regarding S^1 as

$$S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

we note that the differential form $xdy - ydx$ is closed but not exact. However, if π is the covering (3.5), then $\pi^*(ydx - xdy) = d\theta$ for some globallly defined real-valued function θ on $\mathbb{R}$. If $\gamma : S^1 \to S^1$ is smooth, then γ lifts to a smooth map $\tilde{\gamma} : S^1 \to \mathbb{R}$ and

$$\deg(\gamma) = \frac{1}{2\pi} \int_{\tilde{\gamma}} d\theta = \frac{1}{2\pi} \int_{\gamma} (xdy - ydx).$$

By very similar arguments, one could calculate the fundamental groups of many other spaces. For example, if $T^n = \mathbb{E}^n/\mathbb{Z}^n$, the usual n-torus, then

$$\pi_1(T^n, x_0) = \overbrace{\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}}^{n},$$

while if $\mathbb{R}P^n$ is the real projective space obtained by identifying antipodal points on $\mathbb{S}^n(1)$, then $\pi_1(\mathbb{R}P^n, x_0) = \mathbb{Z}_2$.

Finally, the Homotopy Lifting Theorem allows us to finish the argument that a complete simply connected Riemannian manifold which has nonpositive sectional curvatures must be diffeomorphic to $\mathbb{R}^n$:

Covering Theorem. *Suppose that $\pi : \tilde{M} \to M$ is a smooth covering, were $\tilde{M}$ and M are pathwise connected. If M is simply connected, then π is a diffeomorphism.*

To prove this we need only show that π is one-to-one. Suppose that $\tilde{p}$ and $\tilde{q}$ are points in $\tilde{M}$ such that $\pi(\tilde{p}) = \pi(\tilde{q})$. Since $\tilde{M}$ is pathwise connected, there is a continuous path $\tilde{\gamma} : [0,1] \to \tilde{M}$ such that $\tilde{\gamma}(1) = \tilde{p}$ and $\tilde{\gamma}(1) = \tilde{q}$. Let $\gamma = \pi \circ \tilde{\gamma}$. If $p = \pi(\tilde{p}) = \pi(\tilde{q})$, then $\gamma \in P(M,p)$. Since M is simply connected there is a continuous map $\alpha : [0,1] \times [0,1] \to M$ such that

$$\alpha(s,0) = p = \alpha(s,1), \quad \alpha(0,t) = \gamma(t), \quad \alpha(1,t) = p.$$

By the Homotopy Lifting Theorem, there is a continuous map $\tilde{\alpha} : [0,1] \times [0,1] \to \tilde{M}$ such that

$$\tilde{\alpha}(0,t) = \tilde{\gamma}(t), \qquad \pi \circ \tilde{\alpha} = \alpha.$$

The Unique Path Lifting Theorem implies that

$$\tilde{\alpha}(s,0) = \tilde{\alpha}(0,0) = \tilde{\gamma}(0) = \tilde{p} \quad \text{and} \quad \tilde{\alpha}(s,1) = \tilde{\alpha}(0,1) = \tilde{\gamma}(1) = \tilde{q}.$$

On the other hand, the Unique Path Lifting Theorem also implies that $\tilde{\alpha}(1,t)$ is constant. Hence $\tilde{p} = \tilde{q}$ and π is indeed one-to-one, exactly what we wanted to prove.

3.5.3 Universal covers

The final fact we need regarding the fundamental group and covering spaces is the existence of a universal cover.

Universal Cover Theorem. *If M is a connected smooth manifold, there exists a simply connected smooth manifold $\tilde{M}$ together with a smooth covering $\pi : \tilde{M} \to M$. Moreover, if $\tilde{M}_1$ is another simply connected smooth manifold with with smooth covering $\pi_1 : \tilde{M}_1 \to M$, there exists a smooth diffeomorphism $T : \tilde{M} \to \tilde{M}_1$ such that $\pi = \pi_1 \circ T$.*

We sketch the argument. A complete proof can be found in [14].

We start with a point $p_0 \in M$ and let

$$\tilde{M} = \{(p, [\gamma]) : p \in M, [\gamma] \in \pi_1(M; p_0, p)\}.$$

We can then define $\pi : \tilde{M} \to M$ by $\pi(p, [\gamma]) = p$. We need to define a metrizable topology on $\tilde{M}$, check that π is a covering and show that $\tilde{M}$ is simply connected. Then $\tilde{M}$ inherits a unique smooth manifold structure such that π is a local diffeomorphism.

Here is the idea for constructing the topology: Suppose that $\tilde{p} = (p, [\gamma]) \in \tilde{M}$ and let U be a contractible neighborhood of p within M. We then let

$$\tilde{U}_{(p,[\gamma])} = \{(q, [\lambda]) \in \tilde{M} : q \in U, [\lambda] = [\gamma \cdot \alpha], \text{ where } \alpha \text{ lies entirely within } U\ \}.$$

Then π maps $\tilde{U}_{(p,[\gamma])}$ homeomorphically onto U. From this one concludes that π is a covering.

To show that $\tilde{M}$ is simply connected, we suppose that $\tilde{\lambda} : [0, 1] \to \tilde{M}$ is a continuous path with

$$\tilde{\lambda}(0) = \tilde{\lambda}(1) = (p_0, [\varepsilon]),$$

where ε is the constant path at p_0. Then $\lambda = \pi \circ \tilde{\lambda}$ is a closed curve from p_0 to p_0. For $t \in [0, 1]$, we define

$$\lambda_t : [0, 1] \to M \quad \text{by} \quad \lambda_t(s) = \lambda(st),$$

and define

$$\hat{\lambda} : [0, 1] \to \tilde{M} \quad \text{by} \quad \hat{\lambda}(t) = (\lambda(t), [\lambda_t]).$$

Then $\hat{\lambda}(0) = (p_0, [\varepsilon])$ and $\pi \circ \hat{\lambda} = \lambda$. By the Unique Path Lifting Theorem, $\hat{\lambda} = \tilde{\lambda}$. But

$$\hat{\lambda}(1) = \tilde{\lambda}(1) \quad \Rightarrow \quad (p_0, [\lambda]) = (p_0, [\varepsilon]) \quad \Rightarrow \quad [\lambda] = [\varepsilon].$$

Thus λ is homotopic to a constant in M and by the Homotopy Lifting Theorem, $\tilde{\lambda}$ is homotopic to a constant in $\tilde{M}$.

If $\tilde{M}_1$ is another simply connected smooth manifold with with smooth covering $\pi_1 : \tilde{M}_1 \to M$, we choose $\tilde{p}_1 \in \tilde{M}_1$ such that $\pi_1(\tilde{p}_1) = p_0$. We then define

$$T : \tilde{M} \to \tilde{M}_1 \quad \text{by} \quad T(p, [\gamma]) = \tilde{\gamma}(1),$$

where $\tilde{\gamma} : [0, 1] \to \tilde{M}_1$ is the unique lift of γ such that $\tilde{\gamma}(0) = \tilde{p}_1$. It is then relatively straightforward to check that T is a diffeomorphism such that $\pi = \pi_1 \circ T$.

Definition. If M is a connected smooth manifold and $\pi : \tilde{M} \to M$ is a smooth covering with $\tilde{M}$ simply connected, we say that $\pi : \tilde{M} \to M$ (or sometime $\tilde{M}$ itself) is the *universal cover* of M.

Note that the above construction of the universal cover shows that the elements of the fundamental group of M correspond in a one-to-one fashion with the elements of $\pi^{-1}(p_0)$. One can also show that there is a one-to-one correspondence between elements of the fundamental group of M and the group of deck transformations of the universal cover $\pi : \tilde{M} \to M$, where a *deck transformation* is a diffeomorphism

$$T : \tilde{M} \to \tilde{M} \quad \text{such that} \quad \pi \circ T = \pi.$$

3.6 Uniqueness of simply connected space forms

The Hadamard-Cartan Theorem has an important consequence regarding *space forms*, that is, Riemannian manifolds whose sectional curvatures are constant:

Space Form Theorem. *Let k be a given real number. If $(M, \langle \cdot, \cdot \rangle)$ and $(\tilde{M}, \widetilde{\langle \cdot, \cdot \rangle})$ are complete simply connected Riemannian manifolds of constant curvature k, then $(M, \langle \cdot, \cdot \rangle)$ and $(\tilde{M}, \widetilde{\langle \cdot, \cdot \rangle})$ are isometric.*

The proof divides into two cases, the case where $k \leq 0$ and the case where $k > 0$. It is actually the first case that the Hadamard-Cartan Theorem directly applies.

Case I. Suppose that $k \leq 0$. In this case, the idea for the proof is really simple. Let $p \in M$ and $\tilde{p} \in \tilde{M}$. Then

$$\exp_p : T_pM \to M \quad \text{and} \quad \exp_{\tilde{p}} : T_{\tilde{p}}\tilde{M} \to \tilde{M}$$

are both diffeomorphisms by the Hadamard-Cartan Theorem. Let $\tilde{F} : T_pM \to T_{\tilde{p}}\tilde{M}$ be a linear isometry and let

$$F = \exp_{\tilde{p}} \circ \tilde{F} \circ \exp_p^{-1}.$$

Clearly F is a diffeomorphism, and it suffices to show that F is an isometry from M onto $\tilde{M}$.

Suppose that $q \in M$, $v \in T_qM$. Let $\tilde{q}$ be the corresponding point in $\tilde{M}$ and let $\tilde{v}$ be the corresponding vector in $T_{\tilde{q}}\tilde{M}$. It suffices to show that $\|v\| = \|\tilde{v}\|$.

Since M is complete, there is a geodesic $\gamma : [0, 1] \to M_1$ such that $\gamma(0) = p$ and $\gamma(1) = q$. The geodesic γ is the image under $\exp_p$ of a line segment in T_pM. The commutativity of the diagram

$$\begin{array}{ccc} T_pM & \xrightarrow{\tilde{F}} & T_{\tilde{p}}\tilde{M} \\ \exp_p \downarrow & & \exp_{\tilde{p}} \downarrow \\ M & \xrightarrow{F} & \tilde{M} \end{array}$$

shows that F will take γ to a geodesic $\tilde{\gamma}$ from $\tilde{p}$ to $\tilde{q}$. Moreover, F takes Jacobi fields along γ to Jacobi fields along $\tilde{\gamma}$. Let V be the unique Jacobi field along γ which vanishes at p and is equal to v at q. Then $\tilde{V} = F_*V$ is the Jacobi field along $\tilde{\gamma}$ which vanishes at $\tilde{p}$ and is equal to $\tilde{v}$ at $\tilde{q}$. Since $\tilde{F}$ is an isometry,

$$\text{the length of } \nabla_{\gamma'}V(0) = \text{ the length of } \nabla_{\tilde{\gamma}'}\tilde{V}(0).$$

It follows from the explicit formula (3.4) for Jacobi fields that the lengths of V and $\tilde{V}$ are equal at corresponding points, and hence $\|v\| = \|\tilde{v}\|$. This completes the proof when $k \leq 0$.

Case II. Suppose now that $k > 0$ and let $a = 1/\sqrt{k}$. It suffices to show that if $(\tilde{M}, \widetilde{\langle\cdot,\cdot\rangle})$ is an n-dimensional complete simply connected Riemannian manifold of constant curvature k, then it is globally isometric to $(\mathbb{S}^n(a), \langle\cdot,\cdot\rangle)$, where $\langle\cdot,\cdot\rangle$ is the standard metric on $\mathbb{S}^n(a)$. This case is a little more involved than the previous one, because $\mathbb{S}^n(a)$ is not diffeomorphic to its tangent space.

Lemma 1. *Suppose that $p \in \mathbb{S}^n(a)$, $\tilde{p} \in \tilde{M}$ and $\tilde{F} : T_p\mathbb{S}^n(a) \to T_{\tilde{p}}\tilde{M}$ is a linear isometry. If q is the antipodal point to p in $\mathbb{S}^n(a)$, then there is a unique smooth map*

$$F : \mathbb{S}^n(a) - \{q\} \to \tilde{M} \quad \text{such that} \quad (F_*)_p = \tilde{F}.$$

The proof is similar to the construction given for Case I. Note first that $\exp_p$ maps

$$\{v \in T_p\mathbb{S}^n(a) : \sqrt{\langle v, v\rangle} < \pi\} \quad \text{diffeomorphically onto} \quad \mathbb{S}^n(a) - \{q\}.$$

Since we need $(F_*)_p = \tilde{F}$ and F must take geodesics to geodesics, we are forced to define $F : \mathbb{S}^n(a) - \{q\} \to \tilde{M}$ by

$$F = \exp_{\tilde{p}} \circ \tilde{F} \circ \exp_p^{-1},$$

just as in the previous case, establishing uniqueness The argument given in Case I shows that F is indeed an isometric mapping:

$$\widetilde{\langle F_*(v), F_*(w)\rangle} = \langle v, w\rangle, \quad \text{for } v, w \in T_q\mathbb{S}^n(a),$$

establishing existence.

Returning to the proof of the theorem, we choose a point $p \in M$ and apply Lemma 1 to obtain an isometric map $F : \mathbb{S}^n(a) - \{q\} \to \tilde{M}$, where q is the antipodal point to p. Let r be a point in $\mathbb{S}^n(a) - \{p, q\}$. Then $(F_*)_r : T_r\mathbb{S}^n(a) \to \tilde{M}$ is a linear isometry. We obtain Lemma 1 once again to obtain an isometric map $\hat{F} : \mathbb{S}^n(a) - \{s\} \to \tilde{M}$, where s is the antipodal point to r. By uniqueness, $F = \hat{F}$ on overlaps. Hence F extends to a map $\bar{F} : \mathbb{S}^n(a) \to \tilde{M}$. In particular, $\bar{F}$ takes the antipodal point q to p to a single point of $\tilde{M}$.

Clearly, $\bar{F}$ is an immersion and a local isometry. Since $\bar{F}$ maps M onto $\tilde{M}$ by commutativity of the diagram

$$\begin{array}{ccc} T_p\mathbb{S}^n(a) & \xrightarrow{\tilde{F}} & T_{\tilde{p}}\tilde{M} \\ \exp_p \big\downarrow & & \exp_{\tilde{p}} \big\downarrow \\ \mathbb{S}^n(a) & \xrightarrow{\bar{F}} & \tilde{M}, \end{array}$$

we see that $\tilde{M}$ is compact. Thus the rest of the proof will follow from:

Lemma 2. *Suppose that M and $\tilde{M}$ are compact smooth n-dimensional smooth manifolds and $\bar{F} : M \to \tilde{M}$ is an immersion. Then $\bar{F}$ is a smooth covering map.*

The proof is a straightforward exercise in the theory of covering spaces.

In the context of our theorem, $\tilde{M}$ is simply connected, so $\bar{F}$ is a diffeomorphism. Thus we obtain the required diffeomorphism from $(\mathbb{S}^n(a), \langle\cdot,\cdot\rangle)$ to $(\tilde{M}, \widetilde{\langle\cdot,\cdot\rangle})$.

3.7 Non simply connected space forms

There is a wide variety of different space forms which are not simply connected.

In the flat case, we can take a basis $(v_1, \ldots, v_n)$ for $\mathbb{R}^n$ and consider the free abelian subgroup $\mathbb{Z}^n$ of $\mathbb{R}^n$ which is generated by the elements of the basis; thus

$$\mathbb{Z}^n = \{m_1 v_1 + \cdots m_n v_n : m_1, \ldots m_n \in \mathbb{Z}\}.$$

As usual, we let $\mathbb{E}^n$ denote $\mathbb{R}^n$ with the flat Euclidean metric. Then the quotient group $T^n = \mathbb{E}^n/\mathbb{Z}^n$ inherits a flat Riemannian metric; the resulting Riemannian manifold is called a flat n-torus. Note that $\pi_1(T^n) \cong \mathbb{Z}^n$.

In the positive curvature case, we can take identify antipodal points in $\mathbb{S}^n(1)$ obtaining the n-dimensional real projective space $\mathbb{R}P^n$. The obvious projection $\pi : \mathbb{S}^n(1) \to \mathbb{R}P^n$ is a smooth covering. Since the antipodal map is an isometry, there is a unique Riemannian metric on $\mathbb{R}P^n$ which pulls back under π to the metric of constant curvature one on $\mathbb{S}^n(1)$. Thus $\mathbb{R}P^n$ has a metric of constant curvature one and $\pi_1(\mathbb{R}P^n) \cong \mathbb{Z}_2$.

There are many other Riemannian manifolds which have metrics of constant curvature one. To construct further examples of three-dimensional manifolds with constant positive curvature, we make use of Hamilton's quaternions.

A quatenion Q can be regarded as a 2×2 matrix with complex entries,

$$Q = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

The set $\mathbb{H}$ of quaternions can be regarded as a four-dimensional real vector space with basis

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_x = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad E_y = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad E_z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Then if

$$Q = \begin{pmatrix} t + iz & x + iy \\ -x + iy & t - iz \end{pmatrix}, \quad \text{we can write} \quad Q = tI + xE_x + yE_y + zE_z.$$

Note that

$$\det Q = (t + iz)(t - iz) - (-x + iy)(x + iy) = t^2 + x^2 + y^2 + z^2$$

can be taken to be the Euclidean length of the quaternion.

Matrix multiplication makes $\mathbb{H} - \{0\}$ into a noncommutative Lie group. If

$$Q = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad \text{then} \quad Q^{-1} = \frac{1}{|Q|^2} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}, \tag{3.6}$$

where $|Q|^2 = |a|^2 + |b|^2 = \det Q$. Note that the determinant map

$$\det : \mathbb{H} - \{0\} \longrightarrow \mathbb{R}^+ = (\text{positive real numbers})$$

is a group homomorphism when the group operation on $\mathbb{R}^+$ is ordinary multiplication. We can now identify $\mathbb{S}^3(1)$ with the group of unit-length quaternions

$$\{Q \in \mathbb{H} : \det Q = 1\} = \{tI + xE_x + yE_y + zE_z : t^2 + x^2 + y^2 + z^2 = 1\}.$$

Since $\mathbb{S}^3(1)$ is the kernel of the determinant map, it is a Lie subgroup of $\mathbb{H} - \{0\}$. It follows directly from (3.6) that

$$\mathbb{S}^3(1) \cong SU(2) = \{A \in GL(2, \mathbb{C} : A^{-1} = \bar{A}^T\},$$

the special unitary group.

If $A \in \mathbb{S}^3(1)$ and $Q \in \mathbb{H}$, then $\det(AQ) = \det(Q) = \det(QA)$ so the induced metric on $\mathbb{S}^3(1)$ is biinvariant.

Moreover, if $A \in \mathbb{S}^3(1)$, we can define a linear isometry

$$\pi(A) : \mathbb{H} \to \mathbb{H} \quad \text{by} \quad \pi(A)(Q) = AQA^{-1}.$$

Since $\pi(A)$ preserves the t-axis, it can in fact be regarded as an element of

$$SO(3) = \{B \in GL(3, \mathbb{R} : B^T B = I \text{ and } \det B = 1\}.$$

Thus we obtain a group homomorphism $\pi : \mathbb{S}^3(1) \to SO(3)$ and it is an easy exercise to check that the kernel of π is $\{\pm I\}$. It follows that $SO(3)$ is in fact diffeomorphic to $\mathbb{R}P^3$, and we can consider the group of unit-length quaternions as the universal cover of $SO(3)$.

The group $SO(3)$ has many interesting finite subgroups. For example, the group of symmetries of a polygon of n sides is a group of order $2n$ called the dihedral group and denoted by $\mathbf{D}_n$. It is generated by a rotation through an angle $2\pi/n$ in the plane and be a reflection, which can be regarded as a rotation in an ambient $\mathbb{E}^3$. Thus the dihedral group can be regarded as a subgroup of $SO(3)$.

One also has groups of rotations of the five platonic solids, the tetrahedron, the cube, the octahedron, the dodecahedron and the icosahedron. The group of rotations of the tetrahedron $\mathbf{T}$ is just the alternating group on four letters and has order 12. The group of rotations $\mathbf{O}$ of the octahedron is isomorphic to the group of rotations of the cube and has order 24. Finally, the group of rotations $\mathbf{I}$ of the icosahedron is isomorphic to the group of rotations of the dodecahdron and has order 60. It is proven in §2.6 of Wolf [37] that the only finite groups of $SO(3)$ are cyclic and those isomorphic to $\mathbf{D}_n$, $\mathbf{T}$, $\mathbf{O}$ and $\mathbf{I}$.

One can take the preimage of these groups under the projection $\pi : \mathbb{S}^3(1) \to SO(3)$ obtaining the binary dihedral groups $\mathbf{D}_n^*$, the binary tetrahedral group $\mathbf{T}^*$, the binary octahedral group $\mathbf{O}^*$ and the binary icosahedral group $\mathbf{I}^*$. Thus one gets many examples of finite subgroups $G \subseteq \mathbb{S}^3(1)$. For each of these, one has a universal cover

$$\pi : \mathbb{S}^3(1) \to \mathbb{S}^3(1)/G,$$

left translations by elements of G being the deck transformations. Since these left translations are isometries, the quotient space $\mathbb{S}^3(1)/G$ inherits a Riemannian metric of constant curvature one with fundamental group G.

We can produce yet more examples by constructing finite subgroups of $SO(4)$ which act on $\mathbb{S}^3(1)$ without fixed points. For constructing such examples, it is helpful to know that $SU(2) \times SU(2)$ is a double cover of $SO(4)$. Indeed, if $(A_+, A_-) \in SU(2) \times SU(2)$, we can define

$$\pi(A_+, A_-) : \mathbb{H} \to \mathbb{H} \quad \text{by} \quad \pi(A_+, A_-)(Q) = A_+ Q A_-^{-1}.$$

This provides a surjective Lie group homomorphism $\pi : SU(2) \times SU(2) \to SO(4)$ with kernel $\{(I, I), (-I, -I)\}$.

In §4.4 we will show that any compact oriented connected surface of genus $g \geq 2$ possesses a Riemannian metric of constant negative curvature. In higher dimensions, there is an immense variety of nonsimply connected manifolds of constant negative curvature; such manifolds are called *hypberbolic manifolds*, and they possess a rich theory (see [34]).

3.8 Second variation of action

Curvature also affects the topology of M indirectly, through its effect on the stability of geodesics. We recall from §1.3 that geodesics are critical points of the action function $J : \Omega(M; p, q) \to \mathbb{R}$, where

$$\Omega(M; p, q) = \{ \text{ smooth paths } \gamma : [0, 1] \to M : \gamma(0) = p, \gamma(1) = q\},$$

and the action J is defined by

$$J(\gamma) = \frac{1}{2} \int_0^1 \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)} dt.$$

Recall that a *variation* of γ is a map

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega(M; p, q)$$

such that $\bar{\alpha}(0) = \gamma$ and the map

$$\alpha : (-\epsilon, \epsilon) \times [a, b] \to M \quad \text{defined by} \quad \alpha(s, t) = \bar{\alpha}(s)(t),$$

is smooth. Our next goal is to calculate the second derivative of $J(\bar{\alpha}(s))$ at $s = 0$ when $\bar{\alpha}(0)$ is a geodesic, which gives a test for stability because at a local minimum the second derivative must be nonnegative. This second derivative is called the *second variation* of J at γ. We will see that the sectional curvature of M plays a crucial role in the formula for second variation.

The first step in deriving the second variation formula is to differentiate under the integral sign which yields

$$\begin{aligned}\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} &= \frac{d^2}{ds^2}\left[\frac{1}{2}\int_0^1 \left\langle \frac{\partial\alpha}{\partial t}(s,t), \frac{\partial\alpha}{\partial t}(s,t)\right\rangle dt\right]\Big|_{s=0} \\ &= \left[\frac{1}{2}\int_0^1 \frac{\partial^2}{\partial s^2}\left\langle \frac{\partial\alpha}{\partial t}(s,t), \frac{\partial\alpha}{\partial t}(s,t)\right\rangle dt\right]\Big|_{s=0} \\ = \int_0^1 &\left[\left\langle \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial t}\right), \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial t}\right)\right\rangle + \left\langle \nabla_{\frac{\partial}{\partial s}}\nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial t}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle\right] dt\Big|_{s=0} \\ = \int_0^1 &\left[\left\langle \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right), \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right)\right\rangle + \left\langle \nabla_{\frac{\partial}{\partial s}}\nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle\right] dt\Big|_{s=0}.\end{aligned}$$

Using the curvature, we can interchange the order of differentiation to obtain

$$\begin{aligned}\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} &= \int_0^1 \left[\left\langle \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right), \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right)\right\rangle\right. \\ +\left\langle \nabla_{\frac{\partial}{\partial t}}\nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle &- \left.\left\langle R\left(\frac{\partial\alpha}{\partial t}, \frac{\partial\alpha}{\partial s}\right)\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle\right] dt\Big|_{s=0}.\end{aligned} \tag{3.7}$$

Now comes an integration by parts, using the formula

$$\begin{aligned}&\frac{\partial}{\partial t}\left\langle \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle \\ &\qquad = \left\langle \nabla_{\frac{\partial}{\partial t}}\nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle + \left\langle \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial t}\right)\right\rangle.\end{aligned}$$

Note that

$$\int_0^1 \frac{\partial}{\partial t}\left\langle \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle = 0$$

because $\alpha(s,0)$ and $\alpha(s,1)$ are both constant. Hence (3.7) becomes

$$\begin{aligned}\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} &= \int_0^1 \left[\left\langle \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right), \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial s}\right)\right\rangle\right. \\ -\left\langle R\left(\frac{\partial\alpha}{\partial t}, \frac{\partial\alpha}{\partial s}\right)\left(\frac{\partial\alpha}{\partial s}\right), \left(\frac{\partial\alpha}{\partial t}\right)\right\rangle &- \left.\left\langle \nabla_{\frac{\partial}{\partial s}}\left(\frac{\partial\alpha}{\partial s}\right), \nabla_{\frac{\partial}{\partial t}}\left(\frac{\partial\alpha}{\partial t}\right)\right\rangle\right] dt\Big|_{s=0}.\end{aligned}$$

Finally, we evaluate at $s = 0$ and use the fact that $\bar{\alpha}(0) = \gamma$ is a geodesic to obtain

$$\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\Big|_{s=0} = \int_0^1 \left[\langle \nabla_{\gamma'}X, \nabla_{\gamma'}X\rangle - \langle R(X,\gamma')\gamma', X\rangle\right] dt,$$

where X is the variation field defined by $X(t) = (\partial\alpha/\partial s)(0,t)$.

For $\gamma \in Omega(M; p, q)$, we define the "tangent space" to the "infinite-dimensional manifold" $\Omega(M; p, q)$ at the point γ to be

$$T_\gamma\Omega(M; p, q) = \{ \text{ smooth vector fields } X \text{ along } \gamma \ : X(0) = 0 = X(1)\}.$$

Definition. If $\gamma \in \Omega(M; p, q)$ is a geodesic, the *index form* of J at γ is the symmetric bilinear form

$$I : T_\gamma\Omega(M; p, q) \times T_\gamma\Omega(M; p, q) \to \mathbb{R}$$

defined by

$$I(X, Y) = \int_0^1 [\langle \nabla_{\gamma'} X, \nabla_{\gamma'} Y\rangle - \langle R(X, \gamma')\gamma', Y\rangle]\, dt, \tag{3.8}$$

for $X, Y \in T_\gamma\Omega(M; p, q)$.

By the polarization identity, the index form at a geodesic γ is the unique real-valued symmetric bilinear form I on $T_\gamma\Omega(M; p, q)$ such that

$$I(X, X) = \left. \frac{d^2}{ds^2} (J(\bar{\alpha}(s)))\right|_{s=0},$$

whenever $\bar{\alpha} : (-\epsilon, \epsilon) \to \Omega(M; p, q)$ is a smooth variation of γ with variation field X.

We can integrate by parts in (3.8) to obtain

$$I(X, Y) = -\int_0^1 \langle \nabla_{\gamma'} \nabla_{\gamma'} X + R(X, \gamma')\gamma', Y\rangle dt = \int_0^1 \langle L(X), Y\rangle dt,$$

where L is the *Jacobi operator*, defined by

$$L(X) = -\nabla_{\gamma'} \nabla_{\gamma'} X - R(X, \gamma')\gamma'.$$

Thus $I(X, Y) = 0$ for all $Y \in T_\gamma\Omega(M; p, q)$ if and only if X is a Jacobi field in $T_\gamma\Omega(M; p, q)$.

Note that the second variation argument we have given shows that if γ is a minimizing geodesic from p to q, the index form I at γ must be positive semi-definite.

3.9 Myers' Theorem

Recall that the *Ricci curvature* of a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ is the bilinear form

$$\mathrm{Ric} : T_pM \times T_pM \to \mathbb{R} \quad \text{defined by} \quad \mathrm{Ric}(x, y) = (\text{Trace of } v \mapsto R(v, x)y).$$

Myers' Theorem (1941). *If $(M, \langle\cdot,\cdot\rangle)$ is a complete connected n-dimensional Riemannian manifold such that*

$$Ric(v,v) \geq \frac{n-1}{a^2}\langle v,v\rangle, \quad \text{for all } v \in TM, \tag{3.9}$$

where a is a nonzero real number, then M is compact and $d(p,q) \leq \pi a$, for all $p,q \in M$. Moreover, the fundamental group of M is finite.

Proof of Myers' Theorem: It suffices to show that $d(p,q) \leq \pi a$, for all $p,q \in M$, because closed bounded subsets of a complete Riemannian manifold are compact.

Suppose that p and q are points of M with $d(p,q) > \pi a$. Let $\gamma : [0,1] \to M$ be a minimal geodesic with $\gamma(0) = p$ and $\gamma(1) = q$. Let $(E_1, \ldots, E_n)$ be a parallel orthonormal frame along γ with $\gamma' = d(p,q)E_1$. Finally, let

$$X_i(t) = \sin(\pi t)E_i(t), \quad \text{for } t \in [0,1] \text{ and } 2 \leq i \leq n.$$

Then for each i, $2 \leq i \leq n$,

$$\nabla_{\gamma'}X_i = \pi\cos(\pi t)E_i, \qquad \nabla_{\gamma'}\nabla_{\gamma'}X_i = -\pi^2\sin(\pi t)E_i,$$

and hence

$$\langle \nabla_{\gamma'}\nabla_{\gamma'}X_i, X_i\rangle = -\pi^2\sin^2(\pi t).$$

On the other hand,

$$\langle R(X_i,\gamma')\gamma', X_i\rangle = \sin^2(\pi t)d(p,q)^2\langle R(E_i,E_1)E_1,E_i\rangle,$$

so

$$\langle \nabla_{\gamma'}\nabla_{\gamma'}X_i + R(X_i,\gamma')\gamma', X_i\rangle = \sin^2(\pi t)[d(p,q)^2\langle R(E_i,E_1)E_1,E_i\rangle - \pi^2].$$

Hence

$$\begin{aligned}\sum_{i=2}^n I(X_i,X_i) &= -\sum_{i=2}^n \int_0^1 \langle \nabla_{\gamma'}\nabla_{\gamma'}X_i + R(X_i,\gamma')\gamma', X_i\rangle dt \\ &= \int_0^1 \sin^2(\pi t)\left[(n-1)\pi^2 - \pi^2 d(p,q)^2\sum_{i=2}^n\langle R(E_i,E_1)E_1,E_i\rangle\right]dt \\ &= \int_0^1 \pi^2\sin^2(\pi t)\left[(n-1-d(p,q)^2\mathrm{Ric}(E_1,E_1)\right]dt.\end{aligned}$$

Since $\mathrm{Ric}(E_1,E_1) \geq (n-1)/a^2$, we conclude that

$$\sum_{i=2}^n I(X_i,X_i) < \int_0^1 \pi^2\sin^2(\pi t)\left[(n-1-(n-1)\frac{d(p,q)^2}{a^2}\right]dt,$$

and the expression in brackets is negative because $d(p,q) > a$. This contradicts the assumption that γ is a minimal geodesic, by the second variation argument given in the preceding section.

To show that the fundamental group of M is finite, we let $\tilde{M}$ be the universal cover of M, and give $\tilde{M}$ the Riemannian metric $\pi^*\langle\cdot,\cdot\rangle$, where $\pi : \tilde{M} \to M$ is the covering map. The Ricci curvature of $\tilde{M}$ satisfies the same inequality (3.9) as the Ricci curvature of M; moreover $\tilde{M}$ is complete. Thus by the above argument $\tilde{M}$ must also be compact. Hence if $p \in M$, $\pi^{-1}(p)$ is a finite set of points. But by the arguments presented in §3.5.3, the number of points in $\pi^{-1}(p)$ is the order of the fundamental group of M. Thus the fundamental group of M must be finite.

For example, one can apply Myers' Theorem to show that $S^1 \times S^2$ cannot admit a Riemannian metric of positive Ricci curvature, because $\pi_1(S^1 \times S^2, x_0) \cong \mathbb{Z}$, and is therefore not finite. A famous open question posed by Hopf asks whether $S^2 \times S^2$ admits a Riemannian metric with positive sectional curvatures.

Another application is to Lie groups with biinvariant Riemannian metrics. If G is a Lie group with Lie algebra $\mathfrak{g}$, then the *center* of the Lie algebra is

$$\mathfrak{z} = \{X \in \mathfrak{g} : [X, Y] = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

Recall that any compact Lie group possesses a biinvariant Riemannian metric. This fact has a partial converse:

Corollary. *Suppose that G is a Lie group which has a biinvariant Riemannian metric. If the Lie algebra of G has trivial center, then G is compact.*

Proof: We use the explicit formula for curvature of biinvariant Riemannian metrics presented in §1.12. If E_1 is a unit-length element of $\mathfrak{g}$, we can extend E_1 to an orthonormal basis $(E_1, \ldots, E_n)$ for $\mathfrak{g}$, and conclude that

$$\mathrm{Ric}(E_1, E_1) = \sum_{i=2}^{n} \langle R(E_1, E_i)E_i, E_1\rangle = \sum_{i=2}^{n} \frac{1}{4}\langle [E_1, E_i], E_1, E_i]\rangle > 0.$$

As E_1 ranges over the unit sphere in $\mathfrak{g}$, the continuous function $E_1 \mapsto \mathrm{Ric}(E_1, E_1)$ must assume its minimum value. Hence $\mathrm{Ric}(E_1, E_1)$ is bounded below, and it follows from Myers' Theorem that G is compact.

Remark. Although Myers' Theorem puts a major restriction on the topology of compact manifolds of positive Ricci curvature, it is known that any manifold of dimension at least three has a complete Riemannian metric with negative Ricci curvature and finite volume [21].

3.10 Synge's Theorem

Recall from §1.19, smooth closed geodesics can be regarded as critical points for the action function $J : \mathrm{Map}(S^1, M) \to \mathbb{R}$, where $\mathrm{Map}(S^1, M)$ is the space of smooth closed curves and

$$J(\gamma) = \frac{1}{2}\int_0^1 \langle \gamma'(t), \gamma'(t)\rangle_{\gamma(t)} dt.$$

Here S^1 is regarded as being obtained from $[0,1]$ by identifying endpoints of the interval. It is interesting to consider conditions under which such critical points are stable.

In this case a *variation* of a point $\gamma \in \mathrm{Map}(S^1, M)$ is a map

$$\bar{\alpha} : (-\epsilon, \epsilon) \to \mathrm{Map}(S^1, M)$$

such that $\bar{\alpha}(0) = \gamma$ and the map

$$\alpha : (-\epsilon, \epsilon) \times S^1 \to M \quad \text{defined by} \quad \alpha(s,t) = \bar{\alpha}(s)(t),$$

is smooth. We can calculate the second derivative of $J(\bar{\alpha}(s))$ at $s = 0$ when $\bar{\alpha}(0)$ is a smooth closed geodesic, just as we did in §3.8, and in fact the derivation is a little simpler because we do not have to worry about contributions from the boundary of $[0,1]$. Thus we obtain the analogous result

$$\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = \int_0^1 \left[\langle \nabla_{\gamma'} X, \nabla_{\gamma'} X\rangle - \langle R(X,\gamma')\gamma', X\rangle\right] dt, \tag{3.10}$$

where now the variation field X is an element of

$$T_\gamma \mathrm{Map}(S^1, M) = \{ \text{ smooth vector fields } X \text{ along } \gamma : S^1 \to M \ \},$$

and by polarization we have an index form

$$I : T_\gamma \mathrm{Map}(S^1, M) \times T_\gamma \mathrm{Map}(S^1, M) \to \mathbb{R}$$
$$\text{defined by} \quad I(X,Y) = \int_0^1 \left[\langle \nabla_{\gamma'} X, \nabla_{\gamma'} Y\rangle - \langle R(X,\gamma')\gamma', Y\rangle\right] dt,$$

which must be positive semi-definite if the smooth closed geodesic is stable. The fact that the sectional curvature appears in the second variation formula (3.10) implies that there is a relationship between sectional curvatures and the stability of geodesics. This fact can be exploited to yield relationships between curvature and topology, as the following theorem demonstrates.

Synge's Theorem (1936). *Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a compact n-dimensional Riemannian manifold with positive sectional curvatures. If M is even-dimensional and orientable then M is simply connected.*

To prove this theorem, we use the Closed Geodesic Theorem from §1.19. Indeed, the nonconstant geodesic γconstructed in the proof is stable, and hence if $\bar{\alpha}(s)$ is any variation of γ,

$$\frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = \int_0^1 \left[\langle \nabla_{\gamma'} X, \nabla_{\gamma'} X\rangle - \langle R(X,\gamma')\gamma', X\rangle\right] dt \geq 0.$$

To construct an explicit variation that decreases action, we $p = \gamma(0)$ and make use of the parallel transport around γ:

$$\tau_\gamma : T_p M \longrightarrow T_p M.$$

If M is orientable, we know that this is an orientation-preserving isometry of T_pM. If $e_1 = (1/d(p,q))\gamma'(0)$, we can extend to a positively oriented orthonormal frame $(e_1, \ldots, e_n)$ for T_pM. From the canonical form theorem for special orthogonal transformations, if M is even-dimensional, say of dimension $2m$, we can choose the orthonormal basis so that τ_γ is represented by the matrix

$$\begin{pmatrix} \cos\theta_1 & \sin\theta_1 & & & & & & \\ -\sin\theta_1 & \cos\theta_1 & & \cdot & & \cdot & & \cdot \\ & & \cos\theta_2 & \sin\theta_2 & & & & \\ \cdot & & -\sin\theta_2 & \cos\theta_2 & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot & & \cdot \\ & & & & & & \cos\theta_m & \sin\theta_m \\ \cdot & & & \cdot & & \cdot & -\sin\theta_m & \cos\theta_m \end{pmatrix}. \tag{3.11}$$

In the first block, we can take $\theta_1 = 0$ because $\tau_\gamma(e_1) = e_1$. But then it follows that $\tau_\gamma(e_2) = e_2$, so there is a unit-length vector e_2 perpendicular to e_1 which is preserved by parallel transport around γ. We let X be the vector field along γ obtained by parallel transport along γ. Then since M has positive sectional curvatures,

$$I(X,X) = \int_0^1 \left[-\langle R(X,\gamma')\gamma', X\rangle\right] dt < 0.$$

Thus if $\bar{\alpha} : (-\epsilon, \epsilon) \to \mathrm{Map}(S^1, M)$ is a deformation of γ with deformation field X,

$$\frac{d}{ds}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} = 0, \qquad \frac{d^2}{ds^2}\left(J(\bar{\alpha}(s))\right)\bigg|_{s=0} < 0.$$

This contradicts the stability of the minimal geodesic γ, finishing the proof of the theorem.

Remark 1. It follows from Synge's Theorem, that the only even-dimensional complete Riemannian manifolds of constant curvature one are the spheres $\mathbb{S}^{2m}(1)$ and the projective spaces $\mathbb{R}P^{2m}$, the latter being nonorientable.

Exercise XI. Show that an odd-dimensional compact manifold with positive sectional curvatures is automatically orientable by an argument similar to that provided for Synge's Theorem. You can follow the outline:

a. By modifying the proof of the Closed Geodesic Theorem, show that if M were not orientable, one could construct a smallest smooth closed geodesic γ among curves around which parallel transport is orientation-reversing.

b. Show that in this case, the orthogonal matrix representing the parallel transport must have determinant -1, and its standard form is like (3.11) except for an additional 1×1 block containing -1.

c. As before, since the tangent vector to γ gets transported to itself, there is an additional unit vector e_2 perpendicular to γ which is transported to itself, and this implies that there is a nonzero parallel vector field X perpendicular to γ.

Use the second variation formula to show that one can deform in the direction of X to obtain an orientation-reversing curve which is shorter, thereby obtaining a contradiction just as before.

Chapter 4

Cartan's method of moving frames

4.1 An easy method for calculating curvature

Suppose that $(M, \langle \cdot, \cdot \rangle)$ is an n-dimensional Riemannian manifold, U an open subset of M. By a *moving orthonormal frame* on U, we mean an ordered n-tuple $(E_1, \dots, E_n)$ of vector fields on U such that for each $p \in U$, $(E_1(p), \dots, E_n(p))$ is an orthonormal basis for T_pM, for each $p \in M$. Suppose that M is oriented. Then we say that a moving orthonormal frame $(E_1, \dots, E_n)$ is *positively oriented* if $(E_1(p), \dots, E_n(p))$ is a positively basis for T_pM, for each $p \in M$.

Given a moving orthonormal frame on U, we can construct a corresponding *moving orthonormal coframe* $(\theta_1, \dots, \theta_n)$ by requiring that each θ_i be the smooth one-form on U such that

$$\theta_i(E_j) = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{4.1}$$

We can then write the restriction of the Riemannian metric to U as

$$\langle \cdot, \cdot \rangle | U = \sum_{i=1}^{n} \theta_i \otimes \theta_i.$$

Indeed, if

$$v = \sum_{i=1}^{n} a^i E_i(p) \quad \text{and} \quad w = \sum_{i=1}^{n} b^i E_i(p),$$

then

$$\langle v, w \rangle = \sum_{i,j=1}^{n} a^i b^j \langle E_i(p), E_j(p) \rangle = \sum_{i=1}^{n} a^i b^i = \sum_{i=1}^{n} \theta_i \otimes \theta_i(v, w).$$

Conversely, if we can write the Riemannian metric $\langle \cdot, \cdot \rangle$ in the form

$$\langle \cdot, \cdot \rangle | U = \sum_{i=1}^{n} \theta_i \otimes \theta_i,$$

where $\theta^1, \ldots \theta^n$ are smooth one-forms on U, then $(\theta_1, \ldots, \theta_n)$ is a moving orthonormal coframe on U, and one can use (4.1) to define a moving orthonormal frame $(E_1, \ldots, E_n)$ on U.

Corresponding to a given orthonormal frame $(E_1, \ldots, E_n)$ we can define *connection one-forms* ω_{ij} for $1 \leq i, j \leq n$ by

$$\nabla_X E_j = \sum_{i=1}^{n} \omega_{ij}(X) E_i \quad \text{or} \quad \omega_{ij}(X) = \langle E_i, \nabla_X E_j \rangle, \tag{4.2}$$

as well as *curvature two-forms* Ω_{ij} by

$$R(X, Y) E_j = \sum_{i=1}^{n} \Omega_{ij}(X, Y) E_i \quad \text{or} \quad \Omega_{ij}(X, Y) = \langle E_i, R(X, Y) E_j \rangle. \tag{4.3}$$

Since $\langle E_i, E_j \rangle = \delta^i_j$ and the Levi-Civita connection preserves the metric,

$$0 = X \langle E_i, E_j \rangle = \langle \nabla_X E_i, E_j \rangle + \langle E_i, \nabla_X E_j \rangle = \omega_{ji}(X) + \omega_{ij}(X),$$

and hence the matrix $\omega = (\omega_{ij})$ of connection one-forms is skew-symmetric. It follows from the curvature symmetries that the matrix $\Omega = (\Omega_{ij})$ of curvature two-forms is also skew-symmetric.

Theorem. *If $(M, \langle \cdot, \cdot \rangle)$ is a Riemannian manifold and $(E_1, \ldots, E_n)$ is a moving orthonormal coframe defined on an open subset U of M, with dual coframe $(\theta_1, \ldots, \theta_n)$, then the connection and curvature forms satisfy the structure equations of Cartan:*

$$d\theta_i = -\sum_{j=1}^{n} \omega_{ij} \wedge \theta_j, \tag{4.4}$$

$$d\omega_{ij} = -\sum_{j=1}^{n} \omega_{ik} \wedge \omega_{kj} + \Omega_{ij}. \tag{4.5}$$

Moreover, the ω_{ij}'s are the unique collection of one-forms which satisfy (4.4) together with the skew-symmetry condition

$$\omega_{ij} + \omega_{ji} = 0. \tag{4.6}$$

The proof of the two structure equations is based upon the familiar formula for the exterior derivative of a one-form:

$$d\theta(X, Y) = X(\theta(Y)) - Y(\theta(X)) - \theta([X, Y]). \tag{4.7}$$

Indeed, to establish (4.4), we need to verify that

$$d\theta_i(E_k, E_l) = -\sum_{j=1}^{n} (\omega_{ij} \wedge \theta_j)(E_k, E_l).$$

and a straightforward calculation shows that

$$\begin{aligned}
&d\theta_i(E_k, E_l) + \sum_{j=1}^{n} (\omega_{ij} \wedge \theta_j)(E_k, E_l) \\
&\quad = E_k(\theta_i(E_l)) - E_l(\theta_i(E_k)) - \theta_i([E_k, E_l]) + \sum_{j=1}^{n} (\omega_{ij} \wedge \theta_j)(E_k, E_l) \\
&\quad = -\theta_i([E_k, E_l]) + \sum_{j=1}^{n} \omega_{ij}(E_k)\theta_j(E_l) - \sum_{j=1}^{n} \omega_{ij}(E_l)\theta_j(E_k) \\
&\quad = \omega_{il}(E_k) - \omega_{ik}(E_l) - \theta_i([E_k, E_l]) \\
&\quad = \langle E_i, \nabla_{E_k} E_l - \nabla_{E_l} E_k - [E_k, E_l] \rangle = 0,
\end{aligned}$$

where we have used the fact that the Levi-Civita connection is symmetric in the last line of the calculation. Similarly,

$$\begin{aligned}
&d\omega_{ij}(E_k, E_l) + \sum_{r=1}^{n} (\omega_{ir} \wedge \omega_{rj})(E_k, E_l) \\
&\quad = E_k(\omega_{ij}(E_l)) - E_l(\omega_{ij}(E_k)) - \omega_{ij}([E_k, E_l]) \\
&\qquad + \sum_{r=1}^{n} \omega_{ir}(E_k)\omega_{rj}(E_l) - \sum_{r=1}^{n} \omega_{ir}(E_l)\omega_{rj}(E_k) \\
&\quad = E_k(\omega_{ij}(E_l)) - E_l(\omega_{ij}(E_k)) - \omega_{ij}([E_k, E_l]) \\
&\qquad - \sum_{r=1}^{n} \omega_{ri}(E_k)\omega_{rj}(E_l) + \sum_{r=1}^{n} \omega_{ri}(E_l)\omega_{rj}(E_k).
\end{aligned}$$

But

$$\begin{aligned}
E_k(\omega_{ij}(E_l)) = E_k\langle E_i, \nabla_{E_l} E_j \rangle &= \langle \nabla_{E_k} E_i, \nabla_{E_l} E_j \rangle + \langle E_i, \nabla_{E_k} \nabla_{E_l} E_j \rangle \\
&= \sum_{r=1}^{n} \omega_{ri}(E_k)\omega_{rj}(E_l) + \langle E_i, \nabla_{E_k} \nabla_{E_l} E_j \rangle,
\end{aligned}$$

while

$$E_l(\omega_{ij}(E_k)) = \sum_{r=1}^{n} \omega_{ri}(E_l)\omega_{rj}(E_k) + \langle E_i, \nabla_{E_l} \nabla_{E_k} E_j \rangle.$$

Hence

$$d\omega_{ij}(E_k, E_l) + \sum_{r=1}^{n} (\omega_{ir} \wedge \omega_{rj})(E_k, E_l)$$
$$= \langle E_i, \nabla_{E_k} \nabla_{E_l} E_j \rangle - \langle E_i, \nabla_{E_l} \nabla_{E_k} E_j \rangle - \langle E_i, \nabla_{[E_k, E_l]} E_j \rangle$$
$$= \langle E_i, R(E_k, E_l) E_j \rangle,$$

and the second structure equation is established.

Finally, to prove the uniqueness of the ω_{ij}'s, we suppose that we have two matrices of one-forms $\omega = (\omega_{ij})$ and $\tilde{\omega} = (\tilde{\omega}_{ij})$ which satisfy the first structure equation (4.4) and the skew-symmetry condition (4.6). Then the one-forms $\phi_{ij} = \omega_{ij} - \tilde{\omega}_{ij}$ must satisfy

$$\sum_{j=1}^{n} \phi_{ij} \wedge \theta_j = 0, \qquad \phi_{ij} + \phi_{ji} = 0.$$

We can write

$$\phi_{ij} = \sum_{j,k=1}^{n} f_{ijk} \theta_j \wedge \theta_k,$$

where each f_{ijk} is a smooth real-valued function on U. Note that

$$\sum_{j=1}^{n} \phi_{ij} \wedge \theta_j = 0 \quad \Rightarrow \quad f_{ijk} = f_{ikj},$$

while

$$\phi_{ij} + \phi_{ji} = 0 \quad \Rightarrow \quad f_{ijk} = -f_{jik}.$$

Hence

$$f_{ijk} = -f_{jik} = -f_{jki} = f_{kji} = f_{kij} = -f_{ikj} = -f_{ijk}.$$

Thus the functions f_{ijk} must vanish, and uniqueness is established.

The Cartan structure equations often provide a relatively painless procedure for calculating curvature:

Example. Suppose that we let

$$\mathbb{H}^n = \{(x^1, \ldots, x^{n-1}, y) \in \mathbb{R}^n : y > 0\},$$

and give it the Riemannian metric

$$\langle \cdot, \cdot \rangle = \frac{1}{y^2} \left(dx^1 \otimes dx^1 + \cdots + dx^{n-1} \otimes dx^{n-1} + dy \otimes dy \right).$$

In this case, we can set

$$\theta_1 = \frac{1}{y} dx^1, \quad \ldots, \quad \theta_{n-1} = \frac{1}{y} dx^{n-1}, \quad \theta_n = \frac{1}{y} dy, \tag{4.8}$$

thereby obtaining an orthonormal coframe $(\theta^1, \dots, \theta^n)$ on $\mathbb{H}^n$.

Differentiating (4.8) yields

$$d\theta_1 = \frac{1}{y^2} dx^1 \wedge dy = \theta_1 \wedge \theta_n, \quad \cdots$$
$$d\theta_{n-1} = \frac{1}{y^2} dx^{n-1} \wedge dy = \theta_{n-1} \wedge \theta_n, \quad d\theta_n = 0.$$

In other words, if $1 \leq a \leq n-1$,

$$d\theta_a = \theta_a \wedge \theta_n, \quad \text{while} \quad d\theta_n = 0.$$

The previous theorem says that there is a unique collection of one-forms ω_{ij} which satisfy the first structure equation and the skew-symmetry condition. We can solve for these connection forms, obtaining

$$\omega_{ab} = 0, \quad \text{for } 1 \leq a, b \leq n-1, \qquad \omega_{an} = -\theta_a, \quad \text{for } 1 \leq a \leq n-1.$$

From the explicit form of the ω_{ij}'s, it is now quite easy to show that the curvature two-forms are given by

$$\Omega_{ij} = -\theta_i \wedge \theta_j, \quad \text{for } 1 \leq i, j \leq n.$$

It now follows from (4.3) that

$$\begin{aligned}\Omega_{ij}(X, Y) = \langle E_i, R(X, Y)E_j \rangle = -\theta^i \wedge \theta^j (X, Y) \\ = -[\langle E_i, X \rangle \langle E_j, Y \rangle - \langle E_j, X \rangle \langle E_i, Y \rangle],\end{aligned}$$

so that

$$\langle R(X, Y)W, Z \rangle = -[\langle X, Z \rangle \langle Y, W \rangle - \langle X, W \rangle \langle Y, Z \rangle].$$

In other words the Riemannian manifold $(\mathbb{H}^n, \langle \cdot, \cdot \rangle)$ has constant sectional curvatures.

4.2 The curvature of a surface

The preceding theory simplifies considerably when applied to two-dimensional Riemannian manifolds, and yields a particularly efficient method of calculating Gaussian curvature of surfaces (compare §1.10).

Indeed, if $(M, \langle \cdot, \cdot \rangle)$ is an oriented two-dimensional Riemannian manifold, U an open subset of M, then a moving orthonormal frame (E_1, E_2) is uniquely determined up to a rotation: If (E_1, E_2) and $(\tilde{E}_1, \tilde{E}_2)$ are two positively-oriented moving orthonormal frames on a contractible open subset $U \subseteq M$, then

$$\begin{pmatrix} E_1 & E_2 \end{pmatrix} = \begin{pmatrix} \tilde{E}_1 & \tilde{E}_2 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

for some smooth function $\alpha : U \to \mathbb{R}$. It is no surprise that the corresponding moving orthonormal coframes (θ_1, θ_2) and $(\tilde{\theta}_1, \tilde{\theta}_2)$ are related by a similar formula:

$$\begin{pmatrix} \theta_1 & \theta_2 \end{pmatrix} = \begin{pmatrix} \tilde{\theta}_1 & \tilde{\theta}_2 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

It follows that the volume form is invariant under change of positively oriented moving orthonormal frame:

$$\theta_1 \wedge \theta_2 = \tilde{\theta}_1 \wedge \tilde{\theta}_2.$$

We claim that the corresponding skew-symmetric matrix of connection forms

$$\omega = \begin{pmatrix} 0 & \omega_{12} \\ -\omega_{12} & 0 \end{pmatrix},$$

transforms by the rule

$$\omega_{12} = \tilde{\omega}_{12} - d\alpha.$$

To see this, recall that ω_{12} is defined by the formula

$$\nabla_X E_2 = \omega_{12}(X) E_1,$$

and hence

$$\nabla_X(-\sin\alpha \tilde{E}_1 + \cos\alpha \tilde{E}_2 = \omega_{12}(X)(\cos\alpha \tilde{E}_1 + \sin\alpha \tilde{E}_2),$$

which expands to yield

$$\begin{aligned} -\cos\alpha d\alpha(X)\tilde{E}_1 + \sin\alpha d\alpha(X)\tilde{E}_2 - \sin\alpha \nabla_X \tilde{E}_1 + \cos\alpha \nabla_X \tilde{E}_2 & \\ = \omega_{12}(X)(\cos\alpha \tilde{E}_1 + \sin\alpha \tilde{E}_2). & \end{aligned}$$

Taking the inner product with $\tilde{E}_1$ yields

$$-\cos\alpha d\alpha(X) + \cos\alpha \langle \tilde{E}_1, \nabla_X \tilde{E}_2 \rangle = \omega_{12}(X)\cos\alpha,$$

and dividing by $\cos\alpha$ yields the desired formula

$$-d\alpha + \tilde{\omega}_{12} = \omega_{12}.$$

The skew-symmetric matrix of curvature forms

$$\Omega = \begin{pmatrix} 0 & \Omega_{12} \\ -\Omega_{12} & 0 \end{pmatrix}$$

is now determined by the Cartan's second structure equation

$$\Omega_{12} = d\omega_{12} = d\tilde{\omega}_{12} = \tilde{\Omega}_{12}.$$

Note that the curvature form Ω_{12} is independent of the choice of positively oriented moving orthonormal frame. Indeed, it follows from (4.3) that

$$\Omega_{12}(E_1, E_2) = \langle E_1, R(E_1, E_2)E_2 \rangle = K,$$

where K is the Gaussian curvature of $(M, \langle \cdot, \cdot \rangle)$, and hence

$$\Omega_{12} = K\theta_1 \wedge \theta_2.$$

This formula makes it easy to calculate the curvature of a surface using differential forms.

Definition. Suppose that $(M, \langle \cdot, \cdot \rangle)$ is an oriented two-dimensional Riemannian manifold. A positively-oriented coordinate system $(U, (x, y))$ is said to be *isothermal* if on U

$$\langle \cdot, \cdot \rangle = e^{2\lambda}(dx \otimes dx + dy \otimes dy), \tag{4.9}$$

where $\lambda : U \to \mathbb{R}$ is a smooth function.

Here is a deep theorem whose proof lies beyond the scope of the course:

Theorem. *Any oriented two-dimensional Riemannian manifold* $(M, \langle \cdot, \cdot \rangle)$ *has an atlas consisting of isothermal coordinate systems.*

A proof (using regularity theory of elliptic operators) can be found on page 378 of [33]. Assuming the theorem, we can ask: What is the relationship between positively oriented isothermal coordinate systems?

Suppose that (x^1, x^2) and (u^1, u^2) are two positively oriented coordinate systems on U with

$$\langle \cdot, \cdot \rangle = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j = \sum_{i,j=1}^{n} \tilde{g}_{ij} du^i \otimes du^j,$$

where

$$g_{ij} = e^{2\lambda}\delta_{ij}, \qquad \tilde{g}_{ij} = e^{2\mu}\delta_{ij}.$$

Then since the g_{ij}'s transform as the components of a covariant tensor of rank two,

$$g_{ij} = \sum_{k,l=1}^{n} \tilde{g}_{kl} \frac{\partial u^k}{\partial x^i} \frac{\partial u^l}{\partial x^j}$$

or

$$\begin{aligned} \begin{pmatrix} e^{2\lambda} & 0 \\ 0 & e^{2\lambda} \end{pmatrix} &= \begin{pmatrix} \partial u^1/\partial x^1 & \partial u^2/\partial x^1 \\ \partial u^1/\partial x^2 & \partial u^2/\partial x^2 \end{pmatrix} \begin{pmatrix} e^{2\mu} & 0 \\ 0 & e^{2\mu} \end{pmatrix} \begin{pmatrix} \partial u^1/\partial x^1 & \partial u^1/\partial x^2 \\ \partial u^2/\partial x^1 & \partial u^2/\partial x^2 \end{pmatrix} \\ &= J^T \begin{pmatrix} e^{2\mu} & 0 \\ 0 & e^{2\mu} \end{pmatrix} J, \quad \text{where} \quad J = \begin{pmatrix} \partial u^1/\partial x^1 & \partial u^1/\partial x^2 \\ \partial u^2/\partial x^1 & \partial u^2/\partial x^2 \end{pmatrix}. \end{aligned}$$

Hence

$$J^T J = e^{2\lambda - 2\mu} I, \quad \text{or} \quad B^T B = I, \quad \text{where} \quad B = e^{\mu - \lambda} J.$$

Since both coordinates are positively oriented, $B \in SO(2)$, and hence if U is contractible, we can write

$$B = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

for some function $\alpha : U \to \mathbb{R}$. Thus we see that

$$J = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where} \quad a = d \quad \text{and} \quad b = -c.$$

This implies that

$$\frac{\partial u^1}{\partial x^1} = \frac{\partial u^2}{\partial x^2}, \qquad \frac{\partial u^1}{\partial x^2} = -\frac{\partial u^2}{\partial x^1}.$$

These are just the Cauchy-Riemann equations, which express the fact that the complex-valued function $w = u^1 + iu^2$ is a holomorphic function of $z = x^1 + ix^2$.

Thus isothermal coordinates make an oriented two-dimensional Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ into a one-dimensional complex manifold, in accordance with the following definition:

Definition. An *n-dimensional complex manifold* is a second-countable Hausdorff space M, together with a collection $\mathcal{A} = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$ of "charts," where each U_α is an open subset of M and each ϕ_α is a homeomorphism from U_α onto an open subset of $\mathbb{C}^n$, such that $\phi_\alpha \circ \phi_\beta^{-1}$ is holomorphic where defined, for all $\alpha, \beta \in A$. A one-dimensional complex manifold is also called a *Riemann surface*.

We say that $\mathcal{A}$ is the atlas of holomorphic charts. If $(M, \mathcal{A})$ and $(N, \mathcal{B})$ are two complex manifolds, we say that a map $F : M \to N$ is *holomorphic* if $\psi_\beta \circ F \circ \phi_\alpha^{-1}$ is holomorphic where defined, for all charts $(U_\alpha, \phi_\alpha) \in \mathcal{A}$ and $(V_\beta, \psi_\beta) \in \mathcal{B}$. Two complex manifolds M and N are *holomorphically equivalent* if there is a holomorphic map $F : M \to N$ which has a holomorphic inverse $G : N \to M$.

In particular, we can speak of holomorphically equivalent Riemann surfaces. Two Riemannian metrics $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ on an oriented surface M are said to be *conformally equivalent* if there is a smooth function $\lambda : M \to \mathbb{R}$ such that

$$\langle \cdot, \cdot \rangle_1 = e^{2\lambda} \langle \cdot, \cdot \rangle_2.$$

This defines an equivalence relation, and given the existence of isothermal coordinates, it is clear that conformal equivalence classes of Riemannian metrics are in one-to-one correspondence with Riemann surface structures on a given oriented surface M.

Exercise XI. a. Suppose that $(M, \langle \cdot, \cdot \rangle)$ is an oriented two-dimensional Riemannian manifold with isothermal coordinate system $(U, (x, y))$ so that the Riemannian is given by (4.9). Use the method of moving frames to show that the Gaussian curvature of M is given by the formula

$$K = -\frac{1}{e^{2\lambda}} \left(\frac{\partial^2 \lambda}{\partial x^2} + \frac{\partial^2 \lambda}{\partial y^2} \right).$$

Hint: To start with, let $\theta_1 = e^\lambda dx$ and $\theta_2 = e^\lambda dy$. Then calculate ω_{12} and Ω_{12}.

b. Consider the Poincaré disk, the open disk $D = \{(x,y) \in \mathbb{R}^2\}$ with the Poincaré metric

$$\langle \cdot, \cdot \rangle = \frac{4}{[1-(x^2+y^2)]^2}(dx \otimes dx + dy \otimes dy).$$

Show that the Gaussian curvature of $(D, \langle \cdot, \cdot \rangle)$ is given by $K = -1$.

c. Show that reflections through lines passing through the origin are isometries and hence that lines passing through the origin in D are geodesics for the Poincaré metric. Show that the boundary of D is infinitely far away along any of these lines, and hence the geodesics through the origin can be extended indefinitely. Conclude from the Hopf-Rinow theorem that $(D, \langle \cdot, \cdot \rangle)$ is a complete Riemannian manifold, hence isometric to the model of the hyperbolic plane we constructed in §1.8.

4.3 The Gauss-Bonnet formula for surfaces

We now sketch the proof of the Gauss-Bonnet formula for surfaces in a version that suggests how it might be extended to n-dimensional oriented Riemannian manifolds. (See [29] for a more leisurely treatment.)

We start with an oriented two-dimensional Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$ without boundary and a smooth vector field $X : M \to TM$ with finitely many zeros at the points $p_1, p_2, \ldots, p_k$ of M. Let $V = M - \{p_1, \ldots, p_k\}$ and define a unit length vector field $Y : V \to TM$ by $Y = X/\|X\|$.

The *covariant differential* $DY = \nabla.Y$ of Y is the endomorphism of TM defined by $v \mapsto \nabla_v Y$. We will find it convenient to regard $\nabla.Y$ as a one-form with values in TM:

$$DY = \nabla.Y \in \Omega^1(TM).$$

If (E_1, E_2) is a positively oriented orthonormal moving frame defined on an open subset $U \subseteq M$, we can write

$$Y = f_1 E_1 + f_2 E_2 \quad \text{on} \quad U \cap V,$$

and

$$\nabla.Y = (df_1 + \omega_{12} f_2)E_1 + (df_2 - \omega_{12} f_1)E_2.$$

We let J denote counterclockwise rotation through 90 degrees in the tangent bundle, so that

$$JE_1 = E_2, \quad JE_2 = -E_1,$$

and

$$JY = -f_2 E_1 + f_1 E_2 \quad \text{on} \quad U \cap V.$$

Then

$$\psi = \langle JY, DY \rangle = f_1(df_2 - \omega_{12} f_1) - f_2(df_1 + \omega_{12} f_2) = f_1 df_2 - f_2 df_1 - \omega_{12}$$

is a globally defined one-form on $V = M - \{p_1, \ldots, p_k\}$ which depends upon X, and since $d(f_1 df_2 - f_2 df_1) = 0$,

$$d\psi = -\Omega_{12} = -K\theta_1 \wedge \theta_2, \tag{4.10}$$

where $\theta_1 \wedge \theta_2$ is the area form on M. The idea behind the proof of the Gauss-Bonnet formula is to apply Stokes's Theorem to (4.10).

Let ϵ be a small positive number. For each zero p_i of X, we let $C_\epsilon(p_i) = \{q \in M : d(p_i, q) = \epsilon\}$, a circle which inherits an orientation by regarding it as the boundary of

$$D_\epsilon(p_i) = \{q \in M : d(p_i, q) \leq \epsilon\}.$$

Definition. The *rotation index* of X about p_i is

$$\omega(X, p_i) = \frac{1}{2\pi} \lim_{\epsilon \to 0} \int_{C_\epsilon(p_i)} \psi,$$

if this limit exists.

Lemma. *If $(M, \langle \cdot, \cdot \rangle)$ is a two-dimensional compact oriented Riemannian manifold and X is a vector field on M with finitely many isolated zeros, then at each zero p_i the rotation index $\omega(X, p_i)$ exists and depends only on X, not on the choice of Riemannian metric on M.*

To prove this, we make use of the notion of the degree $\deg(F)$ of a continuous map F from S^1 to itself, as described in §3.5.2. Recall that such a map $F : S^1 \to S^1$ determines a homomorphism of fundamental groups

$$F_\sharp : \pi_1(S^1) \to \pi_1(S^1),$$

and since $\pi_1(S^1) \cong \mathbb{Z}$, this group homomorphism must be multiplication by some integer $n \in \mathbb{Z}$. We set $\deg(F) = n$.

Note that for $\epsilon > 0$ sufficiently small, we can define a map

$$F_\epsilon : C_\epsilon(p_i) \to S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \quad \text{by} \quad F_\epsilon(q) = (f_1(q), f_2(q)).$$

Then

$$\begin{aligned}\deg(F_\epsilon) = \frac{1}{2\pi} &\int_{C_\epsilon(p_i)} F_\epsilon^*(x dy - y dx) \\ &= \frac{1}{2\pi} \int_{C_\epsilon(p_i)} f_1 df_2 - f_2 df_1 = \frac{1}{2\pi} \lim_{\epsilon \to 0} \int_{C_\epsilon(p_i)} \psi.\end{aligned}$$

Thus $\omega(X, p_i)$ does indeed exist and is an integer.

To see that this integer is independent of the choice of Riemannian metric, note that any two Riemannian metrics $\langle \cdot, \cdot \rangle_0$ and $\langle \cdot, \cdot \rangle_1$ can be connected by a smooth one-parameter family

$$t \mapsto \langle \cdot, \cdot \rangle_t = (1 - t)\langle \cdot, \cdot \rangle_0 + t\langle \cdot, \cdot \rangle_1.$$

We can let $\omega_t(X, p_i)$ be the degree of X at p_i with respect to $\langle \cdot, \cdot \rangle_t$. Then $\omega_t(X, p_i)$ is a continuously varying integer and must therefore be constant.

It follows from transversality theory (as presented for example in Hirsch [15]) that any compact oriented surface possesses a vector field which has finitely many nondegenerate zeros. If X is any such vector field, we can apply Stokes's Theorem to $W = M - \bigcup_{i=1}^k D_\epsilon(p_i)$:

$$\int_W K\theta_1 \wedge \theta_2 = \int_W -d\psi = \sum_{i=1}^k \int_{C_\epsilon(p_i)} \psi, \tag{4.11}$$

the extra minus sign coming from the fact that the orientation $C_\epsilon(p_i)$ inherits from W is opposite to the orientation it receives as boundary of $D_\epsilon(p_i)$. In the limit as $\epsilon \to 0$, we obtain

$$\int_M K\theta_1 \wedge \theta_2 = 2\pi \sum_{i=1}^k \omega(X, p_i). \tag{4.12}$$

Since the left-hand side of (4.12) does not depend on the vector field while the right-hand side does not depend on the metric, neither side can depend on either the vector field or the metric, so both sides must equal an integer-valued topological invariant of compact oriented smooth surfaces $\chi(M)$, which is called the *Euler characteristic* of M. Thus we obtain two theorems:

Poincaré Index Theorem. *Suppose that M be a two-dimensional compact oriented smooth manifold and that X is a vector field on M with finitely many isolated zeros at the points $p_1, p_2, \ldots, p_k$. Then*

$$\sum_{i=1}^k \omega(X, p_i) = \chi(M).$$

Gauss-Bonnet Theorem. *Let $(M, \langle \cdot, \cdot \rangle)$ be a two-dimensional compact oriented Riemannian manifold with Gaussian curvature K and area form $\theta_1 \wedge \theta_2$. Then*

$$\frac{1}{2\pi} \int_M K\theta_1 \wedge \theta_2 = \chi(M).$$

Recall that a compact connected oriented surface is diffeomorphic to a sphere with h handles M_h. We can imbed M_h into $\mathbb{E}^3$ in such a way that the standard height function has exactly one nondegenerate maximum and one nodegenerate minimum, and $2h$ nondegenerate saddle points. The gradient X of the height function is then a vector field with nondegenerate zeros at the critical points of the height function. The maximum and minimum are zeros with rotation index one while each saddle point is a zero with rotation index -1. Thus $\chi(M_h) = 2 - 2h$.

The previous theorems can be extended to manifolds with boundary. In this case we consider a vector field X which has finitely many zeros at the points p_1,

$p_2, \ldots, p_k$ in the interior of M, is perpendicular to ∂M along ∂M and points outward along ∂M. As before, we let $V = M - \{p_1, \ldots, p_k\}$, define $Y : V \to TM$ by $Y = X/\|X\|$ and set $\psi = \langle JY, DY \rangle$. Then just as before

$$d\psi = -\Omega_{12} = -K\theta_1 \wedge \theta_2.$$

But this time, when we apply Stokes's Theorem to $W = M - \bigcup_{i=1}^k D_\epsilon(p_i)$ we obtain

$$\int_W K\theta_1 \wedge \theta_2 = \int_W -d\psi = -\int_{\partial M} \psi + \sum_{i=1}^k \int_{C_\epsilon(p_i)} \psi.$$

Thus when we let $\epsilon \to 0$, we obtain

$$\int_M K\theta_1 \wedge \theta_2 + \int_\partial M\psi = 2\pi \sum_{i=1}^k \omega(X, p_i). \tag{4.13}$$

Along ∂M, one can show that

$$\langle JY, DY \rangle = \kappa_g ds,$$

where κ_g is known as the geodesic curvature. Note that $\kappa_g = 0$ when ∂M consists of geodesics. As before, the left-hand side of (4.13) does not depend on the vector field while the right-hand side does not depend on the metric, so neither side can depend on either the vector field or the metric. The two sides must therefore equal a topological invariant which we call the *Euler characteristic* of M once again, thereby obtaining two theorems:

Poincaré Index Theorem for Surfaces with Boundary. *Suppose that M be a two-dimensional compact oriented smooth manifold with boundary ∂M and that X is a vector field on M with finitely many isolated zeros at the points $p_1, p_2, \ldots, p_k$ in the interior of M which is perpendicular to ∂M and points out along ∂M. Then*

$$\sum_{i=1}^k \omega(X, p_i) = \chi(M).$$

Gauss-Bonnet Theorem for Surfaces with Boundary. *Let M be a compact oriented smooth surface in with boundary ∂M. Then*

$$\int_M K dA + \int_{\partial S} \kappa_g ds = 2\pi\chi(M),$$

where f is the number of faces, e is the number of edges and v is the number of vertices in $\mathcal{T}$.

The celebrated uniformization theorem for Riemann surfaces shows that any Riemann surface has a complete Riemannian metric in its conformal equivalence class that has constant Gaussian curvature. For compact oriented surfaces, see that

1. the sphere has a Riemannian metric of constant curvature $K = 1$,
2. the torus T^2 has a metric of constant curvature $K = 0$,
3. and we will show in the next section that a sphere with h handles, where $h \geq 2$, has a Riemannian metric with constant curvature $K = -1$.

Of course, one could not expect such simple results for Riemannian manifolds of dimension ≥ 3, but as a first step, one might try to construct analogs of the Gauss-Bonnet formula for Riemannian manifolds of higher dimensions. Such an analog was discovered by Allendoerfer, Weil and Chern and is now called the generalized Gauss-Bonnet Theorem. This formula expresses the Euler characteristic of a compact oriented n-dimensional manifold as an integral of a curvature polynomial. It turns out that there are also several other topological invariants that can be expressed as integrals of curvature polynomials. Our next goal is to present part of the resulting theory, called the theory of characteristic classes, as developed by Chern, Pontrjagin and others [26].

Exercise XII. Use the Gauss-Bonnet Theorem for surfaces with boundary to calculate the Euler characteristic of the disk $D^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

4.4 Application to hyperbolic geometry

The *hyperbolic plane* (also called the Poincaré upper half plane) is the open set $\mathbb{H}^2 = \{(x, y) \in R^2 : y > 0\}$ together with the "Riemannian metric"

$$g = ds^2 = \frac{1}{y^2}[dx \otimes dx + dy \otimes dy] = \left(\frac{dx}{y}\right)^2 + \left(\frac{dy}{y}\right)^2.$$

The geometry of the "Riemannian manifold"$(\mathbb{H}^2, g)$ has many striking similarlities to the geometry of the ordinary Euclidean plane. In fact, the geometry of this Riemannian manifold is exactly the non-Euclidean geometry, which had been studied by Bolyai and Lobachevsky towards the beginning of the nineteenth century. It would be nice indeed if this non-Euclidean geometry could be realized as the geometry on some surface in $\mathbb{R}^3$, but this is not the case because of a famous theorem of David Hilbert (1901): The hyperbolic plane $\mathbb{H}^2$ cannot be realized on a surface in $\mathbb{R}^3$. In fact, a part of the hyperbolic plane can be realized as the geometry of the pseudo-sphere, but according to Hilbert's Theorem, the entire hyperbolic plane cannot be realized as the geometry of a smooth surface in $\mathbb{R}^3$. Thus abstract Riemannian geometry is absolutely essential for putting non-Euclidean geometry into its proper context as an important special case of the differential geometry of surfaces.

To study the Riemannian geometry of the hyperbolic plane in more detail, we can utilize the Darboux-Cartan method of moving frames to calculate the curvature of this metric. We did this before and found that the Gaussian curvature is given by the formula $K \equiv -1$. In particular the Gaussian curvature of the hyperbolic plane is the same at every point, just like in the case of the sphere. (Of course, there is nothing special about the constant -1; any other negative constant could be achieved by rescaling the metric.)

Another quite useful fact is that angles measured via the hyperbolic metric are exactly the same as those measured via the standard Euclidean metric $dx^2 + dy^2$. Indeed, the form of the metric shows that the coordinates (x, y) are isothermal.

Amazingly, the geodesics in the hyperbolic plane are very simple. Indeed, the straight line $x = c$ is the fixed point set of the reflection

$$\phi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2c - x \\ y \end{pmatrix},$$

which is easily seen to be an isometry of the metric:

$$\frac{1}{y^2}[d(2c - x)^2 + dy^2] = \frac{1}{y^2}[dx^2 + dy^2].$$

Thus if γ is the geodesic with initial conditions $\gamma(0) = (c, 1)$, $\gamma'(0) = (0, 1)$ then $\phi \circ \gamma$ is also a geodesic with the same initial conditions. By uniqueness of geodesics satisfying given initial conditions, $\phi \circ \gamma = \gamma$ and γ must lie in the vertical line $x = c$. It therefore follows that each vertical line $x = c$ is a geodesic. We can therefore ask if we can find a function $\alpha : \mathbb{R} \to [0, \infty)$ such that the curve

$$\gamma(t) = (0, \alpha(t))$$

is a unit-speed geodesic.

To solve this problem, note that

$$\frac{\alpha'(t)^2}{\alpha(t)^2} = 1 \quad \Rightarrow \quad \frac{\alpha'(t)}{\alpha(t)} = \pm 1 \quad \Rightarrow \quad \alpha(t) = ce^t \quad \text{or} \quad \alpha(t) = ce^{-t}.$$

We conclude that the x-axis is infinitely far away in terms of the Poincaré metric.

Other geodesics can be found by rewriting the metric in polar coordinates

$$ds^2 = \frac{1}{r^2 \sin^2 \theta}[dr^2 + r^2 d\theta^2] = \frac{1}{\sin^2 \theta}\left[\left(\frac{dr}{r}\right)^2 + d\theta^2\right].$$

The map ϕ which sends $r \mapsto 1/r$ and leaves θ alone is also an isometry:

$$\frac{r^2}{\sin^2 \theta}[(d(1/r))^2 + (1/r)^2 d\theta^2] = \frac{1}{r^2 \sin^2 \theta}[dr^2 + r^2 d\theta^2].$$

From this representation, it is easily seen that the map which sends $r \mapsto R^2/r$ and leaves θ alone is an isometry which fixes the semicircle

$$x^2 + y^2 = R^2, \qquad y > 0,$$

so this semicircle is also a geodesic. Since translation to the right or to the left are isometries of the hyperbolic metric, we see that all circles centered on the

x-axis intersect the hyperbolic plane in geodesics. (Of course, we have not found their constant speed parametrizations.)

Thus semicircles perpendicular to the x-axis and vertical rays are geodesics. Since there is one of these passing through any point and in any direction, we have described all the geodesics on the hyperbolic plane.

There is another property which the hyperbolic plane shares with Euclidean space and the sphere but with no other surfaces. That is, the hyperbolic plane has a large group of isometries, namely enough isometries to rotate through an arbitrary angle about any point and translate any point to any other point.

To study isometries in general, it is useful to utilize complex notation $z = x + iy$. (This is beneficial because the coordinates are isothermal.) Then $\mathbb{H}^2$ is simply the set of complex numbers with positive imaginary part.

Theorem. *The map*

$$z \mapsto \phi(z) = \frac{az+b}{cz+d} \tag{4.14}$$

is an isometry whenever a, b, c and d are real numbers such that $ad - bc > 0$.

Proof: Note that the transformation (4.14) is unchanged if we make the replacements

$$a \mapsto \lambda a, \quad b \mapsto \lambda b, \quad c \mapsto \lambda c, \quad d \mapsto \lambda d,$$

where $\lambda > 0$. So we can assume without loss of generality that $ad - bc = 1$. We can then factor the map

$$\phi(z) = z' = \frac{az+b}{cz+d}$$

into a composition of four transformations

$$z_1 = z + \frac{d}{c}, \quad z_2 = c^2 z_1, \quad z_3 = -\frac{1}{z_2}, \quad z' = z_3 + \frac{a}{c}. \tag{4.15}$$

To see this, note that

$$z_2 = c(cz+d), \quad z_3 = -\frac{1}{c(cz+d)},$$
$$z' = \frac{a(cz+d)}{c(cz+d)} - \frac{1}{c(cz+d)} = \frac{acz + ad - 1}{c(cz+d)} = \frac{az+b}{cz+d}.$$

The first and fourth transformations from (4.15) are translations of the hyperbolic plane which are easily seen to be isometries. It is straightforward to check (using polar coordinates) that the other two are also isometries. Since the composition of isometries is an isometry, ϕ itself is an isometry.

We will call the transformations

$$\phi(z) = \frac{az+b}{cz+d}, \qquad ad - bc = 1,$$

the *linear fractional transformations* and denote the space of linear fractional transformations by $PSL(2, \mathbb{R})$. The reflection $R : \mathbb{H}^2 \to \mathbb{H}^2$ in the line $x =$

0 is also an isometry, but it cannot be written as a special linear fractional transformation. It can be proven that any isometry of $\mathbb{H}^2$ can be written as a special linear fractional transformation, or as the composition $R \circ \phi$, where ϕ is a linear fractional transformation.

Suppose that ϕ is a linear fractional transformation of $\mathbb{H}^2$. If we set $c = 0$, we see that

$$\phi(z) = \frac{az+b}{d} = \frac{a}{d}z + \frac{b}{d} = a^2 z + ab,$$

since $d = 1/a$. This is a radial expansion or contraction about the origin, followed by a translation. Thus we can move any point in $\mathbb{H}^2$ to any other point by means of an isometry. This gives a rigorous proof that the hyperbolic plane is homogeneous; that is, it has the same geometric properties at every point!

Moreover, we can do an arbitrary rotation about a point. Suppose, for example, that we want to fix the point $i = (0,1)$. We impose this condition on the linear fractional transformation

$$\phi(z) = w = \frac{az+b}{cz+d},$$

and do a short calculation

$$i = \frac{ai+b}{ci+d} \qquad \Rightarrow \qquad -c + id = ai + b \qquad \Rightarrow \qquad a = d,\ b = -c.$$

Since

$$ad - bc = 1 \Rightarrow a^2 + b^2 = 1,$$

we can reexpress the linear fractional transformation in terms of sines and cosines:

$$\phi(z) = w = \frac{\cos\theta z + \sin\theta}{-\sin\theta z + \cos\theta}.$$

This isometry fixes the point i but there are clearly not the identity for arbitrary values of θ. An orientation-preserving isometry which fixes a point must act as a rotation on the tangent space at that point. This can be proven in general, but there is also a direct way to see it in our very specific context.

Simply note that

$$dw = \frac{adz(cz+d) - cdz(az+b)}{(cz+d)^2} = \cdots = \frac{dz}{(cz+d)^2}.$$

If we evaluate at our chosen point, we see that if everything is evaluated at the point i,

$$dw = \frac{dz}{(ci+d)^2} = \frac{dz}{(-\sin\theta i + \cos\theta)^2} = \frac{dz}{(e^{-i\theta})^2} = e^{2i\theta}dz.$$

Writing this out in terms of real and imaginary parts by setting

$$dw = du + idv, \qquad dz = dx + idy,$$

we see that the linear fractional transformation rotates the tangent space through an angle of 2θ.

Since $\mathbb{H}^2$ has the same geometric properties at every point, we see that we can realize rotations about any point by means of isometries.

Thus we have a group of hyperbolic motions which just as rich as the group of Euclidean motions in the plane. We can therefore try to develop the geometry of hyperbolic space in exactly the same way as Euclidean geometry of the plane.

In Euclidean geometry, any line can be extended indefinitely. In hyperbolic geometry the geodesic from the point $i = (0, 1)$ straight down to the x-axis has infinite length as we have seen above. Since any geodesic ray can be taken to any other geodesic ray by means of an isometry, we see that *any* geodesic ray from a point in the hyperbolic plane to the x-axis must have infinite length. In other words, geodesics can be extended indefinitely in the hyperbolic plane.

In Euclidean geometry there is a unique straight line between any two points. Here is the hyperbolic analogue:

Proposition. *In hyperbolic geometry, there is a unique geodesic connecting any two points.*

Proof: Existence is easy. Just take a circle perpendicular to the x-axis (or vertical line) which connects the two points.

If two geodesics intersect in more than one point, they would form a geodesic biangle, which is shown to be impossible by the Gauss-Bonnet Theorem:

Exercise XIII. a. Construct a geodesic biangle in S^2 with its standard metric using two geodesics from the north to the south poles, and use the Gauss-Bonnet Theorem to show that area of the geodesic biangle is 2α, where α is the angle between the geodesics.

b. Use the Gauss-Bonnet theorem to show that there is no geodesic biangle in the hyperbolic plane.

Proposition. *In hyperbolic geometry, any isosceles triangle, with angles α, β and β once again, can be constructed, so long as $\alpha + 2\beta < \pi$.*

Proof: Starting from the point i, construct two downward pointing geodesics which approach the x-axis and are on opposite sides of the y-axis. We can arrange that each of these geodesics makes an angle of $\alpha/2$ with the y-axis. Move the same distance d along each geodesic until we reach the points p and q. Connect p and q by a geodesic, thereby forming a geodesic triangle.

The interior angles at p and q must be equal because the triangle is invariant under the reflection R in the y-axis. When $d \to 0$, the Gauss-Bonnet formula shows that the sum of the interior angles of the geodesic approaches π. On the other hand, as the vertices of the triangle approach the x axis, the interior angles β at p and q approach zero.

By the intermediate value theorem from analysis, β can assume any value such that

$$0 < \beta < \frac{1}{2}(\pi - 2\alpha),$$

and the Proposition is proven.

Once we have this Proposition, we can piece together eight congruent isosceles geodesic triangles with angles

$$\alpha = \frac{\pi}{4}, \qquad \beta = \frac{\pi}{8}$$

to form a geodesic octagon. All the sides have the same length so we can identify them in pairs. This prescription identifies all of the vertices on the geodesic octagon. Since the angles at the vertices add up to 2π, a neighborhood of the identified point can be made diffeomorphic to an open subset of $\mathbb{R}^2$.

We need to do some cut and paste geometry to see what kind of surface results. The answer is a sphere with two handles, the compact oriented surface of genus two. Actually, it is easiest to see this by working in reverse, and cutting the sphere with two handles along four circles which emanate from a given point. From this we can see that the sphere with two handles can be cut into an octagon.

This then leads to the following remarkable fact:

Theorem. *A sphere with two handles can be given an abstract Riemannian metric with Gaussian curvature $K \equiv -1$.*

Indeed, a similar construction enables give a metric of constant curvature one on a sphere with g handles, where g is any integer such that $g \geq 2$.

Exercise XIV. Show that any geodesic triangle in the Poincaré upper half plane must have area $\leq \pi$.

Exercise XV. We now return to consider the Riemannian manifold $(\mathbb{H}^n, \langle \cdot, \cdot \rangle)$ described at the end of §**??**. We claim that it is also *homogeneous*, that is, given any two point $p, q \in \mathbb{H}^n$, there is an isometry

$$\phi : (\mathbb{H}^n, \langle \cdot, \cdot \rangle) \longrightarrow (\mathbb{H}^n, \langle \cdot, \cdot \rangle)$$

such that $\phi(p) = q$. Indeed, we have shown this above in the case where $n = 2$.

a. Show that $(\mathbb{H}^n, \langle \cdot, \cdot \rangle)$ is also homogeneous.

b. Use this fact to show that $(\mathbb{H}^n, \langle \cdot, \cdot \rangle)$ is complete. (Hint: If the exponential map $\exp_p$ at p is defined on a ball of radius $\epsilon > 0$ then so is $\phi \circ \exp_p$, which can be taken to be the exponential map at q. Thus no matter how far any geodesic has been extended, it can be extended for a distance at least $\epsilon > 0$ for a fixed choice of ϵ, and this implies that geodesics can be extended indefinitely.

Remark. Thus it follows from the uniqueness of simply connected complete space forms, that $(\mathbb{H}^n, \langle \cdot, \cdot \rangle)$ is isometric to the model of hyperbolic space we constructed in §1.8.

4.5 Vector bundles

The way to put Cartan's method of moving frames into a modern context is to imbed it in the theory of connections on vector bundles. Roughly speaking, a vector bundle of rank m over a manifold M is just a smoothly varying family of m-dimensional vector spaces parametrized by M. We have already seen several examples of smooth vector bundles, but it is helpful to have a precise definition.

Definition. A smooth F*-vector bundle* of rank m (where $F = \mathbb{R}$ or $\mathbb{C}$) is a pair of manifolds E and M, together with a smooth map $\pi : E \to M$, with the following additional structures:

1. For $p \in M$, $E_p = \pi^{-1}(p)$ has the structure of a m-dimensional vector space over F.
2. There is an open cover $\{U_\alpha : \alpha \in A\}$ of M, together with smooth maps
$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F^m,$$
such that $\pi_1 \circ \psi_\alpha = \pi$, where π_1 is the projection from $U_\alpha \times F^m$ to U_α.
3. If $\psi_\alpha = (\pi, \eta_\alpha)$, then $\eta_\alpha | E_p$ is a vector space isomorphism from E_p to F^m.

If $F = \mathbb{R}$, we say that E is a *real vector bundle over* M, while if $F = \mathbb{C}$, it is a *complex vector bundle over* M. We call E the *total space*, M the *base space* and E_p the *fiber* over p. The open cover $\{U_\alpha : \alpha \in A\}$ is called the *trivializing cover* for the vector bundle and the maps ψ_α are called the *trivializations*.

Let $GL(m, F)$ denote the group of $m \times m$ nonsingular matrices with entries in F. If E is a vector bundle of rank m over M, we can define its *transition functions*

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m, F) \quad \text{by} \quad \psi_\alpha \circ \psi_\beta^{-1}(p, v) = (p, g_{\alpha\beta}(p)v),$$

for $p \in U_\alpha \cap U_\beta$. These transition functions satisfy the relations

$$g_{\alpha\alpha} = I. \quad g_{\alpha\beta} g_{\beta\alpha} = I, \quad g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} = I, \tag{4.16}$$

wherever the products are defined.

Conversely, given an open cover $\{U_\alpha : \alpha \in A\}$ of a smooth manifold M, and smooth functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m, F)$ which satisfy the relations (4.16), we can construct a smooth vector bundle of rank m over M as follows: First we let $\hat{E}$ be the collection of all triples

$$(\alpha, p, v) \in A \times M \times F^m \quad \text{such that} \quad p \in U_\alpha.$$

We then define an equivalence relation $\sim$ on $\hat{E}$ by

$$(\alpha, p, v) \sim (\beta, q, w) \quad \Leftrightarrow \quad p = q \in U_\alpha \cap U_\beta \quad \text{and} \quad v = g_{\alpha\beta}(p)w.$$

Denote the equivalence class of (α, p, v) by $[\alpha, p, v]$ and the set of all equivalence classes by E. Define a projection $\pi : E \to M$ by $\pi([\alpha, p, v]) = p$ and a bijection

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F^m \quad \text{by} \quad \psi_\alpha([\alpha, p, v]) = (p, v).$$

There is a unique topology and smooth manifold structure on E such that $\pi^{-1}(U_\alpha)$ is open for each $\alpha \in A$ and each ψ_α is a diffeomorphism from $\pi^{-1}(U_\alpha)$ to $U_\alpha \times F^m$.

One of the most basic examples is the tangent bundle TM of a smooth manifold M. If M has a smooth atlas

$$\{(U_\alpha, \phi_\alpha) : \alpha \in A\}, \quad \text{where} \quad \phi_\alpha = (x^1_\alpha, \ldots, x^n_\alpha),$$

we can define the local trivialization

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n \quad \text{by} \quad \psi_\alpha \left(\sum_{i=1}^n a^i \left. \frac{\partial}{\partial x^i_\alpha} \right|_p \right) = \left(p, \begin{pmatrix} a^1 \\ \cdot \\ a^n \end{pmatrix} \right).$$

One can check that in this case,

$$g_{\alpha\beta} = D(\phi_\alpha \circ \phi_\beta^{-1}).$$

We have encountered many other examples of vector bundles including the cotangent bundle (T^*M, π, M) with transition functions

$$g_{\alpha\beta} = \left[[D(\phi_\alpha \circ \phi_\beta^{-1})]^T \right]^{-1}$$

and the various exterior and tensor powers of the cotangent bundle $\otimes^k T^*M$ and $\Lambda^k T^*M$.

Definition. Suppose that E is a smooth vector bundle over M. A *smooth section* of E is a smooth map $\sigma : M \to E$ such that $\pi \circ \sigma = \mathrm{id}_M$. We let $\Gamma(E)$ denote the space of smooth sections of E. Note that $\Gamma(E)$ is a module over the ring $\mathcal{F}(M)$ of smooth real-valued functions on M.

Similarly, if U is an open subset of M a smooth section of E over U is a smooth map $\sigma : U \to E$ such that $\pi \circ \sigma = \mathrm{id}_U$.

If σ is a smooth section of E and U_α is an element of the trivializing cover $\{U_\alpha : \alpha \in A\}$ for E, then

$$\sigma_\alpha = \eta \circ \sigma : U_\alpha \to F^m$$

is callled a *local representative* of σ. Alternatively, we can write

$$\sigma(p) = [\alpha, p, \sigma_\alpha(p)], \quad \text{for } p \in U_\alpha.$$

Note that two local representatives σ_α and σ_β are related on the overlap $U_\alpha \cap U_\beta$ by

$$\sigma_\alpha = g_{\alpha\beta}\sigma_\beta. \tag{4.17}$$

For example, a section of the tangent bundle TM is just a vector field on M. If U_α is the domain of the local coordinate system $(x^1_\alpha, \ldots, x^n_\alpha)$ and

$$X = \sum_{i=1}^{n} f^i_\alpha \frac{\partial}{\partial x^i_\alpha} \quad \text{on} \quad U_\alpha, \quad \text{then} \quad X_\alpha = \begin{pmatrix} f^1_\alpha \\ \cdot \\ f^n_\alpha \end{pmatrix}$$

is the local representative of X on U_α.

Suppose that (E_1, π_1) and (E_2, π_2) are two smooth vector bundles over M. A *vector bundle morphism* from (E_1, π_1) and (E_2, π_2) is a smooth map $F : E_1 \to E_2$ such that F takes the fiber of E_1 over p to the fiber of E_2 over p and the restriction of F to each fiber is linear. A *vector bundle isomorphism* is an invertible vector bundle morphism.

Given two vector bundles E and F over M, we can form their direct sum $E \oplus F$ and their tensor product $E \otimes F$. We can also construct the dual bundle E^* of a vector bundle and the endomorphism bundle $\mathrm{End}(E)$ whose fiber at a point $p \in M$ is just the space of endomorphisms of E_p.

4.6 Connections on vector bundles

Suppose that E is a smooth vector bundle over M. Let $\mathcal{X}(M)$ denote the space of smooth vector fields on M, $\Gamma(E)$ the space of smooth sections of E. We give several closely related definitions of connections on E.

Definition 1. A *connection* on E is a map

$$\nabla : \mathcal{X}(M) \times \Gamma(E) \to \Gamma(E)$$

which satisfies the following axioms (where $\nabla_X \sigma = \nabla(X, \sigma)$):

$$\nabla_X(f\sigma + \tau) = (Xf)\sigma + f\nabla_X \sigma + \nabla_X \tau, \tag{4.18}$$

$$\nabla_{fX+Y}\sigma = f\nabla_X \sigma + \nabla_Y \sigma. \tag{4.19}$$

Here f is a real-valued function if E is a real vector bundle or a complex-valued function if E is a complex vector bundle.

If E is a vector bundle, we let

$$\Omega^k(E) = \Gamma(\Lambda^k T^* M \otimes E) = \{ \text{ smooth } k\text{-forms with values in } E \}.$$

Given a connection ∇ in E, we can then define the *covariant differential*

$$D : \Omega^0(E) \to \Omega^1(E) \quad \text{by} \quad (D\sigma)(X) = \nabla_X \sigma.$$

Then D satisfies the second definition of connection:

Definition 2. A *connection* on E is a map $D : \Omega^0(E) \to \Omega^1(E)$ which satisfies the following axiom:

$$D(f\sigma + \tau) = df \otimes \sigma + fD\sigma + D\tau, \tag{4.20}$$

whenever f is a smooth function on M and $\sigma, \tau \in \Omega^0(E)$.

Remark. If D_1 and D_2 are two connections in E, then the difference $D_1 - D_2$ satisfies the identity

$$(D_1 - D_2)(f\sigma + g\tau) = f(D_1 - D_2)\sigma + g(D_1 - D_2)\tau,$$

that is the difference $D_1 - D_2$ is linear over functions. Hence we can regard $D_1 - D_2$ as an element of $\Omega^1(\mathrm{End}(E))$. Conversely if D is a connection on E and ϕ is any element of the vector space $\Omega^1(\mathrm{End}(E))$, then $D + \phi$ is also a connection. Thus although the zero operator is not a connection, the space of connections forms an "affine space" $D_0 + \Omega^1(\mathrm{End}(E))$, where D_0 is any particular choice of base connection in E.

Each of the definitions has its advantages, and we will find it convenient to pass back and forth between the two definitions.

The simplest example of a connection occurs on the trivial bundle $E = M \times \mathbb{R}^m$. A section of this bundle can be regarded as an $\mathbb{R}^m$-valued map

$$\sigma = \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix}.$$

In this case, we can use the exterior derivative to define the trivial flat connection on E:

$$D \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix} = \begin{pmatrix} d\sigma^1 \\ \vdots \\ d\sigma^m \end{pmatrix}.$$

More generally, given an $m \times m$ matrix of one-forms on M,

$$\omega = \begin{pmatrix} \omega_1^1 & \cdots & \omega_m^1 \\ \vdots & \ddots & \vdots \\ \omega_1^m & \cdots & \omega_m^m \end{pmatrix},$$

we can define a connection on E by

$$D \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix} = \begin{pmatrix} d\sigma^1 \\ \vdots \\ d\sigma^m \end{pmatrix} + \begin{pmatrix} \omega_1^1 & \cdots & \omega_m^1 \\ \vdots & \ddots & \vdots \\ \omega_1^m & \cdots & \omega_m^m \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix}. \tag{4.21}$$

We can write this last equation in a more abbreviated form as

$$D\sigma = d\sigma + \omega\sigma,$$

matrix multiplication being understood in the last term. Indeed, the axiom (4.20) can be verified directly, using the familiar properties of exterior derivative.

Of course, we can construct a connection in the trivial complex vector bundle $E = M \times \mathbb{C}^m$ in exactly the same way, by choosing ω to be an $m \times m$ matrix of complex-valued one-forms.

It is not difficult to check that any connection on the trivial bundle is of the form (4.21). To see this we apply D to the constant sections

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad E_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

obtaining

$$\begin{pmatrix} DE_1 & \cdots & DE_m \end{pmatrix} = \begin{pmatrix} E_1 & \cdots & E_m \end{pmatrix} \begin{pmatrix} \omega_1^1 & \cdots & \omega_m^1 \\ \vdots & \ddots & \vdots \\ \omega_1^m & \cdots & \omega_m^m \end{pmatrix},$$

for some collection of one-forms ω_j^i. If we then write a section as a linear combination of basis elements,

$$\sigma = \begin{pmatrix} E_1 \dots E_m \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix},$$

it follows directly from (4.20) that

$$\begin{aligned} D\sigma &= \begin{pmatrix} E_1 & \dots & E_m \end{pmatrix} \begin{pmatrix} d\sigma^1 \\ \vdots \\ d\sigma^m \end{pmatrix} + \begin{pmatrix} DE_1 \dots DE_m \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix} \\ &= \begin{pmatrix} E_1 & \dots & E_m \end{pmatrix} \left[\begin{pmatrix} d\sigma^1 \\ \vdots \\ d\sigma^m \end{pmatrix} + \begin{pmatrix} \omega_1^1 & \cdots & \omega_m^1 \\ \vdots & \ddots & \vdots \\ \omega_1^m & \cdots & \omega_m^m \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^m \end{pmatrix} \right], \end{aligned}$$

which is just another way of rewriting (4.21).

Recall that any vector bundle is locally trivial. Suppose, for example, that E is an F vector bundle of rank m defined by the open cover $\{U_\alpha : \alpha \in A\}$ and the transition functions $\{g_{\alpha\beta}\}$. In the notation of the preceding section

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F^m$$

is a vector bundle isomorphism from the restriction of E to U_α onto the trivial bundle over U_α. Such a trivialization defines a corresponding an ordered m-tuple $(E_1^\alpha, \dots E_m^\alpha)$ of smooth sections of E over U_α by the prescription

$$\psi_\alpha(E_i^\alpha(p)) = (p, E_i),$$

where $(E_1, \dots, E_m)$ is the standard basis of F^m. If σ is a smooth section of E over U_α, then

$$\sigma = \sum_{i=1}^{m} \sigma_\alpha^i E_i^\alpha \quad \text{where} \quad \sigma_\alpha = \begin{pmatrix} \sigma_\alpha^1 \\ \cdot \\ \sigma_\alpha^m \end{pmatrix}$$

is the local representative of σ over U_α. If D is a connection on E, then

$$D\sigma = D\left(\sum_{i=1}^{m} \sigma_\alpha^i E_i^\alpha\right) = \sum_{i=1}^{m} \left[d\sigma_\alpha^i + \sum_{j=1}^{m} (\omega_\alpha)_j^i \sigma_\alpha^j \right] E_i^\alpha.$$

We can write this as

$$D\sigma = \sum_{i=1}^{m} (D\sigma_\alpha)^i E_i^\alpha,$$
$$\text{where} \quad \begin{pmatrix} (D\sigma_\alpha)^1 \\ \vdots \\ (D\sigma_\alpha)^m \end{pmatrix} = \begin{pmatrix} d\sigma_\alpha^1 \\ \vdots \\ \sigma_\alpha^m \end{pmatrix} + \begin{pmatrix} (\omega_\alpha)_1^1 & \cdots & (\omega_\alpha)_m^1 \\ \vdots & \ddots & \vdots \\ (\omega_\alpha)_1^m & \cdots & (\omega_\alpha)_m^m \end{pmatrix} \begin{pmatrix} d\sigma_\alpha^1 \\ \vdots \\ \sigma_\alpha^m \end{pmatrix},$$

an m-tuple of ordinary one-forms on U_α. Just as we did above, we can write

$$(D\sigma)_\alpha = d\sigma_\alpha + \omega_\alpha \sigma_\alpha,$$

where ω_α is an $m \times m$ matrix of ordinary one-forms on U_α. We call the matrix operator $d + \omega_\alpha$ the *local representative* of the connection D.

We can use (4.17) to see how the local representatives corresponding to two elements of the distinguished covering are related on overlaps. We note that since the connection is well-defined on overlaps, we must have

$$d\sigma_\alpha + \omega_\alpha \sigma_\alpha = g_{\alpha\beta}(d\sigma_\beta + \omega_\beta \sigma_\beta) \quad \text{on } U_\alpha \cap U_\beta.$$

Since $\sigma_\beta = g_{\alpha\beta}^{-1} \sigma_\alpha$,

$$\begin{aligned} d\sigma_\alpha + \omega_\alpha \sigma_\alpha = g_{\alpha\beta}[d(g_{\alpha\beta}^{-1}\sigma_\alpha) + \omega_\beta g_{\alpha\beta}^{-1} \sigma_\alpha] \\ = d\sigma_\alpha + [g_{\alpha\beta} d g_{\alpha\beta}^{-1} + g_{\alpha\beta}\omega_\beta g_{\alpha\beta}^{-1}]\sigma_\alpha, \end{aligned}$$

and we conclude that

$$\omega_\alpha = g_{\alpha\beta} d g_{\alpha\beta}^{-1} + g_{\alpha\beta}\omega_\beta g_{\alpha\beta}^{-1}. \tag{4.22}$$

This yields yet another definition of connection:

Definition 3. A *connection* on an F-vector bundle E of rank m defined by an trivializing open cover $\{U_\alpha : \alpha \in A\}$ and transition functions $\{g_{\alpha\beta}\}$ is a collection of differential operators

$$\{d + \omega_\alpha, \alpha \in A\},$$

operating on local representatives, where d denotes the usual exterior derivative acting on F^m-valued sections and ω_α is an $m \times m$ matrix of F-valued one-forms, the differential operators transforming in accordance with (4.22).

4.7 Metric connections

A *fiber metric* on a real vector bundle E is a smooth function which assigns to each $p \in M$ an inner product

$$\langle \cdot, \cdot \rangle_p : E_p \times E_p \to \mathbb{R}.$$

Thus a Riemannian metric on a smooth manifold M is a fiber metric in its tangent bundle TM. If σ and τ are sections of E we can define a function

$$\langle \sigma, \tau \rangle : M \to \mathbb{R} \quad \text{by} \quad \langle \sigma, \tau \rangle(p) = \langle \sigma(p), \tau(p) \rangle_p.$$

Similarly, a *Hermitian metric* on a complex vector bundle E is a smooth function which assigns to each $p \in M$ a map

$$\langle \cdot, \cdot \rangle_p : E_p \times E_p \to \mathbb{C}$$

such that $\langle v, w \rangle_p$ is complex linear in v and conjugate linear in w,

$$\langle v, w \rangle_p = \overline{\langle w, v \rangle_p},$$

and $\langle v, v \rangle_p \geq 0$ with equality holding if and only if $v = 0$.

Definition. A *metric connection* on a real vector bundle E with a fiber metric or a complex vector bundle E with a Hermitian metric is a connection $D : \Omega^0(E) \to \Omega^1(E)$ such that

$$d\langle \sigma, \tau \rangle = \langle D\sigma, \tau \rangle + \langle \sigma, D\tau \rangle, \quad \text{for } \sigma, \tau \in \Omega^0(E). \tag{4.23}$$

Here we are using Definition 2 of connection. Of course, we can also use Definition 1 and write this condition as

$$X\langle \sigma, \tau \rangle = \langle \nabla_X \sigma, \tau \rangle + \langle \sigma, \nabla_Y \tau \rangle, \quad \text{for } X \in \mathcal{X}(M) \text{ and } \sigma, \tau \in \Gamma(E).$$

Using a partition of unity, we can show that any real vector bundle with fiber metric or any complex vector bundle with Hermitian metric admits a metric connection.

If E is a real vector bundle of rank m with fiber metric $\langle \cdot, \cdot \rangle$, we can choose trivializations which map each fiber of E_p isometrically onto $\mathbb{R}^m$ with its standard dot product. In this case the transition functions $\{g_{\alpha\beta}\}$ take their values in the orthogonal group $O(m)$. If

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^m$$

is such a trivialization and $(E_1, \ldots, E_m)$ is the standard orthonormal basis for $\mathbb{R}^m$, then the m-tuple $(E_1^\alpha, \ldots E_m^\alpha)$ of smooth sections of E over U_α defined by

$$\psi_\alpha(E_i^\alpha(p)) = (p, E_i)$$

has the property that $(E_1^\alpha(p), \dots E_m^\alpha(p))$ is an orthonormal basis for E_p for each p. It then follows from (4.23) that

$$0 = d\langle E_i^\alpha, E_j^\alpha \rangle = \langle DE_i^\alpha, E_j^\alpha \rangle + \langle E_i^\alpha, DE_j^\alpha \rangle = (\omega_\alpha)_i^j + (\omega_\alpha)_i^j,$$

so that the matrix ω_α is skew-symmetric. Recall that the Lie algebra of $GL(m, \mathbb{R})$ is just

$$\mathfrak{gl}(m, \mathbb{R}) \cong \{ \ m \times m\text{-matrices with entries in } \mathbb{R} \ \},$$

with Lie bracket defined by $[A, B] = AB - BA$, while the Lie algebra of $O(m)$ is the Lie subalgeba

$$\mathfrak{o}(m) = \{A \in \mathfrak{gl}(m, \mathbb{R}) : A^T + A = 0\}.$$

Thus each local representative $d + \omega_\alpha$ has the property that the matrix-valued one-form ω_α is $\mathfrak{o}(m)$-valued, where $\mathfrak{o}(m)$ is the Lie algebra of the orthogonal group.

Remark 1. Of course, the m-tuple $(E_1^\alpha, \dots E_m^\alpha)$ of smooth sections of E over U_α is just an extension of Cartan's notion of moving orthonormal frame from the tangent bundle TM to an arbitrary vector bundle E.

Remark 2. It is customary to lower the index and write ω_{ij} instead of ω_j^i when the matrix-valued form ω is $\mathfrak{o}(m)$-valued.

Similarly, if E is a complex vector bundle with Hermitian metric $\langle \cdot, \cdot \rangle$, we can choose trivializations which map each fiber of E_p isometrically onto $\mathbb{C}^m$ with its standard Hermitian inner product. In this case the transition functions $\{g_{\alpha\beta}\}$ take their values in the unitary group $U(m)$. If

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}^m$$

is such a trivialization and $(E_1, \dots, E_m)$ is the standard orthonormal basis for $\mathbb{C}^m$, then the m-tuple $(E_1^\alpha, \dots E_m^\alpha)$ of smooth sections of E over U_α defined by

$$\psi_\alpha(E_i^\alpha(p)) = (p, E_i)$$

has the property that $(E_1^\alpha(p), \dots E_m^\alpha(p))$ is an orthonormal basis for E_p for each p, and (4.23) implies that

$$0 = d\langle E_i^\alpha, E_j^\alpha \rangle = \langle DE_i^\alpha, E_j^\alpha \rangle + \langle E_i^\alpha, DE_j^\alpha \rangle = (\omega_\alpha)_i^j + \overline{(\omega_\alpha)_i^j},$$

so that the matrix ω_α is skew-Hermitian. The Lie algebra of $GL(m, \mathbb{C})$ is just

$$\mathfrak{gl}(m, \mathbb{C}) \cong \{ \ m \times m\text{-matrices with entries in } \mathbb{C} \ \},$$

with Lie bracket defined by $[A, B] = AB - BA$, while the Lie algebra of $M(m)$ is

$$\mathfrak{u}(m) = \{A \in \mathfrak{gl}(m, \mathbb{C}) : \bar{A}^T + A = 0\}.$$

Thus each local representative $d + \omega_\alpha$ has the property that the matrix-valued one-form ω_α is $\mathfrak{u}(m)$-valued in this case.

More generally, we could consider any Lie subgroup G of $GL(m, F)$, where $F = \mathbb{R}$ or $\mathbb{C}$. We could then define a *G-vector bundle* over M to be a rank m vector bundle over M whose transition functions take their values in G. For example, if $G = SO(m) = \{A \in O(m) : \det A = 1\}$, then an $SO(m)$-vector bundle is a real vector bundle of rank m together with a fiber metric and an orientation.

Suppose that the Lie group G has Lie algebra $\mathfrak{g}$. We could then define a *G-connection* on the G-vector bundle E defined by an trivializing open cover $\{U_\alpha : \alpha \in A\}$ and transition functions $\{g_{\alpha\beta}\}$ is a collection of differential operators

$$\{d + \omega_\alpha, \alpha \in A\}$$

acting on local representatives σ_α of sections, where d denotes the usual exterior derivative and ω_α is an $\mathfrak{g}$-valued matrix of one-forms, the operators transforming in accordance with (4.22). To see that this definition is well-defined, we would need to check that the transformation (4.22) preserves $\mathfrak{g}$-valued matrices of one-forms.

4.8 Curvature of connections

If we use Definition 2 of connection, we can extend the operator $D : \Omega^0(E) \to \Omega^1(E)$ to an operator

$$D : \Omega^*(E) \to \Omega^*(E), \quad \text{where} \quad \Omega^*(E) = \sum_{k=0}^{n} \Omega^k(E)$$

by forcing the Leibniz rule to hold:

$$D(\omega\sigma) = (d\omega) \otimes \sigma + (-1)^k \omega \wedge D\sigma, \quad \text{for} \quad \omega \in \Omega^k(M),\ \sigma \in \Gamma(E). \tag{4.24}$$

Unlike the usual exterior derivative, the extended operator D does not in general satisfy the identity $D \circ D = 0$. However,

$$\begin{aligned} D \circ D(f\sigma + \tau) &= D(df \otimes \sigma + f D\sigma + D\tau) \\ &= d(df) \otimes \sigma - df \wedge D\sigma + df \wedge D\Sigma + f(D \circ D)\sigma + (D \circ D)\tau \\ &= f(D \circ D)\sigma + (D \circ D)\tau, \end{aligned} \tag{4.25}$$

so $D \circ D$ is linear over functions. This implies that $D \circ D$ is actually a tensor field; in other words, there is a two-form with values in $\mathrm{End}(E)$,

$$R \in \Omega^2(\mathrm{End}(E)) \quad \text{such that} \quad D \circ D = R.$$

This $\mathrm{End}(E)$-valued two-form is called the *curvature* of the connection.

Suppose that E is a smooth real vector bundle of rank m over M defined by the open covering $\{U_\alpha : \alpha \in A\}$ and the transition functions $\{g_{\alpha\beta}\}$. Any element

$\sigma \in \Gamma(E)$ possesses local representatives $\{\sigma_\alpha : \alpha \in A\}$, and in accordance with (4.25),

$$(D \circ D\sigma)_\alpha = \Omega_\alpha \sigma_\alpha,$$

where Ω_α is a matrix of two-forms. Since D is represented by the operator $d + \omega_\alpha$ on $\Omega^0(E)$ and the operator $d + \omega_\alpha$ satisfies the Leibniz rule, one readily checks that D must also be represented by $d + \omega_\alpha$ on every $\Omega^k(E)$. We thus conclude that

$$\begin{aligned}(D \circ D\sigma)_\alpha &= (d + \omega_\alpha)(d\sigma_\alpha + \omega_\alpha \sigma_\alpha) \\ &= d(d\sigma_\alpha) + (d\omega_\alpha)\sigma_\alpha - \omega_\alpha \wedge d\sigma_\alpha + \omega_\alpha \wedge d\sigma_\alpha + (\omega_\alpha \wedge \omega_\alpha)\sigma_\alpha \\ &= (d\omega_\alpha + \omega_\alpha \wedge \omega_\alpha)\sigma_\alpha,\end{aligned}$$

and hence we obtain a formula for the local representative of the curvature:

$$\Omega_\alpha = d\omega_\alpha + \omega_\alpha \wedge \omega_\alpha. \tag{4.26}$$

Since $D \circ D$ is independent of trivialization, the matrices of two-forms must satisfy

$$\Omega_\alpha \sigma_\alpha = g_{\alpha\beta} \Omega_\beta \sigma_\beta = g_{\alpha\beta} \Omega_\beta g_{\alpha\beta}^{-1} \sigma_\alpha \qquad \text{on } U_\alpha \cap U_\beta.$$

In other words, the local representatives transform in accordance with the expected rule for two-forms with values in $\mathrm{End}(E)$,

$$\Omega_\alpha = g_{\alpha\beta} \Omega_\beta g_{\alpha\beta}^{-1}. \tag{4.27}$$

Differentiation of (4.26) gives

$$\begin{aligned}d\Omega_\alpha &= d\omega_\alpha \wedge \omega_\alpha - \omega_\alpha \wedge d\omega_\alpha \\ &= (\Omega_\alpha - \omega_\alpha \wedge \omega_\alpha) \wedge \omega_\alpha - \omega_\alpha \wedge (\Omega_\alpha - \omega_\alpha \wedge \omega_\alpha),\end{aligned}$$

which simplifies to yield the so-called *Bianchi identity*,

$$d\Omega_\alpha = \Omega_\alpha \wedge \omega_\alpha - \omega_\alpha \wedge \Omega_\alpha = [\Omega_\alpha, \omega_\alpha]. \tag{4.28}$$

Note that if E is a real vector bundle with metric connection, then the curvature matrices Ω_α are $\mathfrak{o}(m)$-valued, while if E is a complex vector bundle with metric connection, the curvature matrices Ω_α are $\mathfrak{u}(m)$-valued. In the special case where $E = TM$ the tangent bundle of a Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$, and D is the Levi-Civita connection, then (4.26) reduces to the second structure equation of Cartan (4.5) we encountered in Cartan's method of moving frames.

Remark. In some ways, it might have been better to use Definition 1 of connection and define the *curvature* of a connection ∇ on E to be the operator

$$R_1(X, Y)\sigma = \nabla_X \nabla_Y \sigma - \nabla_Y \nabla_X \sigma - \nabla_{[X,Y]}\sigma, \tag{4.29}$$

for $X, Y \in \mathcal{X}(M)$ and $\sigma \in \Gamma(E)$. It would then be straightforward to verify that

$$R_1(fX, Y)\sigma = R_1(X, fY)\sigma = R_1(X, Y)f\sigma = fR_1(X, Y)\sigma\sigma,$$

so that R_1 is an element of $\Omega^2(\mathrm{End}(E))$.

Recall that by means of a local trivialization

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F^m,$$

where $F = \mathbb{R}$ or $\mathbb{C}$, we could define an ordered m-tuple $(E_1^\alpha, \ldots E_m^\alpha)$ of smooth sections of E over U_α by the prescription

$$\psi_\alpha(E_i^\alpha(p)) = (p, E_i),$$

where $(E_1, \ldots, E_m)$ is the standard basis of F^m. The local representative $d+\omega_\alpha$ is then defined so that the entries of the matrix ω_α of one-forms satisfy

$$\nabla_X E_j = \sum_{i=1}^m E_i(\omega_\alpha)_j^i(X),$$

while, just as in §4.1, we could define the matrix Ω_α of curvature two-forms by

$$R_1(X, Y)E_j = \sum_{i=1}^m E_i(\Omega_\alpha)_j^i(X, Y).$$

Just as in the proof of the Theorem from §4.1 we would then be able to use (4.7) to derive Cartan's second structure equation

$$\Omega_\alpha = d\omega_\alpha + \omega_\alpha \wedge \omega_\alpha.$$

Since this agrees with (4.26), R_1 has exactly the same local representatives as R and hence $R = R_1$.

The upshot is that we have two equally valid definitions of curvature for a connection in E; the curvature $R \in \Omega^2(\mathrm{End}(E))$ is either the operator R_1 defined by (4.29) or the operator $R = D^2$, where D is the covariant differential.

4.9 The pullback construction

If (E, π, M) is a smooth vector bundle over M and $F : N \to M$ is a smooth map, we can define the *pullback bundle* with total space

$$F^*E = \{(p, v) \in N \times E : F(p) = \pi(v)\},$$

the projection $\pi : F^*E \to N$ being the projection onto the first factor. If $\{U_\alpha : \alpha \in A\}$ is a trivializing open cover for E, then $\{F^{-1}(U_\alpha) : \alpha \in A\}$ is a trivializing open cover for F^*E, and the transition functions for F^*E are $\{g_{\alpha\beta} \circ F\}$. Note that if E is a G-vector bundle, where G is a Lie subgroup of $GL(m, F)$ then so is F^*E.

If $\sigma \in \Gamma(E)$ then we can define a pullback section

$$F^*\sigma \in \Gamma(F^*E) \quad \text{by} \quad (F^*\sigma)(p) = (p, \sigma(F(p)).$$

More generally, we can define $F^* : \Omega^k(E) \to \Omega^k(F^*E)$ by

$$F^*(\omega \otimes \sigma) = F^*(\omega) \otimes F^*(\sigma) \quad \text{for } \omega \in \Omega^k(M) \text{ and } \sigma \in \Gamma(E).$$

Proposition 1. *If D_E is a connection on the vector bundle E over M and $F : N \to M$ is a smooth map, then there is a unique connection D_{F^*E} in the vector bundle F^*E over N which makes the following diagram commute:*

$$\begin{array}{ccc} \Omega^0(E) & \xrightarrow{D_E} & \Omega^1(E) \\ {\scriptstyle F^*}\downarrow & & {\scriptstyle F^*}\downarrow \\ \Omega^0(F^*E) & \xrightarrow{D_{F^*E}} & \Omega^1(F^*E) \end{array} .$$

If D_E is a G-connection, so is D_{F^*E}.

To prove uniqueness, one simply notes that if $d + \omega_\alpha$ is the local representative for D_E corresponding to the element U_α of the trivializing cover for E, then $d + F^*\omega_\alpha$ must be the local representative for D_{F^*E} corresponding to $F^{-1}(U_\alpha)$. For existence, one simply defines D_{F^*E} by decreeing that the local representative for D_{F^*E} corresponding to $F^{-1}(U_\alpha)$ must be $d + F^*\omega_\alpha$.

We call the connection obtained on F^*E by Proposition 1 the *pullback connection*. Note that if Ω_α is the local representative of curvature corresponding to an element U_α of the trivializing cover for E, then $F^*\Omega_\alpha$ must be the local representative of curvature for D_{F^*E} corresponding to $F^{-1}(U_\alpha)$.

The pullback connection formalizes the constructions we used when calculating first and second variation of the action J in §1.3 and §3.8. It provides a more rigorous means of discussing the parallel transport we introduced in §3.2.

Indeed, if E is a vector bundle over M with connection D_E and $\gamma : [a, b] \to M$ is a smooth map we can consider the connection D_{γ^*E} on the pullback bundle γ^*E. We say that a section σ of γ^*E is *parallel* along γ if $(D_{\gamma^*E})\sigma = 0$. In terms of a local trivialization, we can write such a section as σ_α, where

$$(d + \omega_\alpha)\sigma_\alpha = 0 \quad \text{or} \quad \frac{d}{dt}\sigma_\alpha^i + \sum_{j=1}^{m} (\omega_\alpha)_j^i \left(\frac{d}{dt}\right) \sigma_\alpha^j = 0.$$

Thus a section σ of γ^*E is parallel along γ if and only if its components in any local trivialization satisfy a first-order linear system of ordinary differential equations. It follows from the theory of such systems that given an element $\sigma_0 \in (\gamma^*E)_a$, there is a unique section σ parallel along γ which satisfies the initial condition $\sigma(a) = \sigma_0$. Thus we can define an isomorphism

$$\tau : (\gamma^*E)_a \to (\gamma^*E)_b$$

by setting $\tau(\sigma_0) = \sigma(b)$, where σ is the unique parallel section such that $\sigma(a) = \sigma_0$. Since

$$(\gamma^* E)_a \cong E_{\gamma(a)} \quad \text{and} \quad (\gamma^* E)_b \cong E_{\gamma(b)},$$

this also defines an isomorphism

$$\tau : E_{\gamma(a)} \to E_{\gamma(b)},$$

which we call *parallel transport along* γ *in* E. Of course, this is completely analogous to the construction we carried out in §3.2.

We can sometimes use parallel transport to prove that two vector bundles are isomorphic:

Proposition 2. *If* $F_0, F_1 : N \to M$ *are smoothly homotopic maps and* E *is a vector bundle over* M*, then* $F_0^* E$ *and* $F_1^* E$ *are isomorphic vector bundles over* N.

Sketch of proof: Define smooth maps

$$J_0, J_1 : N \to N \times [0,1] \quad \text{by} \quad J_0(p) = (p, 0), \quad J_1(p) = (p, 1).$$

If F_0 is smoothly homotopic to F_1, there is a smooth map

$$H : N \times [0,1] \to M \quad \text{such that} \quad H \circ J_0 = F_0, \quad H \circ J_1 = F_1.$$

Since $F_0^* E = J_0^* H^* E$ and $F_1^* E = J_1^* H^* E$, it suffices to show that if E is a vector bundle over $N \times [0,1]$, then $J_0^* E$ is isomorphic to $J_1^* E$.

But $J_0^* E$ is just the restriction of E to $N \times \{0\}$, while $J_1^* E$ is the restriction of E to $N \times \{1\}$. Give E a connection and let

$$\tau_p : E_{(p,0)} \to E_{(p,1)} \quad \text{be parallel transport along } t \mapsto (p, t).$$

We can then define a vector bundle isomorphism $\tau : J_0^* E \to J_!^* E$ by

$$\tau(v) = \tau_p(v), \quad \text{for } v \in E_{(p,0)} = J_0^* E_p.$$

Corollary. *Every vector bundle over a contractible manifold is isomorphic to a trivial bundle.*

Proof: If M is a contractible manifold, the identity map on M is homotopic to the constant map, and hence any vector bundle over M is isomorphic to the pullback of a bundle over a point via the constant map. Of course, such a pullback must be trivial.

4.10 Classification of connections in complex line bundles

If L is complex line bundle (that is a complex vector bundle of rank one), then $\mathrm{End}(L)$ is the trivial line bundle, and hence the curvature of any connection is just a complex scalar; it turns out that the curvature is actually purely imaginary if the L has a Hermitian metric and the connection is metric. There is a very simple classification for complex line bundles over a given manifold M and for metric connections in a given line bundle. Thus the differential geometry of complex line bundles over a given smooth manifold M has a relatively simple and complete theory.

Note that since $U(1)$ is isomorphic to $SO(2)$, a rank one complex vector bundle with Hermitian metric is just a rank two oriented real vector bundle with fiber metric. Indeed, given an oriented rank two real vector bundle E with fiber metric we can make it into a complex vector bundle of rank one by defining multiplication by i to be rotation through 90 degrees in the direction determined by the orientation. If (E_1, E_2) is a positively oriented moving frame for E over U, we can define multiplication in the fiber by i by setting $E_2 = iE_1$, thereby obtaining a complex line bundle with Hermitian metric.

If D is a metric connection in E regraded as an oriented real vector bundle, then the corresponding skew-symmetric matrix of connection forms is

$$\begin{pmatrix} 0 & \omega_{12} \\ -\omega_{12} & 0 \end{pmatrix}, \quad \text{where} \quad DE_1 = -\omega_{12}E_2 = (-i\omega_{12})E_1.$$

We can also regard the single section E_1 as defining a unitary frame for the complex line bundle E, and in this case the corresponding skew-Hermitian matrix of connection forms is $(-i\omega_{12})$.

The skew-symmetric curvature matrix of the $SO(2)$-bundle is

$$\begin{pmatrix} 0 & \Omega_{12} \\ -\Omega_{12} & 0 \end{pmatrix}, \quad \text{where} \quad \Omega_{12} = d\omega_{12},$$

while the skew-Hermitian curvature matrix for the $U(1)$-bundle is $(-i\Omega_{12})$. Since $U(1)$ and $SO(2)$ are commutative, it follows from (4.27) that $\Omega_\alpha = \Omega_\beta$ on $U_\alpha \cap U_\beta$, so the locally defined curvature matrices yield a global curvature matrix defined on M which has a globally defined entry Ω_{12}. We let $\mathcal{F} = \Omega_{12}$. It follows from the Bianchi identity that

$$d\Omega_{12} = d\mathcal{F} = 0.$$

Proposition 1. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are the curvatures of two $U(1)$-connections D_1 and D_2 in the same $U(1)$-bundle, then*

$$\left[\frac{1}{2\pi}\mathcal{F}_1\right] = \left[\frac{1}{2\pi}\mathcal{F}_2\right] \in H^2_{dR}(M;\mathbb{R}).$$

The proof is an easy consequence of the fact that the difference of two connections is a tensor field. Since $D_1 - D_2 = \alpha \in \Omega^1(\mathrm{End}(E))$, we see immediately that $\mathcal{F}_1 - \mathcal{F}_2 = d\alpha$, so $(1/2\pi)\mathcal{F}_1$ and $(1/2\pi)\mathcal{F}_2$ lie in the same de Rham cohomology class.

Definition. If L is a complex line bundle with a Hermitian metric, then the *first Chern class of* L or the *Euler class of the underlying* $SO(2)$*-bundle* E is

$$c_1(L) = \left[\frac{1}{2\pi}\mathcal{F}\right] \in H^2(M;\mathbb{R}),$$

where $\mathcal{F}$ is the curvature of any $U(1)$-connection in the $U(1)$-bundle L.

Let us explain the reason for the factor $1/2\pi$. The argument we used in §4.3 for the Poincaré Index Theorem and the Gauss-Bonnet Theorem carries over word for word to an arbitrary $U(1)$-bundle over a compact oriented surface M. Indeed, if $\sigma : M \to L$ is a smooth section with finitely many zeros at the points $p_1, p_2, \ldots, p_k$ of M, we can set $V = M - \{p_1, \ldots, p_k\}$ and define a unit-length section E_1 on V by $E_1 = \sigma/|\sigma|$. We can then define the rotation index $\omega(\sigma, p_i)$ of σ at each zero $p_i \in M$ just as before. The argument presented in §4.3 then shows that

$$\sum_{i=1}^{k} \omega(X, p_i) = \frac{1}{2\pi}\int_M \mathcal{F}, \tag{4.30}$$

and in particular, the differential form $(1/2\pi)\mathcal{F}$ must integrate over the surface M to an integer. We sometimes write

$$\frac{1}{2\pi}\int_M \mathcal{F} = \left\langle \left[\frac{1}{2\pi}\mathcal{F}\right], [M] \right\rangle = \langle c_1(L), [M]\rangle,$$

and call it the pairing of the fundamental class of M with the first Chern class or the Euler class.

The fact that $c_1(L)$ integrates to an integer is a reflection of the fact that $c_1(L)$ is actually an *integral cohomology class*, that is, it is the image of an element, also denoted by $c_1(L)$, under the coefficient homomorphism

$$H^2(X;\mathbb{Z}) \to H^2(X;\mathbb{R}) \tag{4.31}$$

described in §2.8.2. It would be difficult to classify complex vector bundles of higher rank up to isomorphism, but there is a very simple answer for the case of complex line bundles:

Classification Theorem for $U(1)$-bundles. *Let $\mathcal{V}^1_{\mathbb{C}}(M)$ denote the group of isomorphism classes of complex line bundles over a smooth manifold M, the group operation being tensor product. Then the map*

$$\Gamma : \mathcal{V}^1_{\mathbb{C}}(M) \to H^2(M;\mathbb{Z}) \quad \textit{defined by} \quad \Gamma(L) = c_1(L) \in H^2(M;\mathbb{Z})$$

is a group isomorphism.

The proof is a little difficult, so we defer a sketch of the proof to the next section (and, in fact, the reader can skip the proof on a first reading).

Example. The two-sphere $S^2 = \mathbb{C} \cup \{\infty\}$ can be regarded as a Riemann surface with two charts. We let

$$U_0 = \mathbb{C} \quad \text{and} \quad U_\infty = (\mathbb{C} - \{0\}) \cup \{\infty\}$$

and define $z : U_0 \to \mathbb{C}$ to the the identity and $w : U_\infty \to \mathbb{C}$ by $w = 1/z$. Then $\{(U_0, z), (U_\infty, w)\}$ is an atlas on S^2 which makes it into a Riemann surface. We can construct a line bundle H^n over S^2 by specifying

$$g_{0,\infty} : U_0 \cap U_\infty \to GL(1, \mathbb{C}) \quad \text{by} \quad g_{0,\infty}(z) = \frac{1}{z^n}.$$

The complex line bundle H^n will then have two trivializations

$$\psi_0 : \pi^{-1}(U_0) \to U_0 \times \mathbb{C}, \qquad \psi_\infty : \pi^{-1}(U_\infty) \to U_\infty \times \mathbb{C}$$

such that

$$\psi_\infty \circ \psi_0^{-1}(z, c) = \left(z, \frac{1}{z^n} c \right).$$

Define a section $\sigma : S^2 \to H^n$ by

$$\psi_0 \circ \sigma(z) = (z, z^n), \qquad \psi_\infty \circ \sigma(w) = (w, 1).$$

Then σ has a single zero at $z = 0$ and one can easily verify that $\omega(\sigma, 0) = n$. Using this fact together with (4.30) we find that $\langle c_1(H^n), [M] \rangle = n$. Indeed, every complex line bundle over S^2 is isomorphic to H^n for some $n \in \mathbb{Z}$.

We next ask whether it is possible to classify the metric connections in a given line bundle. Let

$$\mathcal{A}(L) = \{ \text{ metric connections in } L \text{ } \}.$$

In a sense we already know how to classify the connections in $\mathcal{A}(L)$. Indeed, suppose that we are given a base connection D_0 on E which is metric for a given choice of fiber metric $\langle \cdot, \cdot \rangle$. If D is any metric connection on L, then $D - D_0$ is a globally defined $\mathfrak{u}(1)$-valued one-form $-i\alpha$, where $\alpha \in \Omega^1(M)$. So elements of $\mathcal{A}(L)$ are in one-to-one correspondence with elements of $\Omega^1(M)$. But the interesting classification is up to "gauge transformation:"

Definition. A *gauge transformation* of a line bundle L with Hermitiian inner product $\langle \cdot, \cdot \rangle$ is a bundle automorphism $g : L \to L$ which preserves the inner product $\langle \cdot \rangle$.

If $g : L \to L$ is a gauge transformation, $g(p) : L_p \to L_p$ is multiplication by some complex number of length one. Thus we can regard g as a map

$$g : M \to U(1) = S^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

Let $\mathcal{G}$ denote the space of gauge transformations, a group under complex multiplication in the range. If p_0 is some choice of base point in M, we let

$$\mathcal{G}_0 = \{g \in \mathcal{G} : g(p_0) = 1\},$$

the subgroup of *based gauge transformations*. Then we have a direct product of groups $\mathcal{G} = \mathcal{G}_0 \times S^1$, where S^1 is the group of constant gauge transformations.

If D is a metric connection on L and $g \in \mathcal{G}$, we can construct a new connection D^g on L by

$$D^g = g \circ D \circ g^{-1} = D + gdg^{-1} : \Omega^0(L) \to \Omega^1(L).$$

We say that D and $\hat{D}$ are *gauge equivalent* if they differ by a bundle automorphism, in other words, if $D = D^g$ for some $g \in \mathcal{G}$, and we let

$$\begin{aligned} \mathcal{B}(L) = \frac{\mathcal{A}(L)}{\mathcal{G}} &= \frac{\mathcal{A}(L)}{\mathcal{G}_0} \\ &= \{ \text{ gauge equivalence classes of metric connections on } L \text{ } \}. \end{aligned}$$

Our goal is to classify elements of $\mathcal{B}(L)$.

In the case where M is simply connected, the classification of elements of $\mathcal{B}(L)$ is relatively easy. In this case, we let

$$\mathcal{C}(L) = \left\{ \mathcal{F} \in \Omega^2(M) : d\mathcal{F} = 0, \; \left[\frac{1}{2\pi}\mathcal{F} \right] = c_1(L) \right\}.$$

Proposition 2. *If $H^1(M;\mathbb{R}) = 0$, the map*

$$\Gamma : \mathcal{B}(L) \to \mathcal{C}(L) \quad \textit{defined by} \quad \Gamma(D) = \frac{1}{2\pi}\mathcal{F},$$

where $-i\mathcal{F}$ is the curvature of D, is a bijection.

Proof: To see that Γ is onto, we let D_0 be a base connection with curvature $\mathcal{F}_0$. If $\mathcal{F} \in \mathcal{C}(L)$, then $d\mathcal{F} = 0$ and since $[\mathcal{F}] = [\mathcal{F}_0]$, there is a one-form α on M such that $\mathcal{F} - \mathcal{F}_0 = d\alpha$. Let $D = D_0 - i\alpha$. Then D is a a connection with curvature $-i\mathcal{F}$.

To show that Γ is one-to-one, we suppose that D and D_0 are two connections with the same curvature, so $D = D_0 - i\alpha$ where $d\alpha = 0$. Since $H^1(M;\mathbb{R}) = 0$, there is a smooth function $u : M \to \mathbb{R}$ such that $du = \alpha$ and we let $g = e^{iu}$. Then

$$(D_0)^g = D_0 + gdg^{-1} = D_0 - idu = D_0 - i\alpha = D,$$

so D and D_0 are gauge equivalent.

We remark that it is possible to establish a similar result when $H^1(M;\mathbb{R}) \neq 0$. For simplicity, let us suppose that $H_1(M;\mathbb{Z})$ is free abelian of rank b_1 and let $\gamma_1, \cdots, \gamma_{b_1}$ be oriented simple closed curves representing generators for $H_1(M;\mathbb{Z})$, all passing through a given point p_0. Given a metric connection in a complex line bundle L, let

$$\tau_j : L_{p_0} \to L_{p_0}$$

be the holonomy around γ_j which is an orientation-preserving isometry and hence a rotation through an angle θ_j. Then τ_j must be rotation through some angle,

$$\tau_j = e^{i\theta_j}.$$

We can then define

$$\Gamma : \mathcal{B}(L) \to \mathcal{C}(L) \times \overbrace{S^1 \times \cdots \times S^1}^{b_1} \quad \text{by} \quad \Gamma(D) = \left(\frac{1}{2\pi}\mathcal{F}, e^{i\theta_1}, \ldots, e^{i\theta_{b_1}} \right),$$

and one can prove that this map is a bijection.

Suppose now that $(M\langle \cdot, \cdot \rangle$ is a Riemannian or Lorentz manifold so that the Hodge star $\star$ can be defined on $\Omega^*(M)$. We can then define the *Yang-Mills function*

$$\mathcal{Y} : \mathcal{A}(L) \to \mathbb{R} \quad \text{by} \quad \mathcal{Y}(D) = \int_M \mathcal{F}_D \wedge \star\mathcal{F}_D,$$

where $\mathcal{F}_D$ is the curvature of the connection D. We say that a connection on L is a *Yang-Mills connection* if it is a critical point for the function $\mathcal{Y}$.

Exercise XVI. Show that a connection D is a Yang-Mills connection if and only if it satisfies the equations

$$d\mathcal{F}_D = 0, \qquad d(\star\mathcal{F}_D) = 0.$$

Hint: Use the fact that if $\mathcal{F}_0$ and $\mathcal{F}_1$ are the curvatures of two $U(1)$-connections D_0 and D_1 in the same $U(1)$-bundle L, then $\mathcal{F}_1 - \mathcal{F}_0 = d\alpha$ for some globally defined one form on M. Moreover, for $t \in \mathcal{R}$, if $D_t = D_0 + t\alpha$, then $\mathcal{F}_t - \mathcal{F}_0 = td\alpha$. Now exploit the fact that if D_0 is a Yang-Mills connection, then

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{Y}(D_t) = 0.$$

In the case where $(M\langle \cdot, \cdot \rangle$ is a Lorentz manifold, these are just Maxwell's equations from electricity and magnetism. In the case where $(M\langle \cdot, \cdot \rangle$ is a compact Riemannian manifold the Yang-Mills connections are just the connections with harmonic curvature forms.

Application. If the base manifold M has a Lorentz metric and is considered to be the space-time of general relativity, we can model the Faraday tensor for an electromagnetic field (described at the end of §2.9) as the curvature $\mathcal{F}$ of a connection in a complex line bundle with a Hermitian metric. Then one of Maxwell's equations $d\mathcal{F} = 0$ is automatically satisfied.

What advantages does this model have over just regarding the Faraday tensor as a closed two-form? Well, for one thing, integrality of the first Chern class implies a quantization condition for $\mathcal{F}$. Second, the connection also provides holonomy around closed curves, and this has been observed experimentally (the Bohm-Aharanov effect). Third, a formulation in terms of connections suggests fruitful generalizations to connections in more general groups, like $SU(2)$ or $SU(3)$ and these have been used to help explain the weak and strong interactions in physics.

Thus while the geometry of pseudo-Riemannian manifolds has important applications to general relativity, the geometry of connections in vector bundles has important applications to the physics underlying the electromagnetic, weak and strong interactions.

4.11 Classification of $U(1)$-bundles*

For completeness, we provide a brief sketch for the proof of the Classification Theorem for $U(1)$-bundles described in the previous section (see [11], page 141). Indeed, this Classification Theorem follows from the proof of the de Rham Isomorphism Theorem presented in §2.8.3.

We start with a good cover of M, an open cover $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ such that any nonempty intersection $U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq 0$ in the cover is contractible. Let

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to U(1)$$

be transition functions for a $U(1)$-bundle L over M. Since the sets $U_\alpha \cap U_\beta$ are contractible, we can define maps

$$h_{\alpha\beta} = \frac{1}{2\pi i} \log g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathbb{R} \quad \text{such that} \quad e^{2\pi i h_{\alpha\beta}} = g_{\alpha\beta}. \tag{4.32}$$

Then

$$z_{\alpha\beta\gamma} = h_{\alpha\beta} + h_{\beta\gamma} + h_{\gamma\alpha} \in \mathbb{Z} \tag{4.33}$$

and one immediately verifies that $(z_{\alpha\beta\gamma}) \in \check{\mathrm{C}}^2(\mathcal{U}, \mathbb{Z})$ is a Čech cocycle. Of course, this also defines a Čech cocycle $(z_{\alpha\beta\gamma}) \in \check{\mathrm{C}}^2(\mathcal{U}, \mathbb{R})$ whose real cohomology class is the image under the coefficient homomorphism (4.31) of an integer class.

We need to check that this cohomology class corresponds to the first Chern class under the de Rham isomorphism.

So we have to follow the zig-zag construction that was used to produce this isomorphism. We start with $z = (z_{\alpha\beta\gamma}) \in \check{\mathrm{C}}^2(\mathcal{U}, \mathbb{R})$ which injects to an element $z = (z_{\alpha\beta\gamma}) \in \check{\mathrm{C}}^2(\mathcal{U}, \Omega^0)$. By (4.33), this image is then a Čech coboundary, $z = \delta(h)$, where $h \in \check{\mathrm{C}}^1(\mathcal{U}, \Omega^0)$ satisfies (4.32), and

$$d(h)_{\alpha\beta} = \frac{i}{2\pi} g_{\alpha\beta} dg_{\alpha\beta}^{-1} \in \check{\mathrm{C}}^1(\mathcal{U}, \Omega^1).$$

Exactness of the rows in the double complex then shows that there is an element $(1/2\pi)A_\alpha \in \check{\mathrm{C}}^0(\mathcal{U}, \Omega^1)$ which goes to $\frac{i}{2\pi} g_{\alpha\beta} dg_{\alpha\beta}^{-1}$ under δ, and hence

$$-iA_\alpha = -iA_\beta + g_{\alpha\beta} dg_{\alpha\beta}^{-1}.$$

But the the one-forms $\omega_\alpha = -iA_\alpha$ define a $U(1)$-connection on L and the element $(1/2\pi)A_\alpha \in \check{\mathrm{C}}^0(\mathcal{U}, \Omega^1)$ goes to its curvature $(i/2\pi)\Omega_\alpha = (1/2\pi)\mathcal{F}_\alpha \in \check{\mathrm{C}}^0(\mathcal{U}, \Omega^2)$. Exactness of the rows in the double complex once again forces the Ω_α's to fit together into a globally defined two-form $(1/2\pi)\mathcal{F}$ which is exactly the curvature of the $U(1)$-connection. The de Rham cohomology class of $(1/2\pi)\mathcal{F}$ is exactly the first Chern class of L. Thus the cohomology class of $(z_{\alpha\beta\gamma}) \in \check{\mathrm{C}}^2(\mathcal{U}, \mathbb{R})$ goes to the first Chern class $[(1/2\pi)\mathcal{F}]$ under the de Rham isomorphism, finishing our sketch of the proof of the Classification Theorem.

4.12 The Pfaffian

Let E be an oriented real vector bundle of rank $n = 2m$ over M with a fiber metric $\langle \cdot, \cdot \rangle$. The key example to keep in mind is the case where $E = TM$, where M has a Riemannian metric $\langle \cdot, \cdot \rangle$.

We let $\Lambda^q E$ denote the q-fold exterior power of E and let

$$\Omega^{p,q}(E) = \Omega^p(\Lambda^q E) = \{ \ p\text{-forms with values in } \Lambda^q E \ \}.$$

The direct sum

$$\Omega^{*,*}(E) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \Omega^{p,q}(E)$$

forms a "bigraded" algebra in which the product is defined by

$$(\phi\alpha) \wedge (\psi\beta) = (-1)^{\deg \psi \deg \alpha} (\phi \wedge \psi)(\alpha \wedge \beta),$$

for $\phi, \psi \in \Omega^{*,0}$, $\alpha, \beta \in \Omega^{0,*}$.

Suppose now that E has not only a fiber metric but also a metric connection $D : \Omega^0(E) \to \Omega^1(E)$. The fiber metric and the connection can be extended to every $\Lambda^q E$. Moreover, the connection on $\Lambda^q E$ can be extended to a first-order differential operator

$$D : \Omega^{p,q}(E) \to \Omega^{p+1,q}(E)$$

by forcing the Leibniz rule to hold.

If $(\Omega_\alpha)_{ij}$ are the curvature forms of the connection in E with respect to a moving frame $(E_1^\alpha, \ldots, E_n^\alpha)$ on an open subset $U_\alpha \subseteq M$, then the element

$$\mathcal{R}_\alpha = -\frac{1}{4} \sum_{i,j=1}^{n} (\Omega_\alpha)_{ij} E_i^\alpha \wedge E_j^\alpha \in \Omega^{2,2}(E|U_\alpha).$$

Moreover, a direct calculation shows that on $U_\alpha \cap U_\beta$, $\mathcal{R}_\alpha = \mathcal{R}_\beta$, so the local representatives fit together to form a globally defined element $\mathcal{R} \in \Omega^{2,2}(E)$. For simplicity, we write this as

$$\mathcal{R} = -\frac{1}{4} \sum_{i,j=1}^{n} \Omega_{ij} E_i \wedge E_j. \tag{4.34}$$

It is then possible to compute the m-th power within the algebra:

$$\begin{aligned}
\mathcal{R}^m &= \frac{(-1)^m}{2^{2m}} \sum_{i_1,\ldots,i_{2m}} \Omega_{i_1 i_2} E_{i_1} \wedge E_{i_2} \cdots \Omega_{i_{2m-1} i_{2m}} E_{i_{2m-1}} \wedge E_{i_{2m}} \\
&= \frac{(-1)^m}{2^{2m}} \sum_{\sigma \in S_{2m}} \Omega_{\sigma(1)\sigma(2)} \wedge \cdots \Omega_{\sigma(2m-1)\sigma(2m)} E_{\sigma(1)} \wedge \cdots \wedge E_{\sigma(2m)} \\
&= \frac{(-1)^m}{2^{2m}} \sum_{\sigma \in S_{2m}} \operatorname{sgn}(\sigma) \Omega_{\sigma(1)\sigma(2)} \wedge \cdots \Omega_{\sigma(2m-1)\sigma(2m)} E_1 \wedge \cdots \wedge E_{2m},
\end{aligned}$$

where S_{2m} is the group of permutations on $2m$ letters and $\mathrm{sgn}(\sigma)$ is the sign of the permutation σ.

Definition. If $A = (a_{ij})$ is a skew-symmetric matrix with $2m$ rows and $2m$ columns, the Pfaffian of A is

$$\mathrm{Pf}(A) = \frac{1}{2^m m!} \sum_{\sigma \in S_{2m}} \mathrm{sgn}(\sigma) a_{\sigma(1),\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)}.$$

The above computation then shows that

$$\frac{1}{m!}\mathcal{R}^m = \frac{(-1)^m}{2^m} Pf(\Omega) E_1 \wedge \cdots \wedge E_{2m}.$$

Since $\mathcal{R}$ is independent of moving frame, the calculation shows that

$$B \in SO(2m) \quad \Rightarrow \quad \mathrm{Pf}(B^{-1}\Omega B) = \mathrm{Pf}(\Omega).$$

The same identity must hold when Ω is replaced by an arbitrary skew-symmetric matrix A, so

$$B \in SO(2m) \quad \Rightarrow \quad \mathrm{Pf}(B^{-1} A B) = \mathrm{Pf}(A), \tag{4.35}$$

for $A \in \mathfrak{so}(2m)$. This identity shows, as we will see later, that the Pfaffian is an example of an "invariant polynomial" for the Lie group $SO(2m)$.

The factor $1/(2^m m!)$ is included in the definition of the Pfaffian so that

$$\mathrm{Pf}\begin{pmatrix} 0 & a_1 & 0 & 0 & \cdots & 0 & 0 \\ -a_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & a_2 & \cdots & 0 & 0 \\ 0 & 0 & -a_2 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_m \\ 0 & 0 & 0 & 0 & \cdots & -a_m & 0 \end{pmatrix} = a_1 a_2 \cdots a_m.$$

One can verify that $\mathrm{Pf}(A)^2 = \det(A)$, so the Pfaffian serves as a square root of the determinant for skew-symmetric matrices.

4.13 The generalized Gauss-Bonnet Theorem

If M is a $(2m)$-dimensional Riemannian manifold and $\Omega = (\Omega_{ij})$ is the matrix of curvature two-forms corresponding to a moving frame, we can construct the Pfaffian $\mathrm{Pf}(\Omega)$, a differential form of degree $2m$ on M, and it follows from (4.35) that $\mathrm{Pf}(\Omega)$ does not depend on the choice of moving frame.

Generalized Gauss-Bonnet Theorem. *If M is a compact oriented $(2m)$-dimensional Riemannian manifold, then*

$$\frac{1}{(2\pi)^m} \int_M Pf(\Omega) = \chi(M),$$

where $\chi(M)$ is the Euler characteristic of M.

Here the Euler characteristic of a compact smooth manifold M is a topological invariant which can be calculated by means of a vector field on M with nondegenerate zeros.

Definition. A vector field $V : M \to TM$ is said to have a *nondegenerate zero* at $p \in M$ if the covariant differential $DV(p) : T_pM \to T_pM$ is an isomorphism. If p is a nondegenerate zero of the vector field $V : M \to TM$, then the *rotation index* of V at p is

$$\omega(V,p) = \begin{cases} 1 & \text{if } \det(DV(p)) > 0, \\ -1 & \text{if } \det(DV(p)) < 0. \end{cases}$$

Thus suppose that p is a zero of V and that $(x^1, \ldots, x^n)$ are normal coordinates centered at p, with $x^i(p) = 0$. Then we can write

$$V = \begin{pmatrix} \frac{\partial}{\partial x^1} & \cdots & \frac{\partial}{\partial x^n} \end{pmatrix} \left[\begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} + \text{higher order terms} \right], \tag{4.36}$$

while the linearization of the vector field V at the zero p is

$$DV(p) = \begin{pmatrix} \left.\frac{\partial}{\partial x^1}\right|_p & \cdots & \left.\frac{\partial}{\partial x^n}\right|_p \end{pmatrix} \begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} dx^1|_p \\ \vdots \\ dx^n|_p \end{pmatrix}.$$

Of course, $DV(p)$ corresponds to a linear system of differential equations which takes the form

$$\begin{array}{ccccccc} \frac{dx^1}{dt} & = & a_1^1 x^1 & + & \cdots & + & a_n^1 x^n, \\ \vdots & & \vdots & & & & \vdots \\ \frac{dx^n}{dt} & = & a_1^n x^1 & + & \cdots & + & a_n^n x^n, \end{array}$$

or

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}, \quad \text{where} \quad A = \begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix}.$$

This linear system is just the linearization of the system of differential equations at the zero p which corresponds to V itself. We see that V has a nondegenerate zero at p if $\det A \neq 0$, and

$$\omega(V,p) = \begin{cases} 1 & \text{if } \det(A) > 0, \\ -1 & \text{if } \det(A) < 0. \end{cases}$$

For example, in the case of a surface, the vector field

$$V = x^1 \frac{\partial}{\partial x^1} + x^2 \frac{\partial}{\partial x^2} = \begin{pmatrix} \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

has rotation index one at the origin, while

$$V = x^1 \frac{\partial}{\partial x^1} - x^2 \frac{\partial}{\partial x^2} = \begin{pmatrix} \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

has rotation index minus one, agreeing with the rotation index as described in §4.3. We can now state the

Poincaré Index Theorem. *Suppose that M be a n-dimensional compact oriented smooth manifold and that V is a vector field on M with finitely many nondegenerate zeros at the points $p_1, p_2, \ldots, p_k$. Then*

$$\sum_{i=1}^{k} \omega(V, p_i) = \chi(M).$$

In our proof of the generalized Gauss-Bonnet Theorem, we will assume the existence of a vector field with nondegenerate zeros, that is, a vector field $V : M \to TM$ which intersects the zero section transversally. (The gradient of a Morse function is a vector field with nondegenerate zeros, and such functions are dense in the space of all real-valued functions on M by Corollary 6.7 in [25].) As in §4.3 our strategy is to prove the generalized Gauss-Bonnet Theorem and the Poincaré Index Theorem at the same time as we define the Euler characteristic of M by establishing the formula

$$\frac{1}{(2\pi)^m} \int_M \mathrm{Pf}(\Omega) = \sum_{i=1}^{k} \omega(V, p_i). \tag{4.37}$$

The left-hand side does not depend on the vector field while the right-hand side does not depend on the Riemannian metric, so neither side can depend on the metric or vector field. Both sides must equal a topological invariant which we call the Euler characteristic of M and denote by $\chi(M)$.

In the proof, we can replace the tangent bundle TM by an oriented real vector bundle E with fiber metric and connection, so long as E has the same rank as the dimension of M. In this case, a zero p of a section V is nondegenerate if $DV(p) : T_pM \to E_p$ is an isomorphism, and the rotation index is plus or minus one if this isomorphism is orientation-preserving or reversing, respectively. For this to work, we must replace (4.36) by

$$V = \begin{pmatrix} E_1 & \cdots & E_n \end{pmatrix} \left[\begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} + \text{higher order terms} \right], \tag{4.38}$$

where $(E_1, \ldots, E_n)$ is a positively oriented moving orthonormal frame for E.

The key idea in the proof is to construct a smooth closed n-form τ on the total space E, where n is both the rank of E and the dimension of M, called the *Thom form*, which is rapidly decreasing in the fiber direction, and therefore after rescaling, concentrated on a small neighborhood of the zero section. If $V : M \to E$ is any section and V_0 is the zero section, then by the homotopy axiom for de Rham theory,

$$\int_M V_0^* \tau = \int_M V^* \tau.$$

After we replace V by $(1/\varepsilon)V$ and let $\varepsilon \to 0$, (4.37) is obtained in the limit.

4.14 Proof of the generalized Gauss-Bonnet Theorem

Our basic object of study is a *geometric vector bundle*: a triple $(E, \langle,\rangle, D)$, where E is an oriented real vector bundle over a smooth manifold M, $\langle,\rangle$ is a fiber metric on E and D is a metric connection on E. We will associate to any such geometric vector bundle a *Thom form*, a smooth closed n-form on the total space E, where n is the rank of E, an n-form which is rapidly decreasing in the fiber direction and integrates to one on each fiber.

The simplest case is that of the trivial bundle $M \times \mathbb{R}^n$, with standard fiber metric and trivial connection. In this case, we can take Euclidean coordinates $(t_1, t_2, \ldots, t_n)$ on $\mathbb{R}^n$, and define the *Thom form* on $M \times \mathbb{R}^n$ by

$$\tau = \left(\frac{1}{\sqrt{\pi}}\right)^n e^{-(t_1^2 + \cdots + t_n^2)} dt_1 \wedge \cdots \wedge dt_n.$$

This form is closed and invariant under the action of the orthogonal group on the fiber, and since

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}, \tag{4.39}$$

it has the important property that its integral over the fiber $\mathbb{R}^n$ is one. Of course, if $\varepsilon > 0$ is a small positive constant, the rescaled Thom form

$$\tau_\varepsilon = \left(\frac{1}{\varepsilon\sqrt{\pi}}\right)^n e^{-(t_1^2 + \cdots + t_n^2)/\varepsilon^2} dt_1 \wedge \cdots \wedge dt_n$$

also integrates to one over the fiber and has the additional property that it is concentrated near the zero section.

In the case of a general geometric vector bundle $(E, \langle,\rangle, D)$, the construction of the Thom form makes use of the metric connection D. We consider the algebra of differential forms with values in the exterior algebra of E as describe in §**??**. (To simplify notation, we often suppress writing the wedge.)

A smooth section V of E gives rise to an element $DV = \nabla.V \in \Omega^{1,1}$, the covariant differential of V. If

$$V = \sum_{i=1}^n v_i E_i,$$

where $(E_1, \ldots, E_n)$ is a moving orthonormal frame defined on an open subset $U \subset M$, the covariant differential can be expressed by the formula

$$DV = \sum_{i=1}^{n} \left(dv_i + \sum_{j=1}^{n} \omega_{ij} v_j E_i \right),$$

where the ω_{ij}'s are the connection one-forms defined by the moving frame. As we have seen, the curvature forms of the connection,

$$\Omega_{ij} = d\omega_{ij} + \omega_{ik} \wedge \omega_{kj}$$

fit together into an element

$$\mathcal{R} = -\frac{1}{4} \sum_{ij} \Omega_{ij} E_i \wedge E_j \in \Omega^{2,2},$$

which like DV, is independent of the choice of moving orthonormal frame. All the summands in the expression

$$\Phi(V, D) = -|V|^2 + DV + \mathcal{R} \in \Omega^{0,0} \oplus \Omega^{1,1} \oplus \Omega^{2,2}$$

have even total degree and must therefore commute with each other. In particular, all terms in the expanded power series

$$\exp(\Phi(V, D)) = I + [-|V|^2 + DV + \mathcal{R}] + (1/2)[-|V|^2 + (DV) + \mathcal{R}]^2 + \cdots$$

commute, and we can write

$$\exp(\Phi(V, D)) = e^{-|V|^2}[I + DV + (1/2)(DV)^2 + \cdots][I + \mathcal{R} + (1/2)\mathcal{R}^2 + \cdots].$$

The two infinite series within brackets have only finitely many terms because $\Omega^{p,q} = 0$ when $p > \dim(M)$ or $q > n = \operatorname{rank}(E)$.

The orientation of E determines a volume element $\star 1 \in \Lambda^n(E)$, which can be expressed in terms of a positively oriented moving orthonormal frame as

$$\star 1 = E_1 \wedge \cdots \wedge E_n.$$

Following Quillen [30], [22], we define the *supertrace*

$$\operatorname{Tr}_s : \Omega^p(\Lambda^* E) \to \Omega^p(M),$$

the space of ordinary p-forms on M, by projection on the $\star 1$-component; thus

$$\operatorname{Tr}_s \left(\sum_{i_1 < \cdots < i_k} \phi_{i_1 \cdots i_k} E_{i_1} \wedge \cdots \wedge E_{i_k} \right) = \phi_{1 \cdots n}.$$

For any choice of section V, we have an n-form on M

$$\tau(V, D) = \frac{(-1)^{[n/2]}}{\pi^{(n/2)}} \operatorname{Tr}_s[\exp(\Phi(V, D))], \tag{4.40}$$

where $[n/2]$ represents the largest integer $\leq n/2$, a form which we will soon see is closed.

It is important to observe that the constructions we have performed are natural. If $F : N \to M$ is a smooth map, F^*E is the pullback of the bundle E to N, F^*V is the pullback section and F^*D is the pullback connection, then

$$\tau(F^*V, F^*D) = F^*\tau(V, D).$$

In particular, we can carry out the construction not only for the bundle E itself, but also for the pullback bundle via the bundle projection $\pi : E \to M$:

$$\pi^*E = \{(e_1, e_2) \in E \times E : \pi(e_1) = \pi(e_2)\}$$

This pullback bundle possesses a "tautological" section

$$T : E \to \pi^*E, \qquad T(e) = (e, e),$$

such that if $V : M \to E$ is any smooth section of E, then $V^*T = V$.

Definition. The *Thom form* on E corresponding to the connection D is the differential form

$$\tau(D) = \frac{(-1)^{[n/2]}}{\pi^{n/2}} \mathrm{Tr}_s[\exp(\Phi(T, \pi^*D))]. \tag{4.41}$$

Note that $\pi^*\omega_{ij}$ restricts to zero on each fiber. Hence if we represent a general point in the fiber in terms of a moving orthonormal frame as $\sum t_i E_i$, the restriction of the tautological section to the fiber is just

$$T(t_1, \ldots, t_n) = t_1 E_1 + \cdots + t_n E_n,$$

while the restriction of the Thom form to the fiber is

$$\frac{(-1)^{[n/2]}}{\pi^{n/2}} e^{-(t_1^2 + \cdots + t_n^2)} \mathrm{Tr}_s \left(\frac{1}{(n)!} \left(\sum dt_i E_i \right)^n \right).$$

In the sum one has $n!$ terms, each of which can be put in the form

$$(-1)^{[n/2]}(dt_1 \wedge dt_2 \wedge \cdots \wedge dt_n)(E_1 \wedge E_2 \wedge \cdots \wedge E_n),$$

since there we must make $n(n-1)/2$ changes of sign when permuting dt_i's and E_j's and $(-1)^{n(n-1)/2} = (-1)^{[n/2]}$. Thus after we apply Tr_s, we obtain

$$\frac{1}{\pi^{n/2}} e^{-(t_1^2 + \cdots + t_n^2)} dt_1 \wedge \cdots \wedge dt_n,$$

and we see that the Thom form integrates to one on each fiber, is rapidly decreasing in the fiber direction, and specializes to the Thom form constructed before in the case of the trivial bundle with trivial connection. By a similar argument, one sees that

$$\tau_\varepsilon(D) = \frac{(-1)^{[n/2]}}{\pi^{n/2}} \mathrm{Tr}_s \left[\exp\left(\Phi\left(\frac{T}{\varepsilon}, \pi^* D \right) \right) \right]$$

specializes to the rescaled Thom form τ_ε in the case of the trivial bundle.

Lemma. *The Thom form on E is closed.*

Proof: It will suffice to show that the differential form $\tau(V, D)$ defined by (4.40) is closed, since the Thom form is just a special case (where V is replaced by the tautological section on $\pi^* E$ and D by the pullback connection).

If V is a section of E, let $\iota(V) : \Omega^{p,q} \to \Omega^{p,q-1}$ be the interior product, the skew-derivation such that $\iota(V)(W) = \langle V, W \rangle$ when W is a section of E. It is then immediate that

$$\iota(V)(-|V|^2) = 0 \quad \text{and} \quad D(-|V|^2) = d(-|V|^2) = 2\iota(V)DV. \tag{4.42}$$

We next calculate

$$\begin{aligned} 2\iota(V)\mathcal{R} = 2\iota(V) &\left(-\frac{1}{4} \sum_{i,j} \Omega_{ij} E_i \wedge E_j \right) \\ &= -\frac{1}{2} \sum_{i,j} \Omega_{ij} \langle V, E_i \rangle E_j + \frac{1}{2} \sum_{i,j} \Omega_{ij} \langle V, E_j \rangle E_i \\ &\qquad = \sum_{i,j} \Omega_{ij} v_j E_i = R(\cdot,\cdot) \left(\sum v_j E_j \right) = R(\cdot,\cdot)V, \end{aligned}$$

so that

$$D(DV) = 2\iota(V)\mathcal{R}. \tag{4.43}$$

Finally, we use the Bianchi identity,

$$d\omega_{ij} = \sum \Omega_{ik} \wedge \omega_{kj} - \sum \omega_{ik} \wedge \Omega_{kj},$$

to calculate

$$\begin{aligned} D(\mathcal{R}) = D &\left(-\frac{1}{4} \sum_{i,j} \Omega_{ij} E_i \wedge E_j \right) \\ &= -\frac{1}{4} \sum_{i,j} d\Omega_{ij} E_i \wedge E_j - \frac{1}{4} \sum_{i,j} \Omega_{ij} DE_i \wedge E_j + \frac{1}{4} \sum_{i,j} \Omega_{ij} E_i \wedge DE_j \\ &\qquad = \cdots = 0. \end{aligned}$$

Thus we conclude that

$$D(\mathcal{R}) = 0. \tag{4.44}$$

It follows from (4.42), (4.43) and (4.44) that

$$D\Phi(V,D) = 2\iota(V)\Phi(V,D),$$

and since D and $\iota(V)$ are both skew-derivations, that

$$D(\exp(\Phi(V,D))) = 2\iota(V)\exp(\Phi(V,D)).$$

Now one applies the identities

$$\mathrm{Tr}_s \circ \iota(V) = 0, \qquad \mathrm{Tr}_s \circ D = d \circ \mathrm{Tr}_s$$

to conclude that $\tau(V,D)$ is closed.

If $V : M \to E$ is a smooth section of E, the form $\tau(V,\nabla)$ is obtained from the Thom form by pulling back via V: $\tau(V,D) = V^*(\tau(D))$. Since any two sections of E are smoothly homotopic, the de Rham cohomology class of $\tau(V,D)$ is independent of V.

In particular, if $n = 2m$,

$$\tau(0,D) = \tau_\varepsilon(0,D) = \frac{(-1)^m}{m!\pi^m}\mathrm{Tr}_s(\mathcal{R}^m) = \cdots = \frac{1}{(2\pi)^m}\mathrm{Pf}(\Omega), \tag{4.45}$$

which is just the Gauss-Bonnet integrand when $E = TM$.

Suppose now that $\dim M = \operatorname{rank} E$ and that V is a section of E with nondegenerate zeros. As $\varepsilon \to 0$,

$$\tau\left(\frac{V}{\varepsilon}, D\right) = (V)^*(\tau_\varepsilon)$$

becomes concentrated near the zeros of V, and we claim that in fact,

$$\lim_{\varepsilon\to 0}\int_M \tau\left(\frac{V}{\varepsilon}, D\right) = \sum(\text{rotation indices of } V \text{ at its zeros}). \tag{4.46}$$

To prove this, we choose a positively oriented moving orthonormal frame $(E_1,\ldots,E_n)$ and normal coordinates $(x^1,\ldots,x^n)$ both defined on a neighborhood $B_\delta(p)$ about a zero p for V. We can then write

$$V = \sum a^i_j x^j E_i + (\text{higher order terms}),$$

and hence

$$DV = \sum a^i_j dx^j E_i + (\text{higher order terms}).$$

Thus

$$\begin{aligned}(DV)^n &= \sum a^{i_1}_{j_1} dx^{i_1} E_{j_1} \cdots \sum a^{i_n}_{j_n} dx^{i_n} E_{j_n} + (\text{higher order terms}) \\ &= (-1)^{[n/2]} \sum a^{i_1}_{j_1} \cdots a^{i_n}_{j_n} (dx^{i_1} \wedge \cdots \wedge dx^{i_1})(E_{j_1} \wedge \cdots \wedge E_{j_n}) \\ &\qquad + (\text{higher order terms}),\end{aligned}$$

the sign coming from commuting the dx_i's and the E_j's. An easy exercise in determinants shows that

$$(DV)^n = (-1)^{[n/2]} n! \det(a^i_j)(dx^1 \wedge \cdots dx^n)(E_1 \wedge \cdots \wedge E_n) \\ + (\text{higher order terms}),$$

and hence

$$\begin{aligned} \tau(V, D) &= \frac{(-1)^{[n/2]}}{\pi^{n/2}} \mathrm{Tr}_s \left(e^{-|V|^2} \frac{1}{n!} (DV)^n \right) \\ &= \frac{1}{\pi^{n/2}} e^{-|V|^2} \det(a^i_j)(dx^1 \wedge \cdots dx^n) \\ &\qquad + (\text{higher order terms}). \end{aligned}$$

After the change of variables $t_i = \sum a^i_j x^j$, this becomes

$$\tau(V, D) = \frac{1}{\pi^{n/2}} e^{-(t_1^2 + \cdots + t_n^2)} (dt_1 \wedge \cdots dt_n) + (\text{higher order terms}).$$

Finally, we replace V by V/ε (or equivalently replace t_i by t_i/ε) and note that the higher order terms go to zero as $\varepsilon \to 0$. Thus it follows from (4.39) that

$$\lim_{\varepsilon \to 0} \int_{B_\delta(p)} \tau\left(\frac{V}{\varepsilon}, D\right) = \begin{cases} 1, & \text{if } \det(a_{ij}) > 0, \\ -1, & \text{if } \det(a_{ij}) < 0, \end{cases}$$

the two cases corresponding to whether the coordinates $(t_1, \ldots, t_n)$ restrict to positively or negatively oriented coordinates on the fiber over p. But this is just the rotation index of V at p. Adding together the contributions at all the zeros of V yields (4.46).

Since the integral of $\tau(V, D)$ over M is independent of V, it follows from (4.45) and (4.46) that if V is a section of E with finitely many nondegenerate zeros at the points p_1, p_2, …, p_k, then (4.37) holds:

$$\frac{1}{(2\pi)^m} \int_M \mathrm{Pf}(\Omega) = \sum_{i=1}^k \omega(V, p_i).$$

If $E = TM$, we can now finish the proof of the generalized Gauss-Bonnet Theorem as sketched at the end of the previous section. If E is an arbitrary oriented real vector bundle over M with $\mathrm{rank}(E) = \dim(M)$, the above formula gives an interpretation for the evaluation of the Euler class of E evaluated on the fundamental class of M.

Remark. If $\dim(M) = \mathrm{rank}(E)$ is odd, the same argument shows that if V is a section of E with finitely many nondegenerate zeros at the points p_1, p_2, …, p_k, then

$$\sum_{i=1}^k \omega(V, p_i) = 0. \tag{4.47}$$

Thus we say that the Euler characteristic of odd-dimensional oriented compact manifolds is zero.

Exercise XVII. a. Suppose that $f : M \to \mathbb{R}$ is a smooth function and that p is a critical point for f, that is $df|p = 0$. Show that we can define a symmetric bilinear form

$$d^2 f(p) : T_p M \times T_p M \to \mathbb{R} \quad \text{by} \quad d^2 f(p)(x, y) = X(Y(f))(p),$$

whenever X and Y are vector fields on M such that $X(p) = x$ and $Y(p) = y$. We call $d^2 f(p)$ the Hessian of f at the critical point p. Moreover, we say that $f : M \to \mathbb{R}$ is *Morse nondegenerate* if $d^2 f(p)$ is a nondegenerate symmetric bilinear form for each critical point p.

b. Suppose now that M has a Riemannian metric $\langle \cdot, \cdot \rangle$ and define a vector field $V = \text{grad}(f)$ on M by

$$df = \langle V, \cdot \rangle.$$

Show that if f is Morse nondegenerate, then the zeros of the gradient V are nondegenerate.

c. We say that a nondegenerate critical point p for f has *Morse index* λ if λ is the maximal dimension of linear subspaces $W \subseteq T_p M$ on which $d^2 f(p)$ is negative definite. Show that if p is a nondegenerate critical point of Morse index λ, then V has rotation index $(-1)^\lambda$ at p.

d. Show that there is a function $f : S^n \to \mathbb{R}$ with exactly two critical points, both nondegenerate, one a local maximum and one a local minimum. Conclude that if n is even, $\chi(S^n) = 2$.

e. Recall that S^n has a Riemannian metric of constant curvature one, for which the curvature forms are given by the formula

$$\Omega_{ij} = \theta_i \wedge \theta_j.$$

Use the Gauss-Bonnet formula to calculate the volume of S^n when it is given this metric.

f. Let M be a compact oriented Riemannian manifold of even dimension $2m$ and constant negative curvature. Show that $(-1)^m \chi(M)$ is positive.

Remark. By triangulating a compact oriented Riemannian manifold M of dimension n, and using an appropriate gradient vector field, one can show that

$$\chi(M) = \sum_{k=0}^{n} (-1)^k \dim H^k(M; \mathbb{R}).$$

Historical remarks. The Generalized Gauss-Bonnet Theorem was first proven by Allendoerfer [1] under the assumption that M admits an isometric imbedding into some Euclidean space $\mathbb{E}^N$. (This would have proven the theorem for general Riemannian manifolds if Nash's imbedding theorem had been known at the time.) Allendoerfer and Weil [2] then gave a proof for a "Riemannian polyhedron" which is isometrically imbedded in $\mathbb{E}^N$, which was complicated by the fact that they used the classical tensor analysis in terms of components. Since it was known that real analytic Riemannian manifolds admit "local" isometric imbeddings into Euclidean space, one could then piece together the result for Riemannian polyhedra to give a proof for all real analytic Riemannian manifolds. In a famous article, Chern [8] gave an intrinsic proof of the Gauss-Bonnet formula for smooth Riemannian manifolds, obtaining a major simplification in the proof by using Cartan's method of moving frames. Indeed, the proof of Chern was a first step towards the theory of characteristic classes, to be described in the next chapter. A more conceptual proof of the generalized Gauss-Bonnet Theorem was later found by Quillen [30] using the theory of Clifford algebras and "superconnections." We have followed the main ideas of Quillen's proof, except that we have used the exterior algebra instead of the Clifford algebra.

Chapter 5

Characteristic classes

5.1 The Chern character

In §4.10 we described how one can classify complex line bundles over a given smooth manifold M up to isomorphism. It is interesting problem in differential topology to classify vector bundles of higher rank over a given manifold. Thus we would like to classify

$$\mathcal{V}^m(M) = \{ \text{ isomorphism classes of real rank } m \text{ vector bundles over } M \ \},$$

$$\mathcal{V}^m_+(M) = \{ \text{ isomorphism classes of oriented rank } m \text{ vector bundles over } M \ \},$$

$$\mathcal{V}^m_{\mathbb{C}}(M) = \{ \text{ isomorphism classes of complex rank } m \text{ vector bundles over } M \ \}.$$

Note that any real vector bundle has a fiber metric (which can be constructed via a partition of unity) and the reader can easily verify that any two fiber metrics in the same vector bundle are equivalent via a vector bundle automorphism. Similar facts hold for Hermitian metrics in complex vector bundles. Thus we can regard $\mathcal{V}^m(M)$ as isomorphism classes of $O(n)$-bundles, element of $\mathcal{V}^m_+(M)$ as isomorphism classes of $SO(n)$-bundles, and elements of $\mathcal{V}^m_{\mathbb{C}}(M)$ as isomorphism classes of $U(n)$-bundles.

Let us first consider the case $\mathcal{V}^m_{\mathbb{C}}(M)$. Our previous experience with classifying line bundles suggests that we should try to develop analogs of the first Chern class for complex vector bundles of higher rank.

In contrast to complex line bundles, the curvature matrices Ω_α for connections in complex vector bundles of higher rank are only locally defined. However, it is possible to construct certain polynomials in Ω_α that are invariant under the transformation (4.27) and these give rise to topological invariants that often enable us to distinguish between nonisomorphic vector bundles over M.

Let us focus first on the case of a $U(m)$-vector bundle E, a complex vector bundle of rank m with a Hermitian metric. We give E a metric connection D which has local representatives $d + \omega_\alpha$ and curvature matrices Ω_α. As we have seen before, the matrices Ω_α are skew-Hermitian, so the matrices

$$\frac{i}{2\pi}\Omega_\alpha \quad \text{and} \quad \left(\frac{i}{2\pi}\Omega_\alpha\right)^k$$

are Hermitian. Hence for each choice of positive integer k, the differential form

$$\tau_k(\Omega_\alpha) = \text{Trace}\left[\left(\frac{i}{2\pi}\Omega_\alpha\right)^k\right]$$

is real-valued, and since trace is invariant under similarity, it follows from (4.27) that

$$\text{Trace}\left[\left(\frac{i}{2\pi}\Omega_\alpha\right)^k\right] = \text{Trace}\left[\left(\frac{i}{2\pi}g_{\alpha\beta}\Omega_\alpha g_{\alpha\beta}^{-1}\right)^k\right] = \text{Trace}\left[\left(\frac{i}{2\pi}\Omega_\beta\right)^k\right]$$

on $U_\alpha \cap U_\beta$. Hence the locally defined forms $\tau_k(\Omega_\alpha)$ fit together into a globally defined real-valued $2k$-form $\tau_k(R)$ on the base space M, where R is the curvature of the connection D. We say that $\tau_k(R)$ is a *characteristic form*.

Lemma. *The differential form $\tau_k(R)$ is closed.*

Proof: It follows from the Bianchi identity that

$$\begin{aligned} d(\tau_k(R)) &= \left(\frac{i}{2\pi}\right)^k d[\text{Trace}(\Omega_\alpha)^k] = \left(\frac{i}{2\pi}\right)^k [\text{Trace} d(\Omega_\alpha)^k] \\ &= \left(\frac{i}{2\pi}\right)^k \text{Trace}(d\Omega_\alpha)\Omega_\alpha^{k-1} + \cdots + \Omega_\alpha^{k-1} d\Omega_\alpha) \\ &= \left(\frac{i}{2\pi}\right)^k \text{Trace}([\Omega_\alpha, \omega_\alpha]\Omega_\alpha^{k-1} + \cdots + \Omega_\alpha^{k-1}[\Omega_\alpha, \omega_\alpha]) = 0, \end{aligned}$$

the last equality coming from the fact that $\text{Trace}(A_1 \cdots A_{k+1})$ is invariant under cyclic permutation of $A_1, \ldots, A_{k+1}$.

This Lemma implies that $\tau_k(R)$ represents a de Rham cohomology class

$$[\tau_k(R)] \in H^{2k}(M; \mathbb{R}),$$

which can be thought of as a generalization of the first Chern class.

Suppose that E is a $U(m)$-vector bundle over M with unitary connection D_E having curvature R_E, and $F : N \to M$ is a smooth map. As we saw in the preceding section, the local representatives Ω_α for the curvature of D_E for a trivializing cover $\{U_\alpha : \alpha \in A\}$ for E pull back to local representatives $F^*\Omega_\alpha$ for the curvature R_{F^*E} of the pullback connection D_{F^*E} on F^*E. Thus the characteristic forms for D_E pull back to the characteristic forms for D_{F^*E}:

$$\tau_k(R_{F^*E}) = F^*\tau_k(R_E) \quad \text{and hence} \quad [\tau_k(R_{F^*E})] = F^*[\tau_k(R_E)] \in H^{2k}N; \mathbb{R}). \tag{5.1}$$

Proposition 1. *The de Rham cohomology class $[\tau_k(R)]$ is independent of the choice of unitary connection on E, as well as the choice of Hermitian metric on E.*

To prove this we use the cylinder construction from topology. From the bundle $\pi : E \to M$ we construct the cylinder bundle

$$\pi \times \mathrm{id} : E \times [0,1] \longrightarrow M \times [0,1].$$

If E possesses the trivializing cover $\{U_\alpha : \alpha \in A\}$, then $E \times [0,1]$ possesses the trivializing cover $\{U_\alpha \times [0,1] : \alpha \in A\}$.

Suppose that D_0 and D_1 are two metric connections on E, with curvatures R_0 and R_1 respectively, which have local representatives $d + \omega_\alpha^0$ and $d + \omega_\alpha^1$ over U_α. If $\pi_1 : U_\alpha \times [0,1] \to U_\alpha$ is the projection on the first factor, then the local representatives

$$d + (1-t)\pi_1^* \omega_\alpha^0 + t\pi_1^* \omega_\alpha^1 \quad \text{over} \quad U_\alpha \times [0,1]$$

fit together to form a metric connection D' on $E \times [0,1]$ with curvature R' such that

$$J_0^*(\tau_k(R')) = R_0 \quad \text{and} \quad J_1^*(\tau_k(R')) = R_1,$$

where $J_0, J_1 : M \to M \times [0,1]$ are the maps defined by $J_0(p) = (p,0)$ and $J_1(p) = (p,1)$. Since J_0 and J_1 are homotopic, it follows from the Homotopy Theorem from de Rham cohomology that

$$[\tau_k(R_0)] = J_0^*[\tau_k(R')] = J_1^*[\tau_k(R')] = [\tau_k(R_1)].$$

This shows that $[\tau_k(R)]$ is independent of the choice of unitary connection.

The proof that $[\tau_k(R)]$ is independent of the choice of Hermitian metric is similar. In this case, we use the fact that any two Hermitian metrics $\langle \cdot, \cdot \rangle_0$ and $\langle \cdot, \cdot \rangle_1$ can be connected by a one-parameter family

$$\langle \cdot, \cdot \rangle_t = (1-t)\langle \cdot, \cdot \rangle_0 + t\langle \cdot, \cdot \rangle_1.$$

According to Proposition 1, to each complex vector bundle E of rank m over M we can associate a collection of cohomology classes

$$\tau_k(E) = [\tau_k(R)] \in H^{2k}(M; \mathbb{R})$$

and by (5.1), these are natural under smooth maps:

$$\tau_k(F^* E) = F^* \tau_k(E).$$

These are called the *characteristic classes* of the complex vector bundle E. We can put these characteristic classes together to form the *Chern character*

$$\begin{aligned} \mathrm{ch}(E) &= \left[\mathrm{Trace}\left(\exp \frac{i}{2\pi} R \right) \right] \\ &= \mathrm{rank}(E) + \tau_1(E) + \frac{1}{2!}\tau_2(E) + \cdots + \frac{1}{k!}\tau_k(E) + \cdots , \end{aligned}$$

which is an element of the de Rham cohomology ring of the base manifold M:

$$\mathrm{ch}(E) \in H^*(M;\mathbb{R}) = H^0(M;\mathbb{R}) \oplus H^1(M;\mathbb{R}) \oplus \cdots \oplus H^k(M;\mathbb{R}) \oplus \cdots .$$

The Chern character collapses to a polynomial because all of the terms of degree larger than $\dim M$ must vanish.

Proposition 2. *If E and F are complex vector bundles over M, then*

$$ch(E \oplus F) = ch(E) + ch(F), \qquad ch(E \otimes F) = ch(E)ch(F).$$

Sketch of proof: Suppose that E and F have Hermitian metrics with metric connections D_E and D_F. Then $E \oplus F$ inherits a direct sum metric and direct sum connection $D_{E\oplus F}$, the latter defined by

$$D_{E\oplus F}(\sigma_E \oplus \sigma_F) = (D_E\sigma_E) \oplus (D_F\sigma_F).$$

It follows that

$$D^2_{E\oplus F}(\sigma_E \oplus \sigma_F) = (D^2_E\sigma_E) \oplus (D^2_F\sigma_F),$$

and hence if $\{U_\alpha : \alpha \in A\}$ is an open cover of M which simultaneously trivialises both E and F,

$$\left(\tfrac{i}{2\pi}\Omega^{E\oplus F}_\alpha\right) = \begin{pmatrix} \frac{i}{2\pi}\Omega^E_\alpha & 0 \\ 0 & \frac{i}{2\pi}\Omega^F_\alpha \end{pmatrix}.$$

It follows that

$$\left(\tfrac{i}{2\pi}\Omega^{E\oplus F}_\alpha\right)^k = \begin{pmatrix} \left(\frac{i}{2\pi}\Omega^E_\alpha\right)^k & 0 \\ 0 & \left(\frac{i}{2\pi}\Omega^F_\alpha\right)^k, \end{pmatrix}$$

and hence

$$\tau_k(R_{E\oplus F}) = \tau_k(R_E) + \tau_k(R_F),$$

which implies the first of the two assertions.

Similarly, $E \otimes F$ inherits a product metric and product connection $D_{E\otimes F}$ defined so that the Leibniz rule is satisfied,

$$D_{E\otimes F}(\sigma_E \otimes \sigma_F) = (D_E\sigma_E) \otimes \sigma_F + \sigma_E \otimes (D_F\sigma_F).$$

Differentiating a second time yields

$$\begin{aligned} D^2_{E\otimes F}(\sigma_E \otimes \sigma_F) = (D^2_E\sigma_E) \otimes \sigma_F &- (D_E\sigma_E) \otimes (D_F\sigma_F) \\ &+ (D_E\sigma_E) \otimes (D_F\sigma_F) + \sigma_E \otimes (D^2_F\sigma_F), \end{aligned}$$

and hence

$$R_{E\otimes F}(\sigma_E \otimes \sigma_F) = (R_E\sigma_E) \otimes \sigma_F + \sigma_E \otimes (R_F\sigma_F).$$

By induction, we establish that

$$R^k_{E\otimes F}(\sigma_E \otimes \sigma_F) = \sum_{j=0}^{k} \binom{k}{j} (R^j_E\sigma_E) \otimes (R^{k-j}_F\sigma_F),$$

or

$$\frac{1}{k!}R^k_{E\otimes F}(\sigma_E \otimes \sigma_F) = \sum_{j=0}^{k} \frac{1}{j!}(R^j_E \sigma_E) \otimes \frac{1}{(k-j)!}(R^{k-j}_F \sigma_F).$$

We conclude that

$$\exp\left(\frac{i}{2\pi}R_{E\otimes F}\right) = \exp\left(\frac{i}{2\pi}R_E\right) \otimes \exp\left(\frac{i}{2\pi}R_F\right),$$

and taking the trace of both sides yields the second assertion.

Remark. The operations of direct sum as addition and tensor product as multiplication make the spaces $\mathcal{V}^m(M)$, $\mathcal{V}^m_+(M)$ and $\mathcal{V}^m_{\mathbb{C}}(M)$ into semirings, where semirings are defined so that they satisfy all of the axioms for rings except for the existence of an additive inverse. There is a process of completing a semiring to form a genuine ring, similar to the construction that creates the integers from the nonnegative integers. When we apply that process to the semiring $\mathcal{V}^m_{\mathbb{C}}(M)$, we obtain a semiring $K(M)$, called the K-theory of M.

In more detail, if M is a smooth manifold, we let $K(M)$ denote the space of equivalence classes of pairs (E, F), where E and F are complex vector bundles and the equivalence relation is defined by

$$(E_1, F_1) \sim (E_2, F_2) \quad \Leftrightarrow \quad E_1 \oplus F_2 \text{ is isomorphic to } E_2 \oplus F_1.$$

The equivalence class of the pair (E, F) is denoted by $[E]-[F] \in K(M)$ and can be thought of as a difference of vector bundles or as a "virtual" vector bundle over M. A smooth map $F : M \to N$ induces a ring homomorphism

$$F^* : K(N) \to K(M)$$

by pulling back vector bundles, and the reader can verify easily that the correspondence

$$M \mapsto K(M), \quad (F : M \to N) \mapsto (F^* : K(N) \to K(M))$$

is a contravariant functor from the category of smooth manifolds and smooth manifold maps to the category of rings and ring homomorphisms. Moreover, it follows from Proposition 2 from §4.9 that if $F_0 : M \to N$ and $F_1 : M \to N$ are smoothly homotopic, then

$$F_0^* = F_1^* : K(N) \to K(M).$$

In particular, if M is contractible, $K(M) \cong \mathbb{Z}$.

Now Proposition 2 implies that the Chern character induces a ring homomorphism

$$\mathrm{ch} : K(M) \to H^*(M; \mathbb{R}).$$

In fact, it was proven by Atiyah and Hirzebruch that the Chern character yields a ring *isomorphism*

$$\mathrm{ch} : K(M) \otimes \mathbb{Q} \to H^{\mathrm{ev}}(M; \mathbb{Q}) \otimes \mathbb{Q}, \quad \text{where} \quad H^{\mathrm{ev}}(M; \mathbb{Q}) = \sum_{i=0}^{\infty} H^{2i}(M; \mathbb{Q}).$$

5.2 Chern classes

There is another useful way of constructing the characteristic classes for $U(m)$-bundles. In this approach, one starts by defining the *Chern polynomials*

$$c_k : \mathfrak{u}(m) \to \mathbb{R} \quad \text{by} \quad \det\left(\frac{itA}{2\pi} + I\right) = \sum_{k=0}^{m} c_k(A)t^k.$$

The next Proposition shows that each of these polynomials can be expressed in terms of the trace polynomials τ_k that we have already constructed:

Proposition 1. *The Chern polynomials can be expressed in terms of the trace polynomials by means of the following* Newton identity*:*

$$kc_k = \sum_{i=1}^{k} (-1)^{i-1} c_{k-i}\tau_i, \quad \text{for } 1 \le k \le m.$$

Proof: For any element of $A \in \mathfrak{u}(m)$, there is an element $B \in U(m)$ such that

$$B^{-1}AB = i \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{pmatrix},$$

it suffices to show that

$$s_k(\lambda_1, \ldots, \lambda_m) = \sum_{i=1}^{k} (-1)^{i-1} s_{k-i}(\lambda_1, \ldots, \lambda_m) p_i(\lambda_1, \ldots, \lambda_m), \tag{5.2}$$

where

$$p_i(\lambda_1, \ldots, \lambda_m) = \lambda_1^i + \cdots + \lambda_m^i,$$

and $s_k(\lambda_1, \ldots, \lambda_m)$ is the k-th elementary symmetric function,

$$s_1(\lambda_1, \ldots, \lambda_m) = \lambda_1 + \cdots + \lambda_m,$$

$$s_2(\lambda_1, \ldots, \lambda_m) = \sum_{i<j} \lambda_i\lambda_j, \quad \cdots$$

$$s_m(\lambda_1, \ldots, \lambda_m) = \lambda_1\lambda_2 \cdots \lambda_m.$$

To prove (5.2) one starts with the identity

$$\Pi_{i=1}^{m}(t - \lambda_i) = \sum_{i=0}^{m} (-1)^{m-i} s_{m-i}(\lambda_1, \ldots, \lambda_m) t^i.$$

Substituting λ_j for t yields

$$0 = \sum_{i=0}^{m} (-1)^{m-i} s_{m-i}(\lambda_1, \ldots, \lambda_m) \lambda_j^i,$$

and summing over j yields

$$0 = (-1)^m s_m(\lambda_1, \dots, \lambda_m) + \sum_{i=1}^{m} (-1)^{m-i} s_{m-i}(\lambda_1, \dots, \lambda_m) p_m(\lambda_1, \dots, \lambda_m).$$

This is just the Newton identity (5.2) for $k = m$, that is in the case where the number of variables equals k.

The general case of the Newton identity follows from the case $k = m$. Indeed, in the identity for s_k one need only show that the coefficients of $\lambda_{i_1} \cdots \lambda_{i_k}$ on both sides of (5.2) are the same. But this follows from the Newton identity when the number of variables is k.

By induction we can show that each c_k is a polynomial in the τ_k's and hence by the Lemma from the previous section, we can conclude that if E is a complex vector bundle over M with Hermitian metric and metric connection D_E, then the globally defined $2k$-form

$$c_k(R_E) = c_k(\Omega_\alpha)$$

is closed. Here R_E is the curvature of D_E and the Ω_α's are local representatives with respect to given trivializations. Moreover, it follows Proposition 1 of the previous section that the de Rham cohomology class

$$c_k(E) = [c_k(R_E)] \in H^{2k}(M; \mathbb{R})$$

is independent of choice of Hermitian metric or metric connection.

Definition. If E is a complex vector bundle over M, then the k-th *Chern class* of E is the cohomology class

$$c_k(E) = [c_k\,(R_E)] \in H^{2k}(M; \mathbb{R})$$

constructed in the previous paragraph.

Proposition 2. *If E and F are complex vector bundles over M, then*

$$c_k(E \oplus F) = \sum_{i=0}^{k} c_i(E) c_{k-i}(F), \tag{5.3}$$

where the multiplication on the right is the cup product.

Note that if we write

$$\begin{aligned} c(E) &= 1 + c_1(E) + c_2(E) + \cdots + c_k(E) + \cdots \\ &\in H^*(M; \mathbb{R}) = H^0(M; \mathbb{R}) \oplus H^1(M; \mathbb{R}) \oplus \cdots \oplus H^k(M; \mathbb{R}) \oplus \cdots , \end{aligned}$$

then (5.3) simplifies to $c(E \oplus F) = c(E)c(F)$.

Here is a sketch of the proof: Suppose that E and F have a common trivilializing open cover $\{U_\alpha : \alpha \in A\}$ and that the curvatures of metric connections in

E and F have local representative Ω_α^E and Ω_α^F respectively. Then the local representatives of the curvature in $E \oplus F$ are

$$\Omega_\alpha^{E\oplus F} = \begin{pmatrix} \Omega_\alpha^E & 0 \\ 0 & \Omega_\alpha^F \end{pmatrix},$$

and by definition of the Chern polynomials

$$\sum_{k=0}^{\infty} c_k(\Omega_\alpha^{E\oplus F}) t^k = \det\left(\frac{it\Omega_\alpha^{E\oplus F}}{2\pi} + I\right)$$
$$= \det\left(\frac{it\Omega_\alpha^{E}}{2\pi} + I\right) \det\left(\frac{it\Omega_\alpha^{F}}{2\pi} + I\right) = \sum_{i=0}^{\infty} c_i(\Omega_\alpha^E) t^i \sum_{j=0}^{\infty} c_j(\Omega_\alpha^F) t^j.$$

Comparing the coefficients of t^k on the two sides of the equation yields (5.3).

Remark. Use of a Hermitian metric on a complex vector bundle E of rank m over M shows that the dual bundle E^* to a $U(m)$-bundle over M is obtained by conjugating the transition functions

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta}.$$

Thus we can think of the dual bundle E^* as the *conjugate* of E. A connection D_E on E defines a conjugate connection D_{E^*} on E^* by conjugating local representatives

$$d + \omega_\alpha \quad \mapsto \quad d + \bar{\omega}_\alpha.$$

Of course, the local representatives of the curvature are also obtained by conjugation,

$$\Omega_\alpha \quad \mapsto \quad \bar{\Omega}_\alpha,$$

from which it follows that

$$c_k(E^*) = (-1)^k c_k(E). \tag{5.4}$$

5.3 Examples of Chern classes

Recall that we can define complex n-dimensional projective space $P^n\mathbb{C}$ as the space of one-dimensional subspaces of the complex vector space $\mathbb{C}^{n+1}$. More precisely, we define an equivalence relation $\sim$ on $\mathbb{C}^{n+1} - \{0\}$ by

$$(z_0, z_1, \ldots, z_n) \sim (w_0, w_1, \ldots, w_n) \quad \Leftrightarrow \quad z_i = \lambda w_i, \quad \text{for some } \lambda \in \mathbb{C} - \{0\}.$$

We let $[z_0, z_1, \ldots, z_n]$ denote the equivalence class of $(z_0, z_1, \ldots, z_n)$ and let

$$U_i = \{[z_0, z_1, \ldots, z_n] : z_i \neq 0\}.$$

We can then define a bijection $\phi_i : U_i \to \mathbb{C}^n$ by

$$\phi_i([z_0, z_1, \ldots, z_n]) = \left(\frac{z_0}{z_i}, \frac{z_1}{z_i}, \ldots, \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, \ldots, \frac{z_n}{z_i}\right).$$

We topologize $P^n\mathbb{C}$ so that every U_i is an open set and each ϕ is a homeomorphism from U_i to $\mathbb{C}^n$. Then $\phi_i \circ \phi_j^{-1}$ is holomorphic where defined, so $P^n\mathbb{C}$ is a compact complex manifold, as defined at the beginning of §4.2. Note that $P^n\mathbb{C} - U_n \cong P^{n-1}\mathbb{C}$. and we have inclusions

$$S^2 = P^1\mathbb{C} \subseteq P^2\mathbb{C} \subseteq \cdots \subseteq P^{n-1}\mathbb{C} \subseteq P^n\mathbb{C} \subseteq \cdots .$$

Exercise XVIII. a. Let V_n be a small neighborhood of $P^{n-1}\mathbb{C}$ in $P^n\mathbb{C}$. Show that $U_n \cap V_n$ is homotopy equivalent to S^{2n-1}.

b. Use induction and the Mayer-Vietoris sequence to establish that

$$H^k_{dR}(P^n\mathbb{C};\mathbb{R}) \cong \begin{cases} \mathbb{R}, & \text{if } k = 0,\ k = 2,\ \ldots,\ k = 2n, \\ 0, & \text{otherwise.} \end{cases}$$

c. Let α be the generator of $H^2_{dR}(P^n\mathbb{C};\mathbb{R})$ such that if $i : S^2 = P^1\mathbb{C} \to P^n\mathbb{C}$ is the inclusion, then $\int_{S^2} i^*\alpha = 1$. Use Poincaré duality and induction to show that

$$\alpha^k = \overbrace{\alpha \cup \alpha \cup \cdots \cup \alpha}^{k} \neq 0, \quad \text{for } k \leq n, \text{ and } \quad \alpha^{n+1} = 0.$$

Thus we can say that the de Rham cohomology algebra of $P^n\mathbb{C}$ is a truncated polynomial ring $p[\alpha]/(\alpha^{n+1} = 0)$, where the generator α lies in $H^2(P^n\mathbb{C};\mathbb{R})$.

We can use the results of Exercise XIV to calculate the Chern classes of the complex projective spaces. To do this, we first note that over $P^n\mathbb{C}$ there are two important bundles,

$$(\text{the universal bundle}) = E_\infty = \{(V, v) \in P^n\mathbb{C} \times \mathbb{C}^{n+1} : v \in V\}$$

and

$$E_\infty^\perp = \{(V, v) \in P^n\mathbb{C} \times \mathbb{C}^{n+1} : v \perp V\},$$

with $E_\infty \oplus E_\infty^\perp \cong \underline{\mathbb{C}}^{n+1}$, where $\underline{\mathbb{C}}^{n+1}$ is the trivial bundle of rank $n+1$ over $P^n\mathbb{C}$.

The transition functions for $TP^n\mathbb{C}$ take values in $GL(n,\mathbb{C})$, so $TP^n\mathbb{C}$ is a complex vector bundle of rank n over $P^n\mathbb{C}$.

Lemma. *We have an isomorphism between complex vector bundles*

$$TP^n\mathbb{C} \cong Hom(E_\infty, E_\infty^\perp)$$

Sketch of proof: Suppose that L is an element in the fiber $\mathrm{Hom}(E_\infty, E_\infty^\perp)_V$ over V, where V is an element of $P^n\mathbb{C}$, so $L : (E_\infty)_V \to (E_\infty^\perp)_V$ is a complex linear map. Define

$$\gamma_L : \mathbb{C} \to P^n\mathbb{C} \quad \text{by} \quad \gamma_L(z) = [v + zL(v)], \quad \text{for } v \in V.$$

Note that γ_L is well-defined because

$$[\lambda v + zL(\lambda v)] = [\lambda(v + zL(v))] = [v + zL(v)],$$

and we can restrict γ_L to a real curve $\mu_L : \mathbb{R} \to P^n\mathbb{C}$. Now we define a vector bundle map

$$F : \mathrm{Hom}(E_\infty, E_\infty^\perp) \to TP^n\mathbb{C} \quad \text{by} \quad F(L) = \frac{d}{dt}\mu_L(t)\bigg|_{t=0},$$

and check that L is in fact a vector bundle isomorphism.

We can use this lemma to calculate the Chern classes of the bundle $TP^n\mathbb{C}$. Indeed,

$$\begin{aligned} TP^n\mathbb{C} \oplus \underline{\mathbb{C}} &\cong \mathrm{Hom}(E_\infty, E_\infty^\perp) \oplus \mathrm{Hom}(E_\infty, E_\infty) \\ &\cong \mathrm{Hom}(E_\infty, E_\infty \oplus E_\infty^\perp) \cong \mathrm{Hom}(E_\infty, \underline{\mathbb{C}}^{n+1}) = \overbrace{H \oplus \cdots \oplus H}^{n+1}, \end{aligned}$$

where H denotes the hyperplane bundle (dual to the universal bundle) defined by

$$H = E_\infty^* = \mathrm{Hom}(E_\infty, \underline{\mathbb{C}}).$$

Thus we find that

$$\begin{aligned} c(TP^n\mathbb{C}) = c(TP^n\mathbb{C} \oplus \underline{\mathbb{C}}) &\cong c\left(\overbrace{H \oplus \cdots \oplus H}^{n+1}\right) \\ &= (c(H)^{n+1}) = (1 + c_1(H))^{n+1} = (1 + k\alpha)^{n+1}, \end{aligned}$$

where $k\alpha = c_1(H)$ and $k \in \mathbb{R}$. To evaluate k, we pull back to $S^2 = P^1\mathbb{C}$ and use the fact that

$$\int_{S^2} c_1(TS^2) = \int_{S^2} c_1(H^2) = 2.$$

We must therefore set $k = 1$, and conclude that

$$c(TP^n\mathbb{C}) = (1 + \alpha)^{n+1}.$$

5.4 Invariant polynomials

Our next goal is to extend the previous theory to real vector bundles or oriented real vector bundles. More generally, we can let G be a Lie subgroup of $GL(m, \mathbb{R})$ with Lie algebra $\mathfrak{g} \subseteq \mathfrak{g}l(m, \mathbb{R})$. A (real-valued) *invariant polynomial* for G is a polynomial function

$$p : \mathfrak{g} \to \mathbb{R} \quad \text{such that} \quad p(B^{-1}AB) = p(A),$$

for $A \in \mathfrak{g}$ and $B \in G$. If E is a G-vector bundle with G-connection D_E which has $\mathfrak{g}$-valued local curvature representatives Ω_α, then just as in the case of $U(m)$, $p(\Omega_\alpha)$ will be a globally defined differential form on the base M.

For $U(m) \subseteq GL(m, \mathbb{C}) \subseteq GL(2m, \mathbb{R})$, the key examples are the trace polynomials

$$\tau_k : \mathfrak{u}(m) \to \mathbb{R} \quad \text{defined by} \quad \tau_k(A) = \text{Trace}\left[\left(\frac{i}{2\pi}A\right)^k\right]$$

or the Chern polynomials

$$c_k : \mathfrak{u}(m) \to \mathbb{R} \quad \text{defined by} \quad \det\left(\frac{itA}{2\pi} + I\right) = \sum_{k=0}^{m} c_k(A)t^k.$$

For general choice of G, if $p : \mathfrak{g} \to \mathbb{R}$ is an invariant polynomial of degree k, we can polarize to obtain a functions of k variables

$$p : \overbrace{\mathfrak{g} \times \cdots \times \mathfrak{g}}^{k} \to \mathbb{R}, \quad \text{defined so that} \quad p(A, \ldots, A) = p(A).$$

Recall that if $\sigma \in G$, we can define an inner automorphism $a(\sigma) \in \text{Aut}(G)$ by $a(\sigma)(\tau) = \sigma\tau\sigma^{-1}$. The differential of $a(\sigma)$ at the identity gives a Lie algebra isomorphism

$$\text{Ad}(\sigma) = R^*_{\sigma^{-1}} \in GL(\mathfrak{g}),$$

where $GL(\mathfrak{g})$ is just the space of all vector space isomorphisms of $\mathfrak{g}$, which is of course a Lie group itself. Then $\text{Ad} : G \to GL(\mathfrak{g})$ is a Lie group homomorphism called the *adjoint representation*. It induces a Lie algebra homomorphism

$$\text{ad} : \mathfrak{g} \to \mathfrak{g}l(\mathfrak{g}),$$

also called the adjoint representation, such that

$$\begin{aligned} \text{ad}(X)(Y) = \frac{d}{dt}(\text{ad}(\exp(tX))\Big|_{t=0}(Y) \\ = \frac{d}{dt}(R^*_{\exp(-tX)}(Y)\Big|_{t=0} = -\frac{d}{dt}(\phi^*_t(Y))\Big|_{t=0}, \end{aligned}$$

where $\{\phi_t : t \in \mathbb{R}\}$ is the one-parameter group of diffeomorphisms corresponding to X. Thus

$$\text{ad}(X)(Y) = -\frac{d}{dt}(\phi^*_t(Y))\Big|_{t=0} = [X, Y].$$

Thus, for example, if $G = U(m)$,

$$a(A)B = ABA^{-1}, \quad \text{for } A, B \in U(m),$$

$$\text{Ad}(A)X = AXA^{-1}, \quad \text{for } A \in U(m) \text{ and } X \in \mathfrak{u}(m),$$

$$\mathrm{ad}(X)Y = [X, Y] = XY - YX, \quad \text{for } X, Y \in \mathfrak{u}(m),$$

If p is an invariant polynomial, its polarization satisfies

$$p(\mathrm{Ad}(\sigma)X_1, \ldots \mathrm{Ad}(\sigma)X_k) = p(A_1, \ldots X_k),$$

for $\sigma \in G$ and $X_1, \ldots, X_k \in \mathfrak{g}$. We can differentiate this identity with respect to σ to obtain

$$p([Y, X_1], X_2, \ldots, X_k) + p(X_1, [Y, X_2], \ldots, X_k) \\ + \cdots + p(A_1, X_2, \ldots, [Y, X_k]) = 0. \quad (5.5)$$

Thus is E is a G-bundle and D is a G-connection, where $G \subseteq GL(m, \mathbb{C})$ or $GL(m, \mathbb{R})$ and p is a real-valued invariant polynomial for G, then $p(\Omega_\alpha, \cdots \Omega_\alpha)$ is a globally defined $2k$-form on M and it follows from the Bianchi identity and (5.5) that

$$d(p(\Omega_\alpha, \cdots \Omega_\alpha) = p(d\Omega_\alpha, \cdots, \Omega_\alpha) + \cdots + p(\Omega_\alpha, \cdots, d\Omega_\alpha) \\ = p([\Omega_\alpha, \omega_\alpha], \cdots, \Omega_\alpha) + \cdots + p(\Omega_\alpha, \cdots, [\Omega_\alpha, \omega_\alpha]) = 0.$$

Thus an invariant polynomial for G determines a closed characteristic form for each G-bundle. If p is an invariant polynomial for G, then the argument in §5.1 shows that the cohomology class $[p(\Omega_\alpha)]$ of the characteristic form does not depend upon choice of connection, and is natural under smooth maps.

Example 1. In §5.1 and §5.2 we have constructed the invariant polynomials for $G = U(m)$. In this case, the invariant polynomials are polynomials in either the trace polynomials τ_k or the Chern polynomials c_k.

Example 2. In a quite similar fashion, we can consider the orthogonal group $O(m)$, and define invariant polynomials s_k by the identity

$$\det\left(\lambda + \frac{1}{2\pi}X\right) = \sum_{k=0}^{m} \lambda^{m-k} s_k(X).$$

One can show that the odd s_k's are automatically zero, but if E is a real bundle of rank m, then

$$p_k(E) = [s_{2k}(\Omega_\alpha)] \in H^{4k}(M : \mathbb{R})$$

is a characteristic class of E, which may be nonzero, called the k-th *Pontrjagin class*.

But their is an alternate approach to the Pontrjagin classes which is even simpler. If E is a real vector bundle of rank m we can construct its complexification $E \otimes \mathbb{C}$. Note that the complexification is isomorphic to its own conjugate so $c_k(E \otimes \mathbb{C}) = 0$ when k is odd. We can then define the k-th *Pontrjagin class* of E to be the cohomology class

$$p_k(E) = (-1)^k c_{2k}(E \otimes \mathbb{C}).$$

It is easily verified that this gives the same result we obtained before. The sign is necessary in order to be able to prove the formula

$$p_k(E \oplus F) = \sum_{i=0}^{k} p_i(E) p_{k-i}(F),$$

in analogy with (5.3).

Exercise XIX. a. Show that the Pontrjagin classes of the tangent bundle TS^n of the sphere are all zero.

b. Suppose that E is a complex vector bundle of rank two (which can be regarded as a real vector bundle of rank four). Show that

$$p_1(E) = c_1(E)^2 - 2c_2(E).$$

Example 3. For the special orthogonal group $SO(2m)$, we have in addition to the Pontrjagin classes, the Pfaffian, which was described in §4.12 and used in the proof of the Generalized Gauss-Beonnet Theorem.

Definition. If E is an oriented real vector bundle of rank $2m$ over M, then the *Euler class* of E is the cohomology class

$$e(E) = \left[\mathrm{Pf}\left(\frac{\Omega}{2\pi} \right) \right] \in H^{2m}(M; \mathbb{R}).$$

Exercise XX. Show that if M is an m-dimensional complex manifold, then $e(TM) = c_m(TM)$.

5.5 The universal bundle*

You may want to skip this section on a first reading. It gives an alternate approach to the theory of Chern classes (and other characteristic classes as well) which relies on the notion of a "universal bundle."

We start with infinite-dimensional Hilbert space C^∞, regarded as the space of infinite sequences of complex numbers

$$z = (z_1, z_2, \ldots, z_i, \ldots) \quad \text{such that} \quad \sum_{i=1}^{\infty} z_i^2 < \infty.$$

Just as in the finite-dimensional case, two nonzero elements

$$z = (z_1, z_2, \ldots, z_i, \ldots) \quad \text{and} \quad w = (w_1, w_2, \ldots, w_i, \ldots)$$

of C^∞ are said to be equivalent if $z_i = \lambda w_i$ for each i, for some choice of $\lambda \in \mathbb{C} - \{0\}$. This defines an equivalence relation on $C^\infty - \{0\}$ and we let

$P^\infty\mathbb{C}$ denote the space of equivalence classes. If $z = (z_1, z_2, \ldots, z_i, \ldots)$ is an element of $C^\infty - \{0\}$, we let $[z] = [z_1, z_2, \ldots, z_i, \ldots]$ denote the equivalence class containing z.

It is not difficult to check that $P^\infty\mathbb{C}$ satisfies the definition of Hilbert manifold as defined for example in [18]. For the convenience of the reader, we recall the definition. Note that although C^∞ is a complex vector space, we can regard it as a real vector space for the purpose of the following definition.

Definition. Suppose that H_1 and H_2 are real Hilbert spaces, and that U is an open subset of H_1. A continuous map $f : U \to H_2$ is said to be *differentiable* at the point $x_0 \in U$ if there exists a continuous linear map $T : H_1 \to H_2$ such that

$$\lim_{\|h\| \to 0} \frac{\|F(x_0 + h) - F(x_0) - T(h)\|}{\|h\|} = 0,$$

where $\| \cdot \|$ denotes both the Hilbert space norms on H_1 and H_2. We will call T the *derivative* of F at x_0 and write $DF(x_0)$ for T. Note that the derivative satisfies the formula

$$Df(x_0)h = \lim_{t \to 0} \frac{F(x_0 + th) - F(x)}{t}.$$

Just as in ordinary calculus, the derivative $Df(x_0)$ determines the *linearization* of f near x_0, which is the affine function

$$\tilde{F}(x) = F(x_0) + DF(x_0)(x - x_0)$$

which most closely approximates F near x_0.

Definition. Let H be a real Hilbert space. A connected *smooth manifold* modeled on H is a connected Hausdorff space $\mathcal{M}$ together with a collection $\mathcal{A} = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$, where each U_α is an open subset of $\mathcal{M}$ and each ϕ_α is a homeomorphism from U_α onto an open subset $\phi_\alpha(U_\alpha) \subseteq H$ such that

1. $\bigcup\{U_\alpha : \alpha \in A\} = \mathcal{M}$.
2. $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is C^∞, for all $\alpha, \beta \in A$.

We say that $\mathcal{A} = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$ is the *atlas* defining the smooth structure on $\mathcal{M}$, and each (U_α, ϕ_α) is one of the *charts* in the atlas.

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be smooth manifolds modeled on Hilbert spaces H_1 and H_2 respectively. Suppose that $\mathcal{M}_1$ and $\mathcal{M}_2$ have atlases $\mathcal{A}_1 = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$ and $\mathcal{A}_2 = \{(V_\beta, \psi_\beta) : \beta \in B\}$. A continuous map $F : \mathcal{M}_1 \to \mathcal{M}_2$ is said to be *smooth* if $\psi_\beta \circ F \circ \phi_\alpha^{-1}$ is C^∞, where defined, for $\alpha \in A$ and $\beta \in B$. It follows from the chain rule that the composition of smooth maps is smooth. As in the case of finite-dimensional manifolds, a *diffeomorphism* between Hilbert manifolds is a smooth map between manifolds with smooth inverse.

In the case of $P^\infty\mathbb{C}$ we can construct a countable atlas of smooth charts

$$\{(U_1,\phi_1),(U_2,\phi_2),\ldots,(U_i,\phi_i),\ldots\}$$

by setting

$$U_i = \{\{[z_1,z_2,\ldots,z_i,\ldots] \in P^\infty\mathbb{C} : z_i \neq 0\}$$

and defining

$$\phi_i : U_i \to C^\infty \quad \text{by} \quad \phi_i([z_1,z_2,\ldots,z_i,\ldots]) = \left(\frac{z_1}{z_i},\frac{z_2}{z_i},\ldots,\frac{z_{i-1}}{z_i},\frac{z_{i+1}}{z_i},\ldots\right).$$

It is quickly verified that $\phi_j \circ \phi_i^{-1}$ is smooth where defined, just as in the case of finite-dimensional projective spaces.

We can regard $P^\infty\mathbb{C}$ as the space of one-dimensional subspaces $V \subset H$. There is a *universal bundle* over $P^\infty\mathbb{C}$ whose total space is

$$E_\infty = \{(v,z) \in P^\infty\mathbb{C} \times H : v \in V\}.$$

The trivialization $\psi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{C}$ over U_i is defined by

$$\psi_i(V,z_1,z_2,\ldots,z_i,\ldots]) = (v,z_i),$$

and the corresponding transition functions are just the maps

$$g_{ij} : U_i \cap U_j \to GL(1,\mathbb{C}) \quad \text{defined by} \quad g_{ij} = \frac{z_i}{z_j}.$$

We have just described a special case of a more general construcction. Just like infinite-dimensional projective space, the *infinite Grassmannian* $G_m(C^\infty)$, which is the space of m-dimensional subspaces of infinite-dimensional Hilbert space C^∞. One can show that this is also a smooth Hilbert manifold and that

$$E_\infty = \{(V,v) \in G_m(C^\infty) \times C^\infty : v \in V\}$$

is the total space of a smooth vector bundle over $G_m(C^\infty)$, called the *universal bundle*.

If M is a smooth finite-dimensional manifold, we let $[M, G_m(C^\infty)]$ denote the space of smooth homotopy classes of maps from M to $G_m(C^\infty)]$. The following theorem then gives a homotopy-theoretic classification of complex vector bundles:

Universal Bundle Theorem I. *If M is a finite-dimensional smooth manifold, there is a bijection*

$$\Gamma : [M, G_m(C^\infty)] \to \mathcal{V}^m_{\mathbb{C}}(M),$$

where $\mathcal{V}^m_{\mathbb{C}}(M)$ is the set of isomorphism classes of smooth complex vector bundles of rank m over M.

Sketch of proof: First note that Γ is well-defined by Proposition 2 of §4.9.

To see that Γ is surjective, we suppose that the base space M has dimension n. We claim that M can be covered by $n+1$ open sets (not usually connected) which have the property that each component of each open set is contractible.

To construct such a cover we triangulate M and take the first barycentric subdivision. The star of each vertex in the barycentric subdivision is a contractible set. We let U_k be the union of the stars of vertices in the barycentric subdivision which correspond to k-simplices in the original triangulation. Then $(U_0, U_1, \ldots, U_n)$ is an open cover with the desired properties.

By the Corollary at the end of S 4.9, we can use this cover as a trivializing cover for the bundle E. Let

$$\psi_k : \pi^{-1}(U_k) \to U_k \times \mathbb{C}^m$$

be a trivialization over U_k and compose it with projection on the second factor to obtain maps

$$\eta_k = \pi_2 \circ \psi_k : \pi^{-1}(U_k) \to \mathbb{C}^m.$$

Finally, let $(\zeta_0, \zeta_1, \ldots, \zeta_n)$ be a partition of unity subordinate to the open cover $(U_0, U_1, \ldots, U_n)$, and define $\tilde{F} : E \to \mathbb{C}^{(n+1)m} \subset \mathbb{C}^\infty$ by

$$\tilde{F}(e) = (\zeta_0(\pi(e))\eta_0(e), \zeta_1(\pi(e))\eta_1(e), \ldots \zeta_n(\pi(e))\eta_n(e)).$$

Since $\tilde{F}$ is injective on each fiber, it induces a smooth map $F : M \to G_m(C^\infty)$ such that $F(p) = \tilde{F}(E_p)$, where E_p is the fiber of E over $p \in M$. Then $F^* E_\infty = e$ and surjectivity is established.

For injectivity, we need two subspaces of complex Hilbert space C^∞:

$$C_e^\infty = \{(z_1, z_2, \ldots, z_i, \ldots) \in C^\infty : \ z_{2i-1} = 0 \text{ for } i \in \mathbb{N}\ \}$$
$$\text{and} \qquad C_o^\infty = \{(z_1, z_2, \ldots, z_i, \ldots) \in C^\infty : \ z_{2i} = 0 \text{ for } i \in \mathbb{N}\ \}.$$

We define linear maps

$$\tilde{T}_e : C^\infty \to C_e^\infty \quad \text{by} \quad T_e(z_1, z_2, z_3, z_4, \ldots) = (0, z_1, 0, z_2, 0, z_3, 0, z_4, \ldots)$$

and

$$\tilde{T}_o C^\infty \to C_o^\infty \quad \text{by} \quad T_e(z_1, z_2, z_3, z_4, \ldots) = (z_1, 0, z_2, 0, z_3, 0, z_4, 0, \ldots).$$

These induce maps

$$T_e : G_m(C^\infty) \to G_m(C_e^\infty) \subseteq G_m(C^\infty) \quad \text{and} \quad T_o : G_m(H) \to G_m(C_o^\infty) \subseteq G_m(C^\infty),$$

and we claim that these maps are homotopic to the identity. Indeed, we can define

$$\tilde{H}_e : C^\infty \times [0,1] \to C^\infty \quad \text{by} \quad \tilde{H}_e(z,t) = tz + (1-t)\tilde{H}_e(z).$$

If $(e_1, \ldots e_m)$ are linearly independent elements of H, then so are

$$(\tilde{H}_e(e_1, t), \ldots, \tilde{H}_e(e_m, t))$$

for every choice of $t \in [0,1]$. Hence $\tilde{H}_e$ induces a smooth homotopy

$$G_e : G_m(C^\infty) \times [0,1] \longrightarrow G_m(C^\infty)$$

from T_e to the identity. A similar construction shows that T_0 is homotopic to the identity.

To show that Γ is injective, we need to show that if $E \in \mathcal{V}_{\mathbb{C}}^m(M)$ and $F, G : M \to G_m(C^\infty)$ are two smooth maps such that $E = F^*E_\infty = G^*E_\infty$, then F and G are homotopic. To do this it suffices to show that $T_e \circ F$ and $T_o \circ G$ are homotopic. But the maps $T_e \circ F$ and $T_o \circ F$ are covered by maps

$$\tilde{T}_e \circ \tilde{F} : E \to C_e^\infty, \qquad \tilde{T}_o \circ \tilde{G} : E \to C_o^\infty.$$

Thus we can define

$$\tilde{H} : E \times [0,1] \to C^\infty \quad \text{by} \quad \tilde{H}(e,t) = t\tilde{T}_e \circ \tilde{F}(e,t) + (1-t)\tilde{T}_o \circ \tilde{G}(e,t).$$

Then $\tilde{H}$ induces a map

$$H : M \times [0,1] \to G_m(C^\infty) \quad \text{b} \quad H(p,t) = \tilde{H}(E_p,t).$$

This is the desired homotopy from $T_e \circ F$ to $T_o \circ G$, and injectivity is established.

Remark. One can calculate the real cohomology ring of $G_m(C^\infty)$, and we find that it is a polynomial algebra on generators

$$c_1 \in H^2(G_m(C^\infty)), \quad c_2 \in H^4(G_m(C^\infty)), \quad \cdots, \quad c_m \in H^{2m}(G_m(C^\infty)),$$

these generators being exactly the Chern classes of the universal bundle. Given a complex vector bundle E of rank m over a smooth manifold M, the Chern classes of E are just the pullbacks of the Chern classes of the universal bundle E_∞ via a map $F : M \to G_m(C^\infty)$ such that $\Gamma^*(E_\infty) = E$. Thus we could have defined the Chern classes in terms of the cohomology of the infinite Grassmannian.

By exactly the same procedure, one can prove real analogs of the Universal Bundle Theorem. We consider the *infinite Grassmannians*

$$G_m(R^\infty) = \{ \text{ m-dimensional subspaces of infinite-dimensional Hilbert space } R^\infty \}$$

and

$$G_m^+(R^\infty) = \{ \text{ m-dimensional subspaces of infinite-dimensional Hilbert space } R^\infty \}$$

We let

$$E_\infty = \{(V,v) \in G_m(R^\infty) \times C^\infty : v \in V\},$$
$$\text{and} \quad E_\infty^+ = \{(V,v) \in G_m^+(R^\infty) \times C^\infty : v \in V\},$$

the total space of smooth vector universal vector bundles over $G_m(R^\infty)$ and $G_m^+(R^\infty)$.

Universal Bundle Theorem II. *If M is a finite-dimensional smooth manifold, there is a bijection*

$$\Gamma : [M, G_m(R^\infty)] \to \mathcal{V}^m(M), \quad \Gamma(F) = F^* E_\infty$$

where $\mathcal{V}^m(M)$ is the set of isomorphism classes of smooth real vector bundles of rank m over M.

Universal Bundle Theorem III. *If M is a finite-dimensional smooth manifold, there is a bijection*

$$\Gamma : [M, G_m^+(C^\infty)] \to \mathcal{V}_+^m(M), \quad \Gamma(F) = F^* E_\infty^+$$

where $\mathcal{V}_+^m$ is the set of isomorphism classes of smooth complex vector bundles of rank m over M.

5.6 The Clifford algebra

Many of the most subtle modern developments in contemporary Riemannian geometry make use of the Clifford algebra, and the theory of spinors which it renders accessible. We provide an introduction to that theory here; the reader can find much more detail in [19], and a quick survey in §1.11 of [16].

To a finite-dimensional real vector space V with inner product $\langle \cdot, \cdot \rangle$ we can associate its *Clifford algebra*. This is an associative algebra $\mathrm{Cl}(V)$ with unit 1 together with a monomorphism $\theta : V \to \mathrm{Cl}(V)$ such that

$$\theta(v) \cdot \theta(v) = -\langle v, v \rangle 1,$$

which satisfies the universal property: Given a linear map $h : V \to A$, where A is an associative algebra with unit, such that $h(v) \cdot h(v) = -\langle v, v \rangle 1$ for $v \in V$, there is a *unique* algebra homomorphism $\tilde{h} : \mathrm{Cl}(V) \to A$ such that $\tilde{h} \circ \theta = h$.

Using the universal property, one can show that such an associative algebra with unit is unique up to isomorphism, if it exists. For existence, one takes $\mathrm{Cl}(V)$ to be the quotient of the tensor algebra $\otimes^* V/I$, where I is the two-sided ideal generated by $v \otimes v + \langle v, v \rangle 1$, for $v \in V$. One then defines $\theta : V \to \mathrm{Cl}(V)$ to be the composition of inclusion $V \subseteq \otimes^* V$ with the projection to the quotient $\otimes^* V/I$. One can then show that $\theta : V \to \mathrm{Cl}(V)$ is a monomorphism.

More informally, we could define the Clifford algebra to be the algebra generated by the vector space V, with its multiplication being subject to the relations

$$v \cdot w + w \cdot v = -2\langle v, w \rangle, \quad \text{for } v, w \in V.$$

If $(e_1, \ldots, e_n)$ is an orthonormal basis for V_n, the unique inner product space of dimension n, then the Clifford algebra $\mathrm{Cl}(V_n)$ is generated by the elements e_i subject to the relations

$$e_i \cdot e_j + e_j \cdot e_i = -2\delta_{ij}.$$

We can think of the product on the Clifford algebra as a refinement of the wedge product in the exterior product $\Lambda^* V$. To see how this works, we define, for $v \in V$,

$$\varepsilon_v : \Lambda^* V \to \Lambda^* V \quad \text{by} \quad \varepsilon_v(\omega) = v \wedge \omega$$

and let $\iota_v : \Lambda^* V \to \Lambda^* V$ be the skew-derivation such that

$$\iota_v(w) = \langle v, w \rangle, \quad \text{for } w \in V.$$

We then constsruct a linear map $h : V \to \mathrm{End}(\Lambda^* V)$ by

$$h(v)(\omega) = \varepsilon_v(\omega) - \iota_v(\omega).$$

For arbitrary $v \in V$,

$$h(v) \circ h(v) = -\langle v, v \rangle,$$

so by the universal property of Clifford algebras, h induces an algebra homomorphism

$$\tilde{h} : \mathrm{Cl}(V) \to \mathrm{End}(\Lambda^* V).$$

This enables us to define a map

$$\sigma : \mathrm{Cl}(V) \to \Lambda^* V \quad \text{by} \quad \sigma(\phi) = \tilde{h}(\phi)1.$$

If $(e_1, \ldots, e_n)$ is an orthonormal basis for V, then one immediately verifies that

$$\sigma(e_{i_1} \cdots e_{i_k}) = e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

From this one easily sees that σ is a vector space isomorphism from $\mathrm{Cl}(V)$ to $\Lambda^* V$. Thus we can regard $\Lambda^* V$ and $\mathrm{Cl}(V)$ as the same vector space with two different products, a wedge product or a Clifford product. Note that if $\phi \in \Lambda^k V$ and $\psi \in \Lambda^l V$, then

$$\phi \wedge \psi = (\text{component of } \phi \cdot \psi \text{ of degree } k + l).$$

We let $\mathrm{Cl}^k(V)$ denote the preimage of $\Lambda^k V$ under the vector space isomorphism σ.

The universal property shows that Clifford multiplication preserves a $\mathbb{Z}_2$-grading. Indeed, we can define a linear map

$$\alpha : V \to V \quad \text{by} \quad \alpha(v) = -v.$$

Note that $\alpha^2 = \mathrm{id}$. Composition with θ yields a linear map $\alpha : V \to \mathrm{Cl}(V)$ such that

$$\alpha(v) \cdot \alpha(v) = -\langle v, v \rangle 1, \quad \text{for } v \in V,$$

and hence α induces an algebra homomorphism

$$\alpha : \mathrm{Cl}(V) \to \mathrm{Cl}(V) \quad \text{such that} \quad \alpha^2 = \mathrm{id}.$$

We let

$$\mathrm{Cl}_0(V) = \{\phi \in \mathrm{Cl}(V) : \alpha(\phi) = \phi\}, \qquad \mathrm{Cl}_1(V) = \{\phi \in \mathrm{Cl}(V) : \alpha(\phi) = -\phi\}.$$

Then $\mathrm{Cl}_i(V)\mathrm{Cl}_j(V) \subset \mathrm{Cl}_{i+j}(V)$ where the addition of i and j is modulo two. We can also write

$$\mathrm{Cl}_0(V) = \sum \mathrm{Cl}^{2k}V, \qquad \mathrm{Cl}_1(V) = \sum \mathrm{Cl}^{2k+1}V.$$

To better understand the structure of Clifford algebras, it is convenient to study the eigenvalues and eigenvectors of the linear operator

$$L : \mathrm{Cl}(V) \to \mathrm{Cl}(V) \quad \text{defined by} \quad L(\phi) = -\sum_{i=1}^{n} e_i \cdot \phi \cdot e_i, \tag{5.6}$$

where $(e_1, \ldots, e_n)$ is an orthonormal basis for V. We claim that

$$\phi \in \mathrm{Cl}^k V \quad \Rightarrow \quad L(\phi) = (-1)^k(n-2k)\phi.$$

Indeed, to prove the claim, one can assume that $\phi = e_1 \cdots e_k$. Then a direct calculation shows that

$$\begin{aligned} L(\phi) = -\sum_{i=1}^{k} e_i \cdot e_1 \cdots e_k \cdot e_i - \sum_{i=k+1}^{n} e_i \cdot e_1 \cdots e_k \cdot e_i \\ = (-1)^{k+1} k e_1 \cdots e_k + (-1)^k (n-k) e_1 \cdots e_k = (-1)^k(n-2k)\phi, \end{aligned}$$

From this one can determine the eigenvalues λ for L and the eigenspace W_λ for a given choice of λ. If $n = 2m$ is even, then the eigenvalues are

$$2m, \quad -2(m-1), \quad 2(m-2), \quad \ldots, \quad -2m$$

and the corresponding eigenspaces are

$$W_{2m} = \mathrm{Cl}^0 V, \quad W_{-2(m-1)} = \mathrm{Cl}^1 V, \quad \ldots, W_{-2m} = \mathrm{Cl}^{2m} V.$$

On the other hand when $n = 2m+1$, the eigenvalues are

$$2m+1, \quad -(2m-1), \quad , \ldots$$

and the corresponding eigenspaces are

$$W_{2m+1} = \mathrm{Cl}^0 V \oplus \mathrm{Cl}^{2m+1} V, \quad W_{-(2m-1)} = \mathrm{Cl}^1 V \oplus \mathrm{Cl}^{2m} V, \ldots.$$

An immediate applications is:

Lemma 1. *Suppose that $(V, \langle \cdot, \cdot \rangle)$ is an inner product space of dimension n. If n is even, the center of the Clifford algebra $Cl(V_n)$ consists of the scalar multiples of 1, while if n is odd the center is generated by the scalar multiples of 1 and $e_1 \cdots e_n$.*

Indeed, if σ is in the center of $\mathrm{Cl}(V)$, $L(\sigma) = n\sigma$. Thus if $n = 2m$ is even, σ must lie in $W_{2m} = \mathrm{Cl}^0(V)$. On the other hand, if $n = 2m+1$ is odd, σ must be an element of $W_{2m+1} = \mathrm{Cl}^0(V) \oplus \mathrm{Cl}^n(V)$, which is spanned by 1 and $e_1 \cdots e_n$.

Lemma 2. *If V is even-dimensional, then the Clifford algebra of V has no nontrivial ideals.*

Indeed, suppose that I is a nonzero ideal in $\mathrm{Cl}(V)$. If $\phi \in I$, then so is

$$\Pi_{k \neq p}[L - (-1)^k(n-2k)I](\phi) = c\phi_p,$$

where ϕ_p is the component of ϕ in $\mathrm{Cl}^p V$, c being a nonzero constant. Thus if $\phi \in I$, so is its homogeneous component ϕ_p of degree p. We can suppose that

$$\phi_p = ae_1 \cdots e_p + (\text{linear combination of other terms in } \mathrm{Cl}^p V),$$

where $a \in \mathbb{R} - \{0\}$. Then

$$\frac{1}{a}\phi_p \cdot e_{p+1} \cdots e_n \in I,$$

so its homogeneous component $e_1 \cdots e_n$ of degree n must be in I. Hence $(e_1 \cdots e_n)^2 = \pm 1 \in I$, and $I = \mathrm{Cl}(V)$, finishing the proof.

Remark. It follows immediately that the complexified Clifford algebra $\mathrm{Cl}(V) \otimes \mathbb{C}$ has no nontrivial ideals as a complex algebra, when V is even-dimensional.

Example 1. The Clifford algebra is easily constructed for a two-dimensional inner product space $(V, \langle \cdot, \cdot \rangle)$ with orthonormal basis (e_1, e_2). In this case, the Clifford algebra $\mathrm{Cl}(V)$ is generated as a vector space by

$$1, \quad e_2, \quad e_2, \quad e_1 \cdot e_2.$$

We could adopt the notation

$$e_1 = \mathbf{i}, \quad e_2 = \mathbf{j}, \quad e_1 \cdot e_2 = \mathbf{k}.$$

Then the rules of Clifford multiplication show that

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j},$$

so in this case $\mathrm{Cl}(V)$ is isomorphic to the space of quaternions.

Let $M_n(\mathbb{C})$ denote the space of $n \times n$ complex matrices, an associative algebra with identity over $\mathbb{C}$ with matrix multiplication as the product. We can then define a linear map $\theta : V \to M_2(\mathbb{C})$ by

$$\theta(e_1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \theta(e_2) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Alternatively, if $z = xe_1 + ye_2 = x + iy \in V$, we set

$$\theta(z) = \begin{pmatrix} 0 & z \\ -\bar{z} & 0 \end{pmatrix}.$$

Then

$$\theta(z) \cdot \theta(z) = \begin{pmatrix} -|z|^2 & 0 \\ 0 & -|z|^2 \end{pmatrix} = -\langle z, z \rangle I,$$

so θ induces an injective algebra homomorphism $\theta : \mathrm{Cl}(V) \to M_2(\mathbb{C})$. If we complexify the Clifford algebra, we obtain an isomorphism

$$\theta : \mathrm{Cl}(V) \otimes \mathbb{C} \to M_2(\mathbb{C}).$$

Note that

$$\mathrm{Cl}_0(V) \otimes \mathbb{C} \text{ consists of complex matrices of the form } \begin{pmatrix} z & 0 \\ 0 & w \end{pmatrix},$$

$$\mathrm{Cl}_1(V) \otimes \mathbb{C} \text{ consists of complex matrices of the form } \begin{pmatrix} 0 & w \\ z & 0 \end{pmatrix}.$$

We let W_0 and W_1 be the complex vector subspaces of $\mathbb{C}^2$ spanned by the vectors

$$\epsilon_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \epsilon_2 = \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

respectively. Then $\mathrm{Cl}_i(V)W_j \subset W_{i+j}$, where the addition is modulo two once again.

Example 2. We can also give a quite explicit expression for the Clifford algebra in the case of a four-dimensional inner product space $(V, \langle \cdot, \cdot \rangle)$ with orthonormal basis (e_1, e_2, e_3, e_4). In this case, we can let V be the space of 2×2 complex matrices Q such that $Q = \lambda A$ for some $\lambda \in \mathbb{R}$ and some $A \in SU(2)$. Thus if $Q \in V$, we can write

$$\begin{aligned} Q = &\begin{pmatrix} t + iz & x + iy \\ -x + iy & t - iz \end{pmatrix} \\ &= x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + z \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \end{aligned}$$

where x, y, z and t are real numbers. Moreover, we can assume that the inner product on V is defined so that

$$\det Q = t^2 + x^2 + y^2 + z^2 = \langle Q, Q \rangle.$$

We then define

$$\theta : V \to M_4(\mathbb{C}) \quad \text{by} \quad \theta(Q) = \begin{pmatrix} 0 & Q \\ -\bar{Q}^T & 0 \end{pmatrix},$$

and immediately verify that

$$\theta(Q) \cdot \theta(Q) = \begin{pmatrix} -Q\bar{Q}^T & 0 \\ 0 & -\bar{Q}^T Q \end{pmatrix} = \langle Q, Q \rangle I,$$

since $Q = \lambda A$, where $A \in SU(2)$ and hence $\bar{A}^T A = I$. It follows that θ induces an algebra homomorphism $\theta : \mathrm{Cl}(V_4) \to M_4(\mathbb{C})$ which must have trivial kernel by Lemma 2. The complexified map must also have trivial kernel and since image and range have the same dimension, we obtain an algebra isomorphism

$$\theta : \mathrm{Cl}(V_4) \otimes \mathbb{C} \to M_4(\mathbb{C}).$$

Note that $\theta(e_1)$, $\theta(e_2)$, $\theta(e_3)$ and $\theta(e_4)$ are all skew-Hermitian matrices lying in the off-diagonal blocks. We let W_0 and W_1 be the complex vector subspaces of $\mathbb{C}^4$ spanned by the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

respectively, and once again $\mathrm{Cl}_i(V) W_j \subset W_{i+j}$.

A similar structure holds for Clifford algebras of even-dimensional inner product spaces:

Theorem. *If V is even-dimensional, say $\dim V = 2m$, then*

$$Cl_{\mathbb{C}}(V) = Cl(V) \otimes \mathbb{C} \cong End_{\mathbb{C}}(W),$$

where W is a complex vector space of dimension 2^m.

We call W the *irreducible Clifford module* for the Clifford algebra $\mathrm{Cl}(V) \otimes \mathbb{C}$. It can be shown that the irreducible Clifford module W is unique up to isomorphism, and that any Clifford module is isomorphic to

$$\overbrace{W \oplus \cdots \oplus W}^{k} = W \otimes \mathbb{C}^k, \quad \text{for some } k \in \mathbb{Z}.$$

When $\dim V = 2m$, we can define a complex volume element

$$\Theta_{\mathbb{C}} = i^m e_1 \cdot e_2 \cdots e_{2m} \quad \text{such that} \quad \Theta_{\mathbb{C}}^2 = 1. \tag{5.7}$$

Hence W divides into a direct sum decomposition, $W = W_0 \oplus W_1$, where

$$W_0 = \{w \in W : \Theta_{\mathbb{C}}(w) = w\}, \qquad W_1 = \{w \in W : \Theta_{\mathbb{C}}(w) = -w\},$$

and it follows from the relation

$$(e_1 \cdots e_k)\omega = (-1)^k \omega (e_1 \cdots e_k)$$

that

$$\mathrm{Cl}_i(V) W_j \subset W_{i+j}, \tag{5.8}$$

where the sum $i + j$ is taken modulo two. Thus we can say that $W = W_0 \oplus W_1$ is a $\mathbb{Z}_2$-graded module over the Clifford algebra.

One can prove the above theorem is by induction on m using properties of tensor products of $\mathbb{Z}_2$-graded complex algebras. An algebra A is said to be $\mathbb{Z}_2$-graded if it is written as a direct sum $A = A_0 \oplus A_1$ in such a way that $A_i \cdot A_j \subseteq A_{i+j}$ where the sum $i + j$ is taken modulo two. For example, $\mathrm{Cl}_{\mathbb{C}}(V)$ is $\mathbb{Z}_2$-graded, and so is $\mathrm{End}_{\mathbb{C}}(W)$, with

$$\mathrm{End}_{\mathbb{C}}(W)_0 = \mathrm{Hom}_{\mathbb{C}}(W_0, W_0) \oplus \mathrm{Hom}_{\mathbb{C}}(W_1, W_1),$$
$$\mathrm{End}_{\mathbb{C}}(W)_1 = \mathrm{Hom}_{\mathbb{C}}(W_1, W_0) \oplus \mathrm{Hom}_{\mathbb{C}}(W_0, W_1).$$

If $A = A_0 \oplus A_1$ and $B = B_0 \oplus B_1$ are graded algebras, we can make the tensor product $A \otimes B$ into a $\mathbb{Z}_2$-graded algebra, called the *graded tensor product* and denoted by $A\hat{\otimes}B$, by setting

$$(A\hat{\otimes}B)_0 = A_0 \otimes B_0 \oplus A_1 \otimes B_1, \quad (A\hat{\otimes}B)_1 = A_0 \otimes B_1 \oplus A_1 \otimes B_0,$$

and defining the product by

$$(a \otimes b) \cdot (a' \otimes b') = (-1)^{(\deg b)(\deg a')}(aa') \otimes (bb').$$

Following [19], page 11, we can then prove:

Lemma 3. If $(V_1, \langle \cdot, \cdot \rangle)$ and $(V_2', \langle \cdot, \cdot \rangle)$ are even-dimensional inner product spaces, then

$$\mathrm{Cl}_{\mathbb{C}}(V_1 \oplus V_2) \cong \mathrm{Cl}_{\mathbb{C}}(V_1)\hat{\otimes}\mathrm{Cl}_{\mathbb{C}}(V_2).$$

Indeed, if

$$\theta_1 : V_1 \to \mathrm{Cl}_{\mathbb{C}}(V_1) \quad \text{and} \quad \theta_2 : V_2 \to \mathrm{Cl}_{\mathbb{C}}(V_2)$$

are the usual inclusions into the Clifford algebras, we simply define

$$\theta : V_1 \oplus V_2 \to \mathrm{Cl}_{\mathbb{C}}(V_1)\hat{\otimes}\mathrm{Cl}_{\mathbb{C}}(V_2) \quad \text{by} \quad \theta(v_1 + v_2) = \theta_1(v_1) \otimes 1 + 1 \otimes \theta_2(v_2),$$

and check that

$$\begin{aligned}\theta(v_1 + v_2) \cdot \theta(v_1 + v_2) &= (\theta_1(v_1) \otimes 1 + 1 \otimes \theta_2(v_2)) \cdot (\theta_1(v_1) \otimes 1 + 1 \otimes \theta_2(v_2)) \\ &= (\theta_1(v_1) \cdot \theta_1(v_1)) \otimes 1 + 1 \otimes (\theta_2(v_2) \cdot \theta_2(v_2)) = -\langle v_1, v_1 \rangle - \langle v_2, v_2 \rangle.\end{aligned}$$

Then θ induces a map from $\mathrm{Cl}_{\mathbb{C}}(V_1 \oplus V_2)$ into $\mathrm{Cl}_{\mathbb{C}}(V_1)\hat{\otimes}\mathrm{Cl}_{\mathbb{C}}(V_2)$ with trivial kernel, and since both range and domain have the same dimension, it must be an isomorphism.

The Theorem now follows from Lemma 3 and the relation

$$\mathrm{End}_{\mathbb{C}}(W \otimes W') \cong \mathrm{End}_{\mathbb{C}}(W) \otimes \mathrm{End}_{\mathbb{C}}(W'),$$

by an induction which starts with Examples 1 and 2.

What about Clifford algebras of inner product spaces of odd dimensions? These can be reduced to the even-dimensional case, yielding:

Corollary. *If V is odd-dimensional, say $\dim V = 2m+1$, then*

$$Cl_{\mathbb{C}}(V) = Cl(V) \otimes \mathbb{C} \cong End_{\mathbb{C}}(W_0) \oplus End_{\mathbb{C}}(W_1),$$

where W_0 and W_1 are complex vector spaces of dimension 2^m.

To prove the Corollary, suppose that V has orthonormal basis $(e_1, \ldots, e_{2m+1})$ and consider V as a hyperplane in an inner product space V' with orthonormal basis $(e_1, \ldots, e_{2m+2})$. We can then construct an isomorphism

$$h : \mathrm{Cl}_{\mathbb{C}}(V) \to \mathrm{Cl}_0(V') \otimes \mathbb{C} \quad \text{by setting} \quad h(\phi) = \phi \cdot e_{2m+2}.$$

5.7 The spin group

In the preceding section, we saw that if $(V, \langle \cdot, \cdot \rangle)$ is an inner product space of even dimension, then the complexification of its Clifford algebra $\mathrm{Cl}(V) \otimes \mathbb{C}$ is of the form $\mathrm{End}(W)$, for some complex vector space W, while if it is of odd dimension, it is a subalgebra of $\mathrm{End}(W)$. Within $\mathrm{End}(W)$ is the Lie group

$$GL(W) = \{T \in \mathrm{End}(W) : \ T \text{ is an isomorphism } \}.$$

According to the Lie group—Lie algebra correspondence, there is a one-to-one correspondence between Lie subalgebras of $\mathrm{End}(W)$ and Lie subgroups of $GL(W)$. This is explained in most references on Lie groups; see in particular, Theorem 8.7, page 158 of [5].

One can check directly that the usual bracket operation

$$[e_i \cdot e_j, e_k \cdot e_l] = e_i \cdot e_j \cdot e_k \cdot e_l - e_k \cdot e_l \cdot e_i \cdot e_j$$

makes $\mathrm{Cl}^2(V)$ into a Lie subalgebra of $\mathrm{End}(W)$. The corresponding Lie group is the spin group $\mathrm{Spin}(n)$, and we will construct it in the following paragraphs.

First we describe the transpose operation on the Clifford algebra. If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space of dimension n with Clifford algebra $\mathrm{Cl}(V)$, the transformation

$$v_1 \otimes \cdots \otimes v_k \mapsto v_k \otimes \cdots \otimes v_1$$

determines an antiautomorphism of $\otimes^* V$ which descends an antiautomorphism of the Clifford algebra $\mathrm{Cl}(V)$,

$$v_1 \cdots v_k \mapsto (v_1 \cdots v_k)^T = v_k \cdots v_1.$$

Note that if v_1, v_2, ..., v_k are unit-length elements of V, then

$$(v_1 \cdots v_k)(v_1 \cdots v_k)^T = \begin{cases} -1, & \text{if } k \text{ is odd,} \\ 1, & \text{if } k \text{ is even.} \end{cases}$$

We conclude that the elements of

$$\mathrm{Pin}(n) = \{v_1 \cdots v_k : v_1, \ldots, v_k \text{ are unit length elements of } V \ \}$$

have multiplicative inverses and hence $\mathrm{Pin}(n)$ is a subgroup of the group of units in the Clifford algebra, as is the *spin group*

$$\mathrm{Spin}(n) = \mathrm{Pin}(n) \cap \mathrm{Cl}_0(V). \tag{5.9}$$

Note that if $\sigma \in \mathrm{Pin}(n)$ and $v \in V$, then

$$\rho(\sigma)(v) = \sigma v \sigma^T \in V \quad \text{and} \quad |\rho(\sigma)(v)| = |v|.$$

Thus we have a group homomorphism

$$\rho : \mathrm{Pin}(n) \to O(n) = \{T \in GL(V) : |T(v)| = |v|, \text{ for all } v \in V\ \}.$$

Following Theorem 1.11.1 of [16], we present:

Proposition. *The homomorphism ρ restricts to a surjective homomorphism*

$$\rho : Spin(n) \to SO(n)$$

which has kernel ± 1.

To prove this, we note that if v is a unit-length element of V, then $\rho(v) : V \to V$ is a reflection in the hyperplane perpendicular to v. It follows that

$$\rho(v_1 \cdot v_2 \cdots v_k) = \rho(v_1) \circ \rho(v_2) \circ \cdots \circ \rho(v_k)$$

is the composition of the reflection in k hyperplanes. Any element of the orthogonal group can be represented as a composition of reflections in hyperplanes and the element lies in the special orthogonal group if and only if the number of reflections is even. Thus we see that ρ maps $\mathrm{Spin}(n)$ onto $SO(n)$.

Moreover, since $\sigma^{-1} = \sigma^T$ for $\sigma \in \mathrm{Spin}(n)$, we see that an element $\sigma \in \mathrm{Spin}(n)$ lies in the kernel of ρ if and only if $\sigma v = v\sigma$ for all $v \in V$, and hence if and only if

$$\sigma\phi = \phi\sigma \quad \text{for all } \phi \in \mathrm{Cl}(V),$$

or equivalently, σ lies in the center of $\mathrm{Cl}(V)$. Since no scalar multiple of $e_1 \cdots e_n$ lies in $\mathrm{Spin}(n)$ when n is odd, it follows from Lemma 1 from the previous section that the kernel of ρ is just $\{\pm 1\}$, as claimed.

We claim that $\mathrm{Spin}(n)$ is the Lie group which has Lie algebra $\mathrm{Cl}^2(V)$. To see this, we note that

$$\begin{aligned} e^{te_1 \cdot e_2} &= 1 + te_1 \cdot e_2 + \frac{t^2}{2!}(e_1 \cdot e_2)^2 + \frac{t^3}{3!}(e_1 \cdot e_2)^3 + \frac{t^4}{4!}(e_1 \cdot e_2)^4 + \cdots \\ &= 1 + te_1 \cdot e_2 - \frac{t^2}{2!} - \frac{t^3}{3!} e_1 \cdot e_2 + \frac{t^4}{4!} + \cdots \\ &= \cos t + \sin t e_1 \cdot e_2 = e_1 \cdot (-\cos t e_1 + \sin t e_2) \in \mathrm{Spin}(n). \end{aligned}$$

Thus $t \mapsto e^{te_1 \cdot e_2}$ is a one-parameter subgroup of $\mathrm{Spin}(n)$ and its tangent vector at $t = 0$ must be an element of the Lie algebra $\mathfrak{spin}(n)$ of $\mathrm{Spin}(n)$. Thus $\mathrm{Cl}^2(V)$

is a subspace of $\mathfrak{spin}(n)$ and since $\dim(\mathrm{Cl}^2(V)) = \dim(\mathfrak{spin}(n))$, we conclude that $\mathfrak{spin}(n) = \mathrm{Cl}^2(V)$.

We let $\mathrm{Spin}^c(n)$ denote the group of units in $\mathrm{Cl}(V) \otimes \mathbb{C}$ generated by the products $e_1 \cdot e_2$, where e_1 and e_2 are unit-length elements of V, and by the complex scalars λ of length one. As described more fully on page 71 of [16], we have an isomorphism

$$\mathrm{Spin}^c(n) = \frac{\mathrm{Spin}(n) \times S^1}{\mathbb{Z}_2},$$

where the $\mathbb{Z}_2$-action is described by $(\sigma, e^{i\theta}) \mapsto (-\sigma, -e^{i\theta})$. In addition to the homomorphism $\rho : \mathrm{Spin}^c(n) \to SO(n)$ we have a homomorphism

$$\pi : \mathrm{Spin}^c(n) \to S^1 \quad \text{defined by} \quad \pi(\sigma, e^{i\theta}) = e^{2i\theta}.$$

The key advantage to the groups $\mathrm{Spin}(n)$ and $\mathrm{Spin}^c(n)$ is that they have representations that are more basic than the representations of $SO(n)$ on Euclidean space V itself. These are the representations

$$\rho_W : \mathrm{Spin}(n) \to \mathrm{End}(W), \qquad \rho_W : \mathrm{Spin}^c(n) \to \mathrm{End}(W)$$

on the space W defined by the inclusion of $\mathrm{Cl}(V)$ into $\mathrm{End}(W)$. Note that this representation preserves the direct sum decomposition $W = W_0 \oplus W_1$.

Example 1. As we saw in the previous section, if $(V, \langle \cdot, \cdot \rangle)$ is two-dimensional, we can represent a typical element of V as a complex number $z \in \mathbb{C}$ and the map

$$\theta : V \to M_2(\mathbb{C}) \quad \text{defined by} \quad \theta(z) = \begin{pmatrix} 0 & z \\ -\bar{z} & 0 \end{pmatrix}$$

allows us to identify the the complexified Clifford algebra $\mathrm{Cl}(V) \otimes \mathbb{C}$ with $M_2(\mathbb{C})$. The spin group $\mathrm{Spin}(2)$ is simply the group of matrices of the form

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}, \quad \text{where} \quad e^{it} \in U(1),$$

while the group $\mathrm{Spin}^c(2)$ consists of matrices

$$\begin{pmatrix} e^{i(s+t)} & 0 \\ 0 & e^{i(s-t)} \end{pmatrix}, \quad \text{where} \quad e^{is}, e^{it} \in U(1).$$

Example 2. On the other hand, if $(V, \langle \cdot, \cdot \rangle)$ is four-dimensional, we can represent a typical element of V as a matrix

$$Q = \begin{pmatrix} t + iz & x + iy \\ -x + iy & t - iz \end{pmatrix}, \quad \text{where} \quad Q\overline{Q}^T = \lambda^2 I, \quad \lambda = \det Q.$$

Then the map

$$\theta : V \to M_4(\mathbb{C}) \quad \text{defined by} \quad \theta(Q) = \begin{pmatrix} 0 & Q \\ -\overline{Q}^T & 0 \end{pmatrix}$$

allows us to identify $\mathrm{Cl}(V) \otimes \mathbb{C}$ with $M_4(\mathbb{C})$. When this is done the spin group $\mathrm{Spin}(4)$ is simply the group of matrices of the form

$$\begin{pmatrix} A_+ & 0 \\ 0 & A_- \end{pmatrix}, \quad \text{where} \quad A_+, A_- \in SU(2)$$

and the map $\rho : \mathrm{Spin}(4) \to GL(V)$ is given by

$$\rho(A_=, A_-)Q = A_+ Q A_-^{-1}.$$

Similarly the group $\mathrm{Spin}^c(4)$ is the group of matrices

$$\begin{pmatrix} \lambda A_+ & 0 \\ 0 & \lambda A_- \end{pmatrix}, \quad \text{where} \quad A_+, A_- \in SU(2), \quad \lambda \in S^1.$$

Thus when $m = 1$ or 2, the group $\mathrm{Spin}(2m)$ preserves an Hermitian inner product on $W = W_0 \oplus W_1$ which respects the direct sum decomposition, such that W is isomorphic to its complex dual W^*. By induction based on the Theorem from §5.6 one can show that $\mathrm{Spin}(2m)$ preserves an Hermitian inner product on $W = W_0 \oplus W_1$ with similar properties for every positive integer m.

Similarly, in the odd-dimensional case, $\mathrm{Spin}(2m + 1)$ preserves Hermitian inner products on W_0 and W_1.

5.8 Spin structures and spin connections

Suppose now that $(M, \langle\cdot,\cdot\rangle)$ is an oriented smooth Riemannian manifold. The Riemannian metric allows us to regard the tangent bundle TM as an $SO(n)$-bundle. Thus we can choose a trivializing open cover $\{U_\alpha : \alpha \in A\}$ for TM such that the corresponding transition functions take their values in $SO(n)$:

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to SO(n).$$

In the preceding section, we constructed a Lie group $\mathrm{Spin}(n)$ together with a surjective Lie group homomorphism $\rho : \mathrm{Spin}(n) \to SO(n)$ which has kernel $\mathbb{Z}_2$.

Definition. A *spin structure* on $(M, \langle\cdot,\cdot\rangle)$ is defined by an open covering $\{U_\alpha : \alpha \in A\}$ of M and a collection of transition functions

$$\tilde{g}_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}(n)$$

such that the projections $g_{\alpha\beta} = \rho \circ \tilde{g}_{\alpha\beta}$ define the $SO(n)$-structure on TM and

$$\tilde{g}_{\alpha\beta}\tilde{g}_{\beta\gamma}\tilde{g}_{\gamma\alpha} = 1 \quad \text{on} \quad U_\alpha \cap U_\beta \cap U_\gamma,$$

for all α, β and γ in A.

We say that an oriented manifold M is a *spin manifold* if it possesses a Riemannian metric that admits a spin structure.

Although we cannot go into the details here, we mention that topologists have found that an oriented manifold M admits a spin structure if and only if "the second Stiefel-Whitney class" of its tangent bundle

$$w_2(TM) \in H^2(TM; \mathbb{Z}_2)$$

vanishes. (The definition and properties of Stiefel-Whitney classes are given in [26], and their application to spin structures is described in [19].)

If $(M, \langle \cdot, \cdot \rangle)$ is an oriented Riemannian manifold, then for each $p \in M$, T_pM is an inner product space and we can form the Clifford algebra $\mathrm{Cl}(T_pM)$. The Clifford algebras at the various points piece together to form bundle of Clifford algebras, which is a vector bundle $\mathrm{Cl}(TM)$ over M with additional structure.

We can think of the various tangent spaces as being modeled on a fixed inner product space $(V, \langle \cdot, \cdot \rangle)$. Thus, given a trivializing open cover $\{U_\alpha : \alpha \in A\}$, we can think of the tangent bundle TM as equivalence classes of triples $(\alpha, p, v) \in A \times M \times V$ with equivalence relation

$$(\alpha, p, v_\alpha) \sim (\beta, q, v_\beta) \quad \Leftrightarrow \quad p = q \in U_\alpha \cap U_\beta \quad \text{and} \quad v_\alpha = g_{\alpha\beta}(p) v_\beta.$$

As we saw in §5.6, the Clifford algebra $\mathrm{Cl}(V)$ of V and hence the Lie group $\mathrm{Spin}(n)$ lies within the space $\mathrm{End}(W)$ of complex endomorphisms of some fixed complex vector space W. Thus if M has a spin structure, we can construct a complex vector bundle with fiber isomorphic to W by the same construction. Thus we consider all triples $(\alpha, p, w) \in A \times M \times W$ with equivalence relation

$$(\alpha, p, w_\alpha) \sim (\beta, q, w_\beta) \quad \Leftrightarrow \quad p = q \in U_\alpha \cap U_\beta \quad \text{and} \quad w_\alpha = \tilde{g}_{\alpha\beta}(p) w_\beta,$$

where now $\tilde{g}_{\alpha\beta}$ takes values in $\mathrm{Spin}(n)$. This yields a complex vector bundle over M called the *spin bundle*, and we will denote it by S. Since by definition (5.9) the group $\mathrm{Spin}(n)$ lies in $\mathrm{Cl}_0(V)$, it follows that the direct sum decomposition $W = W_0 \oplus W_1$ yields a direct sum of spin bundles,

$$S = S_+ \oplus S_-,$$

called the spin bundles of *positive and negative chirality*. Sections of these bundles are called *spinor fields*.

Note that by the remarks at the end of the previous section, the spin bundles S_+ and S_- are endowed with Hermitian inner products.

If M is even-dimensional, the complexified bundle $\mathrm{Cl}(TM) \otimes \mathbb{C}$ of Clifford algebras over M can be regarded as isomorphic to the endomorphism bundle $\mathrm{End}(S)$. If M is odd-dimensional, $\mathrm{Cl}(TM) \otimes \mathbb{C}$ can be regarded as the direct sum $\mathrm{End}(S_+) \oplus \mathrm{End}(S_-)$.

Note that the Levi-Civita connection on TM induces a metric connection on $\Lambda^* TM$ and hence a connection on $\mathrm{Cl}(TM)$ and $\mathrm{Cl}(TM) \otimes \mathbb{C}$. Thus if M is even-dimensional, the Levi-Civita connection induces a Levi-Civita connection on $\mathrm{End}(S)$ which preserves the direct sum decomposition $S = S_+ \oplus S_-$, while if M is odd-dimensional it induces a connection on $\mathrm{End}(S_+) \oplus \mathrm{End}(S_-)$.

Theorem. *If $(M, \langle \cdot, \cdot \rangle)$ is an oriented Riemannian manifold with spin structure, then $S = S_+ \oplus S_-$ possesses a unique Spin(n)-connection which preserves the direct sum decomposition and induces the Levi-Civita connection on End$(S) =$ End$(S_+) \oplus$ End(S_-).*

We say that a connection D on S induces the Levi-Civita connection D_{LC} on End(S) if whenever $\omega \in \Gamma(\mathrm{End}(S))$ and $\sigma \in \Gamma(S)$, then

$$D(\omega\sigma) = (D_{LC}\omega)\sigma + \omega(D\sigma), \tag{5.10}$$

which is just the Leibniz rule.

To prove the Theorem, we let $\tilde{\psi}$ be a trivialization of the spin bundle S over $U \subseteq M$. This trivialization determines a trivialization of End(S) as well as trivializations of the subbundles of End(S) which correspond to subspaces of the Clifford algebra which are left fixed by the action of Spin(n). In particular, $\tilde{\psi}$ determines a trivialization ψ of TM over U.

Our strategy is to prove existence and uniqueness of the Spin(n)-connection over U. By local uniqueness, the locally defined Spin(n)-connections will then piece together to form a globally defined Spin(n)-connection over M which induces the Levi-Civita connection on the bundle of Clifford algebras.

Let $(e_1, \ldots, e_n)$ be the standard orthonormal basis of the model space V and let $(\epsilon_1, \ldots, \epsilon_{2^m})$ be the standard basis of W. We can then define an orthonormal moving frame $(E_1, \ldots, E_n)$ of $TM|U$ and corresponding orthonormal sections $(\varepsilon_1, \ldots, \varepsilon_{2^m})$ of $S|U$ such that

$$\psi \circ E_i(p) = (p, e_i), \qquad \tilde{\psi} \circ \varepsilon_\lambda(p) = (p, \epsilon_\lambda).$$

We can regard $(E_1, \ldots, E_n)$ as sections of End(S) and hence they act on sections of S. Moreover,

$$E_i \cdot \varepsilon_\lambda = \sum_{\mu=1}^{2^m} c^\mu_{i\lambda} \varepsilon_\mu,$$

where the $c^\mu_{i\lambda}$'s are constants.

Since the Lie algebra $\mathfrak{spin}(n)$ of Spin(n) is $\mathrm{Cl}^2(V)$, a Spin(n)-connection D over U must have an explicit expression in terms of our local trivilialization,

$$D = d + (\mathfrak{spin}(n)\text{-valued one-form}).$$

Since the Lie algebra $\mathfrak{spin}(n)$ is just $\mathrm{Cl}^2(V)$ which is generated by $e_i \cdot e_j$ with $i < j$, we must in fact have

$$D = d + \sum_{i,j=1}^{n} \phi_{ij} E_i \cdot E_j \cdot = d + \sum_{i,j=1}^{n} \phi_{ij} d_i \cdot d_j \cdot,$$

where the $\phi_{ij} = \phi_{ji}$ and the ϕ_{ij}'s are ordinary real valued one-forms, and the sections $(E_1, \ldots, E_n)$ are constantly equal to $(e_1, \ldots, e_n)$ in terms of the trivializations. Since the sections ε_λ and $E_k\varepsilon_\lambda$ have constant representatives in terms of the trivialization,

$$D(\varepsilon_\lambda) = \sum_{i,j=1}^{n} \phi_{ij} E_i \cdot E_j \cdot \varepsilon_\lambda, \quad D(E_k \varepsilon_\lambda) = \sum_{i,j=1}^{n} \phi_{ij} E_i \cdot E_j \cdot E_k \cdot \varepsilon_\lambda,$$

and hence it follows from (5.10) that

$$\sum_{i,j=1}^{n} \phi_{ij} E_i \cdot E_j \cdot E_k \cdot \varepsilon_\lambda = (D_{LC} E_k)\varepsilon_\lambda + E_k \cdot \sum_{i,j=1}^{n} \phi_{ij} E_i \cdot E_j \cdot \varepsilon_\lambda,$$

or equivalently,

$$(D_{LC} e_k) \cdot \varepsilon_\lambda = \sum_{i,j=1}^{n} (E_i \cdot E_j \cdot E_k - E_k \cdot E_i \cdot E_j)\varepsilon_\lambda.$$

The only terms that survive in the sum on the right are those in which $i \neq j$ and $k = 1$ or $k = j$. A short calculation shows that

$$D_{LC} e_k = -4 \sum_{i=1}^{n} \phi_{ik} E_i.$$

But the connection forms ω_{ij} of the Levi-Civita connection are defined by the equation

$$D_{LC} e_j = \sum_{i=1}^{n} \omega_{ij} E_i.$$

Thus we conclude that $\phi_{ij} = -(1/4)\omega_{ij}$, and the spin connection must be given in the local trivialization by the formula

$$D = d - \frac{1}{4} \sum_{i,j=1}^{n} \omega_{ij} E_i \cdot E_j \cdot, \tag{5.11}$$

where the differential forms ω_{ij} are the components of the Levi-Civita connection on TM. This proves uniqueness of the Spin(n)-connection.

For existence, we note that (5.11) does define a Spin(n)-connection over U and we can check that it induces the Levi-Civita connection, first on the constant sections E_i of End(S), then on products $E_i \cdot E_j$, and so forth. Thus we have established both existence and uniqueness of the Spin(n)-connection over any open neighborhood U for which we have associated trivializations of both the tangent bundle TM and the spin bundle S, exactly what we needed to prove.

We now ask: What is the curvature R of the Spin(n)-connection on the spin bundle S? To obtain this, we just set $R = D^2$, and obtain

$$D^2 = \left(d - \frac{1}{4} \sum_{i,j=1}^{n} \omega_{ij} E_i \cdot E_j \cdot \right)^2.$$

Expanding this yields

$$R = d\left(-\frac{1}{4}\sum_{i,j=1}^{n}\omega_{ij}E_i\cdot E_j\cdot\right) + \left(-\frac{1}{4}\sum_{i,j=1}^{n}\omega_{ij}E_i\cdot E_j\cdot\right)^2 = -\frac{1}{4}\sum_{i,j=1}^{n}\Omega_{ij}E_i\cdot E_j,$$

where

$$\Omega_{ij} = d\omega_{ij} + \sum_{k=1}^{m}\omega_{ik}\wedge\omega_{kj}.$$

In other words, we obtain the remarkable fact that the curvature R of the spin connection is exactly the element $\mathcal{R}$ encountered in (4.34) in the construction of the Euler form for an $SO(n)$ bundle.

We can define the components R_{ijkl} of the curvature with respect to the moving orthonormal frame $(E_1,\ldots,E_n)$ by

$$R_{ijkl} = \Omega_{ij}(E_k,E_l) \quad \text{so that} \quad \Omega_{ij} = \frac{1}{2}\sum_{k.l=1}^{n}R_{ijkl}\theta_k\wedge\theta_l,$$

where $(\theta_1,\ldots,\theta_n)$ is the dual moving orthonormal coframe.

Definition. If $(M,\langle\cdot,\cdot\rangle)$ is a Riemannian manifold, the *scalar curvature* of M is given by the formula

$$s = \sum_{i,j=1}^{n}R_{ijij}.$$

The scalar curvature is a weaker invariant than the sectional curvature or the Ricci curvature. It is natural to ask whether there are any topological obstructions to the existence of metrics of positive scalar curvature.

We claim that if R is the curvature of the spin connection, then

$$\sum_{i,j=1}^{n}E_i\cdot E_j\cdot R(E_iE_j) = \frac{s}{2}. \tag{5.12}$$

Indeed,

$$\begin{aligned}\sum_{i,j=1}^{n}E_i\cdot E_j\cdot R(E_iE_j) &= -\frac{1}{4}\sum_{i,j,k,l=1}^{n}E_i\cdot E_j\cdot E_k\cdot E_l\Omega_{kl}(E_iE_j)\\ &= -\frac{1}{4}\sum_{i,j,k,l=1}^{n}R_{ijkl}E_i\cdot E_j\cdot E_k\cdot E_l.\end{aligned}$$

If prime denotes the sum over those indices for which j, k and l are distinct, then

$$\sum{}' R_{ijkl}E_i\cdot E_j\cdot E_k\cdot E_l = \frac{1}{3}\sum{}'(R_{ijkl}+R_{iklj}+R_{iljk})E_i\cdot E_j\cdot E_k\cdot E_l = 0,$$

by one of the curvature symmetries. Similarly, we get zero if any three indices are distinct. Thus the only terms surviving are those for which i, j, k and l assume at most (and hence exactly) two values. Thus

$$\begin{aligned}
&\sum_{i,j=1}^{n} E_i \cdot E_j \cdot R(E_i E_j) \\
&\qquad = -\frac{1}{4}\sum_{i,j=1}^{n} R_{ijij} E_i \cdot E_j \cdot E_i \cdot E_j - \frac{1}{4}\sum_{i,j=1}^{n} R_{ijji} E_i \cdot E_j \cdot E_j \cdot E_i \\
&\qquad\qquad = \frac{1}{2}\sum_{i,j=1}^{n} R_{ijij} = \frac{s}{2},
\end{aligned}$$

proving our claim.

5.9 The Dirac operator

Suppose now that $(M, \langle\cdot,\cdot\rangle)$ is an oriented Riemannian manifold with a spin structure and let $S \to M$ be the spin bundle of M. We give S the spin connection ∇ inherited from the Levi-Civita connection on TM as described in the preceding section.

Definition. The *Dirac operator* on S is the first-order differential operator $\mathcal{D} : \Gamma(S) \to \Gamma(S)$ defined in terms of a local moving orthonormal frame $(E_1, \ldots, E_n)$ by

$$\mathcal{D}\psi = \sum_{i=1}^{n} E_i \nabla_{E_i} \psi.$$

Note that in the case in which the spin manifold M is simply Euclidean space $\mathbb{E}^n$ with standard Euclidean coordinates $(x_1, \ldots, x_n)$ and standard orthonormal frame $E_1 = (\partial)(\partial x_1)$, ..., $E_n = (\partial)(\partial x_n)$, the spin bundle S is trivial and

$$\mathcal{D}\psi = \sum_{i=1}^{n} E_i \frac{\partial \psi}{\partial x_i}.$$

In this case, the E_i's can be regarded as constant matrices which satisfy the identities

$$E_i \cdot E_j + E_j \cdot E_i = \begin{cases} 2, & \text{for } i = j, \\ 0, & \text{for } i \neq j, \end{cases}$$

and hence

$$\mathcal{D}^2\psi = (\mathcal{D} \circ \mathcal{D})\psi = -\sum_{i=1}^{n} \frac{\partial^2 \psi}{\partial x_i^2}.$$

In other words, the Dirac operator is (up to sign) the square root of the Euclidean Laplace operator. Indeed, the spin bundle is the bundle of smallest rank for which a square root of the Laplace operator can be constructed.

On a general curved spin manifold, the square of the Dirac operator is not minus the most obvious "rough" Laplace operator constructed by means of the Spin(n)-connection. We can proceed just as in §2.13 and define an operator

$$\nabla_{X,Y} : \Gamma(S) \to \Gamma(S), \quad \text{by} \quad \nabla_{X,Y}\psi = \nabla_X \nabla_Y \psi - \nabla_{\nabla_X Y}\psi,$$

for $X, Y \in \mathcal{X}(M)$. The point of this construction is that

$$\nabla_{fX,Y}\psi = f\nabla_{X,Y}\psi = \nabla_{X,fY}\psi,$$

so that $\nabla_{X,Y}\psi(p)$ depends only on $X(p)$, $Y(p)$ and $\psi(p)$. In other words, $\nabla_{X,Y}\psi$ is a tensor field. The *rough Laplace operator* Δ_R is then defined by

$$\Delta_R(\psi) = \sum_{i=1}^{n} \nabla_{E_i,E_i}\psi = \sum_{i=1}^{n} [\nabla_{E_i} \circ \nabla_{E_i} - \nabla_{\nabla_{E_i}E_i}]\psi,$$

where $(E_1, \ldots, E_n)$ is any choice of local moving orthonormal frame.

Proposition 1. *The rough Laplace operator on the spin bundle is related to the Dirac operator by the formula*

$$\mathcal{D}^2\psi = -\Delta_R\psi + \frac{s}{4}\psi. \tag{5.13}$$

To prove this we choose a moving orthonormal frame at a given point p so that $(\nabla_{E_i}E_j)(p) = 0$. Then at the point p,

$$\begin{aligned}
\mathcal{D}^2\psi = \left(\sum_{i=1}^{n} E_i \nabla_{E_i}\right)\left(\sum_{i=1}^{n} E_i \nabla_{E_i}\right)\psi &= \sum_{i,j=1}^{n} E_i \cdot E_j \nabla_{E_i}\nabla_{E_j}\psi \\
&= -\sum_{i=1}^{n} \nabla_{E_i}\nabla_{E_i}\psi + \frac{1}{2}\sum_{i,j=1}^{n} E_i \cdot E_j (\nabla_{E_i}\nabla_{E_j} - \nabla_{E_j}\nabla_{E_i})\psi \\
&= -\Delta_R\psi + \frac{1}{2}\sum_{i,j=1}^{n} E_i \cdot E_j R(E_iE_j)\psi = -\Delta_R\psi + \frac{s}{4}\psi,
\end{aligned}$$

where R is the curvature of the spin connection and we have utilized the curvature identity (5.12) at the very last step.

Recall that the spin bundle S has an Hermitian inner product $\langle \cdot, \cdot \rangle$, which is complex linear in the first variable, conjugate linear in the second, and that the Spin(n)-connection is unitary with respect to this Hermitian inner product. Moreover, the inner product preserves the direct sum decompositions $S = S_+ \oplus S_-$. Finally, Clifford multiplication by unit-lenght vectors preserves the inner product, so

$$\langle E_i \cdot \eta, E_i \cdot \phi \rangle = \langle \eta, \phi \rangle \quad \Rightarrow \quad \langle E_i \cdot \psi, \eta \rangle + \langle \psi, E_i \cdot \eta \rangle = 0.$$

We claim that the Dirac operator is "formally self-adjoint":

Proposition 2. *If ψ and η are sections of the spin bundle S, then*

$$\int_M \langle \mathcal{D}\psi, \eta \rangle \Theta_M = \int_M \langle \psi, \mathcal{D}\eta \rangle \Theta_M. \tag{5.14}$$

We describe the proof as given in [19]; see page 114. We choose a p moving orthonormal frame $(E_1, \ldots, E_n)$ on a neighborhood of a point p so that $(\nabla_{E_i} E_j)(p) = 0$. Then

$$\begin{aligned}\langle \mathcal{D}\psi, \eta \rangle(p) &= \sum_{i=1}^{n} \langle E_i \cdot \nabla_{E_i}\psi, \eta \rangle(p) = -\sum_{i=1}^{n} \langle \nabla_{E_i}\psi, E_i \cdot \eta \rangle(p) \\ &= -\sum_{i=1}^{n} \left[E_i \langle \psi, E_i \cdot \eta \rangle(p) - \langle \psi, \nabla_{E_i}(E_i \cdot \eta) \rangle(p) \right] \\ &= -\sum_{i=1}^{n} E_i \langle \psi, E_i \cdot \eta \rangle(p) + \langle \psi, \mathcal{D}\eta \rangle(p),\end{aligned}$$

because ∇ is a unitary connection. Thus (5.14) follows from the Divergence Theorem.

Suppose now that M is of even dimension $2m$. Then the spin connection preserves the direct sum decomposition $S = S_+ \oplus S_-$, while multiplication by E_i interchanges S_+ and S_-, so the Dirac operator $\mathcal{D}$ induces two operators

$$\mathcal{D}^+ : \Gamma(S_+) \to \Gamma(S_-) \quad \text{and} \quad \mathcal{D}^- : \Gamma(S_-) \to \Gamma(S_+).$$

It follows from Proposition 2 that

$$\int_M \langle \mathcal{D}^+\psi, \eta \rangle \Theta_M = \int_M \langle \psi, \mathcal{D}^-\eta \rangle \Theta_M;$$

in other words, $\mathcal{D}^+$ and $\mathcal{D}^-$ are "formal adjoints" of each other.

We can also integrate by parts in terms of the rough Laplace operator:

Proposition 3. *If $(M, \langle \cdot, \cdot \rangle)$ is an Riemannian manifold with spin structure, then for any $\psi, \eta \in \Gamma(S)$,*

$$\int_M \langle -\Delta_R(\psi), \eta \rangle \Theta_M = \int_M \langle D\psi, D\eta \rangle \Theta_M = \int_M \langle \psi, -\Delta_R(\eta) \rangle \Theta_M.$$

The proof is based upon Stokes's Theorem and is virtually identical to the proof of the Proposition in §2.13 for the rough Laplace operator on k-forms. Proposition 3 states that the rough Laplace operator Δ_R is formally self-adjoint.

We say that a section $\psi \in \Gamma(S)$ is a *harmonic spinor field* if it satisfies the equation $\mathcal{D}\psi = 0$. It follows from elliptic regularity theory (just as in the case of Hodge theory) that harmonic spinor fields are smooth and that the dimension of the space of harmonic spinor fields on a compact manifold is finite. (See e.g. [10].)

It follows from the preceding Propositions that if ψ is a harmonic spinor field, then

$$0 = \int_M \|\nabla\psi\|^2 \Theta_M + \int_M \frac{s}{4} \Theta_M,$$

and hence an oriented Riemannian manifold with positive scalar curvature cannot admit any nonzero harmonic spinor fields. We will exploit this fact in the next section to show that certain spin manifolds cannot have Riemannian metrics with positive scalar curvature.

5.10 The Atiyah-Singer Index Theorem

5.10.1 Index of the Dirac operator

Let M be an oriented manifold of dimension $2m$ with a spin structure and spin bundle $S = S_+ \oplus S_-$.

Definition. The *index* or more precisely the *analytic index* of the Dirac operator $\mathcal{D}^+ : \Gamma(S_+) \to \Gamma(S_-)$ is

$$\text{Index}(\mathcal{D}^+) = \dim \text{Ker}(\mathcal{D}^+) - \dim \text{Ker}(\mathcal{D}^-).$$

The Atiyah-Singer index theorem gives a topological expression for the index of the Dirac operator, and more general types of Dirac operators, such as Dirac operators with coefficients. The statement of this index theorem for the usual Dirac operator makes use of the notion of the $\hat{A}$-polynomial in the Pontrjagin classes.

To describe this, we let E be a vector bundle over M, and for simplicity, we start by assuming that the complexification $E \otimes \mathbb{C}$ can be written as a direct sum of complex line bundles,

$$E \otimes \mathbb{C} \cong L_1 \oplus \bar{L}_1 \oplus \cdots \oplus L_m \oplus \bar{L}_m.$$

If we set $x_i = c_1(L_i)$, then it follows from Proposition 2 of §**??** that

$$\begin{aligned} c(E \otimes \mathbb{C}) &= 1 + c_1(E \otimes \mathbb{C}) + c_2(E \otimes \mathbb{C}) = \cdots \\ &= (1 + x_1)(1 - x_1) \cdots (1 + x_m)(1 - x_m) = (1 - x_1^2) \cdots (1 - x_m)^2, \end{aligned}$$

and since $p_k(E) = (-1)^k c_k(E \otimes \mathbb{C})$,

$$p(E) = 1 + p_1(E) + p_2(E) + \cdots = (1 + x_1^2) \cdots (1 + x_m)^2,$$

which implies that

$$p_k(E) = s_k(x_1^2, \ldots, x_m^2),$$

where s_k denotes the k-th elementary symmetric function.

Next one determines the power series for the function

$$\frac{x/2}{\sinh(x/2)} = 1 - \frac{1}{3!}\left(\frac{x}{2}\right)^2 + \frac{7}{5 \cdot 3 \cdot 2^3}\left(\frac{x}{2}\right)^4 + \cdots .$$

Then the power series expansion of

$$\hat{A}(x_1, \ldots, x_m) = \Pi_{j=1}^m \frac{x_j/2}{\sinh(x_j/2)}$$

only involves the symmetric functions of $x_1^2, \ldots, x_m^2$, so we can write

$$\hat{A}(E) = \Pi_{j=1}^k \frac{x_j/2}{\sinh(x_j/2} = 1 + \hat{A}_1(p_1(E)) + \hat{A}_2(p_1(E)), p_2(E)) + \cdots$$

where each $\hat{A}_k(p_1(E)), \ldots, p_k(E))$ is a polynomial in $H^{4k}(M;\mathbb{R})$ in the Pontrjagin classes $p_j(E)$ of the vector bundle E. One thus obtains

$$\hat{A}_1(p_1) = -\frac{1}{24}p_1, \quad \hat{A}_2(p_1, p_2) = \frac{1}{2^7 \cdot 3^2 \cdot 5}(-4p_2 + 7p_1^2), \quad \ldots .$$

Finally, note that each polynomial $\hat{A}_k(p_1(E)), \ldots, p_k(E))$ in the Pontrjagin classes is well-defined, even if the direct sum decomposition of $E \otimes \mathbb{C}$ does not exist, because the $\hat{A}$ polynomials involve only the symmetric functions of the x_i^2's, that is, they can be defined directly in terms of the Pontrjagin classes. It is not difficult to show that

$$\hat{A}(E_1 \oplus E_2) = \hat{A}(E_1)\hat{A}(E_2).$$

Index Theorem for the Standard Dirac Operator. *If $(M, \langle \cdot, \cdot \rangle)$ is a compact spin manifold of dimension $4k$, then the index of the Dirac operator $\mathcal{D}^+ : \Gamma(S_+) \to \Gamma(S_-)$ is given by*

$$\text{Index}(\mathcal{D}^+) = \int_M \hat{A}_k(p_1(TM), \ldots, p_k(TM)) = \hat{A}(M), \tag{5.15}$$

where $\hat{A}(M)$ is called the $\hat{A}$-genus of M.

The proof of this theorem requires techniques such as the theory of pseudo-differential operators which are beyond the scope of this course (but are treated throughly in [19]). However, we note that from this theorem and the Weitzenböck formula, we immediately obtain the consequence:

Lichnerowicz Theorem. *A compact spin manifold with nonzero $\hat{A}$-genus cannot admit a metric with positive scalar curvature.*

We remark that the $\hat{A}$-genus can be refined to a topological invariant which is nonzero exactly for those compact simply connected manifolds of dimension at least five which do not admit metrics of positive scalar curvature [32]. In the special case where M has dimension four,

$$\mathrm{Index}(\mathcal{D}^+) = -\frac{1}{24}\int_M p_1(TM).$$

Thus the index is nonzero so long as $p_1(TM)) \neq 0$ and we see that a compact four-dimensional spin manifold with nontrivial first Pontrjagin class cannot admit a metric with positive scalar curvature.

Dirac operators with coefficients. Suppose that E is a complex vector bundle over the spin manifold M with Hermitian metric and unitary connection D_E. We can then define a connection on $S \otimes E$ by forcing the Leibniz rule to hold; thus for example,

$$D_E(\psi \otimes \sigma) = D\psi \otimes \sigma + \psi \otimes D_E\sigma,$$

when $\psi \in \Gamma(S)$ and $\sigma \in \Gamma(E)$. Moreover, we can construct the *Dirac operator with coefficients* in E,

$$\mathcal{D}_E : \Gamma(S \otimes E) \to \Gamma(S \otimes E), \quad \text{by} \quad \mathcal{D}_E(\psi \otimes \sigma) = \sum_{i=1}^n E_i \cdot \nabla_{E_i}(\psi \otimes \sigma)$$

which interchanges the summands in the direct sum decomposition

$$S \otimes E = S_+ \otimes E \oplus S_- \otimes E.$$

Let E^* be the dual bundle to E with the induced Hermitian inner product and unitary connection D_{E^*}. We can then restrict to $S_+ \otimes E$ obtaining the first-order elliptic operator

$$\mathcal{D}_E^+ : \Gamma(S_+ \otimes E) \to \Gamma(S_- \otimes E).$$

One can show that this operator has the formal adjoint

$$\mathcal{D}_{E^*}^- : \Gamma(S_- \otimes E^*) \to \Gamma(S_+ \otimes E^*).$$

The *index* of the operator $\mathcal{D}_E^+$ is therefore given by the formula

$$\mathrm{Index}(\mathcal{D}_E^+) = \dim \mathrm{Ker}(\mathcal{D}_E^+) - \dim \mathrm{Ker}(\mathcal{D}_{E^*}^-).$$

Atiyah-Singer Index Theorem for Spin Manifolds. *If $(M, \langle \cdot, \cdot \rangle)$ is an even-dimensional compact spin manifold and E is a complex vector bundles over M with Hermitian metric and unitary connection D_E, then the index of the Dirac operator with coefficients in E is given by*

$$\mathit{Index}(\mathcal{D}^+) = \int_M \hat{A}(p_1(TM), \ldots, p_k(TM))\mathit{ch}(E).$$

In this formula, both $\hat{A}(p_1(TM), \ldots, p_k(TM))$ and $\mathrm{ch}(E)$ are polynomials in the de Rham cohomology ring $H^*(M; \mathbb{R})$. To integrate such an expression over M one takes a representative of the homogeneous component in this polynomial in $H^n(M; \mathbb{R})$, where n is the dimension of M.

5.10.2 Spin^c structures*

The question now arises: Can we construct a Dirac operator on manifolds which are not necessarily spin? It turns out that we can extend the theory of the Dirac operator to more general manifolds by considering Dirac operators with coefficients in a complex vector bundle, or a "virtual complex vector bundle." An important special case is a spin^c-structure on a smooth manifold M, which allows us to define a Dirac operator with coefficients in certain virtual line bundles over M.

The definition of spin^c structure is quite similar to that of spin structure, but spin^c structures can be put on more manifolds. For example, every four-dimensional oriented Riemannian manifold admits a spin^c-structure.

To define the notion of spin^c structure, we suppose that $(M, \langle \cdot, \cdot \rangle)$ is an oriented smooth Riemannian manifold, and we can choose a trivializing open cover $\{U_\alpha : \alpha \in A\}$ for TM such that the corresponding transition functions take their values in $SO(n)$:

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to SO(n).$$

We can assume that $\{U_\alpha : \alpha \in A\}$ is a *good* open cover, meaning that any nonempty intersection $U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq 0$ of sets in the cover is contractible.

Recall that in §5.7 we constructed the Lie group

$$\mathrm{Spin}^c(n) = \frac{\mathrm{Spin}(n) \times S^1}{\mathbb{Z}_2},$$

where the $\mathbb{Z}_2$-action is described by $(\sigma, e^{i\theta}) \mapsto (-\sigma, -e^{i\theta})$. We have two Lie group homomorphisms

$$\rho : \mathrm{Spin}^c(n) \to SO(n), \quad \rho([\sigma, e^{i\theta}]) = \rho(\sigma),$$
$$\pi : \mathrm{Spin}^c(n) \to S^1, \quad \pi([\sigma, e^{i\theta}]) = e^{2i\theta}.$$

Definition. A *spin^c structure* on $(M, \langle \cdot, \cdot \rangle)$ is defined by an open covering $\{U_\alpha : \alpha \in A\}$ of M and a collection of transition functions

$$\sigma_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}^c(n)$$

such that the projections $g_{\alpha\beta} = \rho \circ \sigma_{\alpha\beta}$ define the $SO(n)$-structure on TM, and

$$\sigma_{\alpha\beta}\sigma_{\beta\gamma}\sigma_{\gamma\alpha} = 1 \quad \text{on} \quad U_\alpha \cap U_\beta \cap U_\gamma,$$

for all α, β and γ in A.

If $(M, \langle \cdot, \cdot \rangle)$ has a genuine spin structure given by transition functions

$$\tilde{g}_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}(n)$$

and L is a $U(1)$-bundle over M with transition functions

$$h_{\alpha\beta} : U_\alpha \cap U_\beta \to U(1),$$

we can construct a spin^c structure on M by taking the transition functions to be

$$\sigma_{\alpha\beta} = h_{\alpha\beta}\tilde{g}_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}^c(n),$$

and in fact all spin^c structures on spin manifolds arise in this way. Note that $\pi \circ \sigma_{\alpha\beta} = (h_{\alpha\beta})^2$.

If M is only a spin^c manifold, we can construct transition functions

$$\tilde{g}_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}(n) \quad \text{and} \quad h_{\alpha\beta} : U_\alpha \cap U_\beta \to U(1),$$

but now they satisfy only the weaker conditions

$$\tilde{g}_{\alpha\beta}\tilde{g}_{\beta\gamma}\tilde{g}_{\gamma\alpha} = h_{\alpha\beta}h_{\beta\gamma}h_{\gamma\alpha} = \pm 1 \quad \text{on} \quad U_\alpha \cap U_\beta \cap U_\gamma.$$

It is only the product $\sigma_{\alpha\beta} = h_{\alpha\beta}\tilde{g}_{\alpha\beta}$ which satisfies the "cocycle condition" allowing the construction of genuine vector bundles. Thus although one cannot define the spin bundles S_+ and S_-, or the line bundle L with transition functions $h_{\alpha\beta}$ on a spin^c manifold, one can define the tensor product

$$S \otimes L = (S_+ \otimes L) \oplus (S_- \otimes L). \tag{5.16}$$

Although the line bundle L itself is not well-defined on a spin^c manifold which is not spin, the transition functions

$$\pi \circ \sigma_{\alpha\beta} : U_\alpha \cap U_\beta \to U(1)$$

always define a complex line bundle over M, called the *determinant line bundle* of the spin^c structure, and denoted by L^2.

We will use somewhat nonstandard terminology and call L a *virtual line bundle* over M.

Just as we can construct a Spin(n)-connection on S_+ and S_- we can construct a unique Spin$^c(n)$-connection D_L on $S \otimes L$ which induces the Levi-Civita connection on TM and a given $U(1)$-connection on the determinant line bundle L. We can then define the Dirac operator with coefficients in L,

$$\mathcal{D}_L : \Gamma(S \otimes L) \to \Gamma(S \otimes L), \quad \text{by} \quad \mathcal{D}_L(\psi) = \sum_{i=1}^{n} E_i \cdot \nabla_{E_i}(\psi),$$

whenever $\psi \in \Gamma(S \otimes L)$, and this Dirac operator interchanges the two summands in the direct sum decomposition (5.16). The determinant line bundle L^2 has a dual line bundle $(L^*)^2$, and a $U(1)$-connection in L^2 induces a corresponding connection in $(L^*)^2$, so one can provide the bundle $S \otimes L^*$ with a dual Spin$^c(n)$-connection. The argument for Proposition 2 from §5.9 shows that

$$\int_M \langle \mathcal{D}_L \psi, \eta \rangle \Theta_M = \int_M \langle \psi, \mathcal{D}_{L^*} \eta \rangle \Theta_M, \tag{5.17}$$

whenever $\psi \in \Gamma(S \otimes L)$ and $\eta \in \Gamma(S \otimes L^*)$. In other words, the Dirac operator $\mathcal{D}_{L^*}$ is the formal adjoint to $\mathcal{D}_L$.

We can then restrict to $S_+ \otimes L$ obtaining a Dirac operator with coefficients in L, obtaining

$$\mathcal{D}_L^+ : \Gamma(S_+ \otimes L) \to \Gamma(S_- \otimes L),$$

an operator which has formal adjoint

$$\mathcal{D}_{L^*}^- : \Gamma(S_- \otimes L^*) \to \Gamma(S_+ \otimes L^*).$$

The *index* of this operator is defined by

$$\mathrm{Index}(\mathcal{D}_L^+) = \dim \mathrm{Ker}(\mathcal{D}_L^+) - \dim \mathrm{Ker}(\mathcal{D}_{L^*}^-).$$

Index Theorem for Spin^c Manifolds. *Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a compact even-dimensional oriented Riemannian manifold and L^2 is the determinant line bundle of a spin^c structure on M with Hermitian metric and unitary connection D_L. Then the index of the Dirac operator $\mathcal{D}_L^+$ with coefficients in L is given by*

$$\mathit{Index}(\mathcal{D}_L^+) = \int_M \hat{A}_k(p_1(TM), \ldots, p_k(TM)) ch(L).$$

5.10.3 Dirac operators on general manifolds*

We can now state the general version of the Atiyah-Singer index theorem, which does not require the base manifold M to have a spin or spin^c structure.

To do this, we must generalize the notion of Dirac operator with coefficients from virtual line bundle to virtual vector bundles. As before, we suppose that $(M, \langle \cdot, \cdot \rangle)$ is an oriented smooth Riemannian manifold, and that $\{U_\alpha : \alpha \in A\}$ is a trivializing cover for TM such that the corresponding transition functions take their values in $SO(n)$:

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to SO(n).$$

Since each $U_\alpha \cap U_\beta$ is contractible, we can lift these transition functions to maps

$$\tilde{g}_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}(n)$$

such that the projections $g_{\alpha\beta} = \rho \circ \tilde{g}_{\alpha\beta}$ define the $SO(n)$-structure on TM and

$$\tilde{g}_{\alpha\beta} \tilde{g}_{\beta\gamma} \tilde{g}_{\gamma\alpha} = \eta_{\alpha\beta\gamma} \quad \text{on} \quad U_\alpha \cap U_\beta \cap U_\gamma,$$

for all α, β and γ in A, where $\eta_{\alpha\beta\gamma} = \pm 1$. A *virtual complex vector bundle* E of rank m with Hermitian metric over M can then be defined by a collection of transition functions

$$h_{\alpha\beta} : U_\alpha \cap U_\beta \to U(m) \quad \text{such that} \quad h_{\alpha\beta} h_{\beta\gamma} h_{\gamma\alpha} = \eta_{\alpha\beta\gamma}.$$

The key point is that although E itself cannot be defined as a genuine vector bundle over M, the tensor products $S_+ \otimes E$ and $S_- \otimes E$ are well-defined vector bundles, because the $\mathrm{Spin}(n) \times U(m)$-valued transition functions $g_{\alpha\beta} \times h_{\alpha\beta}$ satisfy the relations

$$(g_{\alpha\beta} \times h_{\alpha\beta})(g_{\beta\gamma} \times h_{\beta\gamma})(g_{\gamma\alpha} \times h_{\gamma\alpha}) = 1,$$

for all α, β and γ in A. Moreover, the Chern character $ch(E)$ of a virtual complex vector bundle E can be defined as the square root of the Chern character of $E \otimes E$, which is a genuine complex vector bundle.

Just as in the case of spin^c structures, if the virtual $U(m)$-bundle E has a unitary connection D_E, we can define a Dirac operator with coefficients in E,

$$\mathcal{D}_E : \Gamma(S \otimes E) \to \Gamma(S \otimes E), \quad \text{by} \quad \mathcal{D}_E(\psi) = \sum_{i=1}^{n} E_i \cdot \nabla_{E_i}(\psi),$$

whenever $\psi \in \Gamma(S \otimes E)$. Moreover, the restriction to $S_+ \otimes E$ yields a Dirac operator

$$\mathcal{D}_E^+ : \Gamma(S_+ \otimes E) \to \Gamma(S_- \otimes E)$$

which has formal adjoint

$$\mathcal{D}_{E^*}^- : \Gamma(S_- \otimes E^*) \to \Gamma(S_+ \otimes E^*).$$

The *index* of $\mathcal{D}_E^+$ is defined by

$$\mathrm{Index}(\mathcal{D}_E^+) = \dim \mathrm{Ker}(\mathcal{D}_E^+) - \dim \mathrm{Ker}(\mathcal{D}_{E^*}^-).$$

Atiyah-Singer Index Theorem. *Suppose that $(M, \langle \cdot, \cdot \rangle)$ is a compact even-dimensional oriented Riemannian manifold and E is a virtual complex vector bundle over M with Hermitian metric and unitary connection D_E. Then the index of the Dirac operator $\mathcal{D}_E^+$ with coefficients in E is given by*

$$\mathit{Index}(\mathcal{D}_E^+) = \int_M \hat{A}_k(p_1(TM), \ldots, p_k(TM)) ch(E).$$

There is an alternate way of describing Dirac operators with coefficients on general manifolds. Let M be an oriented Riemannian manifold of dimension $2m$ with or without a spin structure, and let E be a direct sum of complex vector bundles over M, say $E = E_0 \oplus E_1$, which we regard as a $\mathbb{Z}_2$-graded vector bundle. We can then also regard $\mathrm{End}(E)$ as a $\mathbb{Z}_2$-graded vector bundle with direct sum decomposition

$$\mathrm{End}(E) = \mathrm{End}(E)_0 \oplus \mathrm{End}(E)_1,$$

where

$$\mathrm{End}(E)_0 = \mathrm{Hom}(E_0, E_0) \oplus \mathrm{Hom}(E_1, E_1),$$
$$\mathrm{End}(E)_1 = \mathrm{Hom}(E_0, E_1) \oplus \mathrm{Hom}(E_1, E_0).$$

Following [4], we say that E is a *bundle of Clifford modules* if there is a vector bundle homomorphism

$$c : TM \to \mathrm{End}(E)_1 \quad \text{such that} \quad c(v)c(w) + c(w)c(v) = -2\langle v, w\rangle.$$

Note that c induces a vector bundle map from the bundle of Clifford algebras into the vector bundle $\mathrm{End}(E)$,

$$c : \mathrm{Cl}(TM) \to \mathrm{End}(E) \quad \text{such that} \quad c : \mathrm{Cl}_i(TM) \subseteq \mathrm{End}(E)_i.$$

A connection ∇^E on the $\mathbb{Z}_2$-graded vector bundle $E = E_0 \oplus E_1$ is called a *Clifford connection* if it preserves the direct sum decomposition and satisfies the condition that

$$[\nabla^E_X, c(Y)] = c\left(\nabla_X Y\right),$$

where the connection on the right-hand side is the Levi-Civita connection on the bundle $\mathrm{Cl}(TM)$ of Clifford algebras. Given a Clifford connection ∇^E, we can define a corresponding *generalized Dirac operator*

$$\mathcal{D}_E : \Gamma(E) \to \Gamma(E) \quad \text{by} \quad \mathcal{D}_E(\psi) = \sum_{i=1}^{n} E_i \cdot \nabla_{E_i}(\psi \otimes \sigma),$$

where $(E_1, \ldots, E_n)$ is a locally defined moving orthonormal frame defined on M. This Dirac operator restricts to operators

$$\mathcal{D}^+_E : \Gamma(E_0) \to \Gamma(E_1) \quad \text{and} \quad \mathcal{D}^-_E : \Gamma(E_1) \to \Gamma(E_0).$$

We suppose now that the $\mathbb{Z}_2$-graded vector bundle $E = E_0 \oplus E_1$ has a Hermitian metric preserving the direct sum decomposition. Then given any generalized Dirac operator $\mathcal{D}^+_E : \Gamma(E_0) \to \Gamma(E_1)$ we can construct its *formal adjoint*, a first-order operator $(\mathcal{D}^+_E) : \Gamma(E_1) \to \Gamma(E_0)$ which satisfies the relation

$$\int_M \langle \mathcal{D}^+_E \psi, \eta\rangle \Theta_M = \int_M \langle \psi, (\mathcal{D}^+_E \eta)^*\rangle \Theta_M, , \tag{5.18}$$

whenever $\psi \in \Gamma(E_0)$ and $\eta \in \Gamma(E_1)$. The exact form of the formal adjoint is obtained by integration by parts. We can then define the *index* of $\mathcal{D}^+_E$ by the formula

$$\text{index of } \mathcal{D}^+_E = \dim \mathrm{Ker}(\mathcal{D}^+_E) - \dim \mathrm{Ker}((\mathcal{D}^-_E)^*).$$

The Atiyah-Singer Index Theorem stated above then gives a formula for the index of these generalized Dirac operators $\mathcal{D}^+_E$ in terms of characteristic classes of the base manifold M.

For a key example, we make $E = \Lambda^* TM \otimes \mathbb{C}$ into a $\mathbb{Z}_2$-graded vector bundle. To define the decomposition $E = E_0 \oplus E_1$, it is convenient to make us of the isomorphism $\Lambda^* TM \cong \mathrm{Cl}(TM)$ as well as use the complex volume element

$$\Theta_{\mathbb{C}} = i^m e_1 \cdot e_2 \cdots e_{2m}, \quad \text{satisfying} \quad \Theta_{\mathbb{C}}^2 = 1.$$

introduced before; see (5.7).

Let us suppose that M is $4k$-dimensional so that $\omega = (-1)^k e_1 \cdot e_2 \cdots e_{4k}$, and let

$$E_0 = \phi \in \mathrm{Cl}(TM) : \Theta_{\mathbb{C}} \cdot \phi = \phi\}, \quad E_1 = \phi \in \mathrm{Cl}(TM) : \Theta_{\mathbb{C}} \cdot \phi = -\phi\}.$$

It is easily verified that $e_i \Theta_{\mathbb{C}} = \Theta_{\mathbb{C}} e_i$ and hence

$$\mathrm{Cl}_i(TM) E_j \subseteq E_{i+j},$$

so once again the vector bundle morphism c makes $E = E_0 \oplus E_1$ into a bundle of Clifford modules. Once obtains the same Dirac operator $d + \delta$, but the index is different because of the different decomposition of $\Lambda^* TM \otimes \mathbb{C}$.

Exercise XXI. a. Show that if M is an oriented $(4k)$-dimensional Riemannian manifold then for $\phi \in \Lambda^k(TM) \cong \mathrm{Cl}^k(TM)$, $\Theta_{\mathbb{C}} \cdot \phi = \star\phi$, where $\star$ is the Hodge star.

b. Conclude that on an oriented $4k$-dimensional Riemannian manifold, $\star^2 = -1$.

We let

$$\Omega^{2k}_+(M) = \{\omega \in \Omega^{2k}(M) : \star\omega = \omega\}, \qquad \Omega^{2k}_-(M) = \{\omega \in \Omega^{2k}(M) : \star\omega = -\omega\},$$

and call $\Omega^{2k}_+(M)$ the space of *self-dual* $(2k)$-forms and $\Omega^{2k}_-(M)$ the space of *anti-self-dual* $(2k)$-forms. It then follows from the exercise that

$$\Omega^*_+ = \Gamma(E_0) = \sum_{i=0}^{2k-1} \{\omega + \star\omega : \omega \in \Omega^i(M)\} \oplus \Omega^{2k}_+(M),$$

$$\Omega^*_- = \Gamma(E_1) = \sum_{i=0}^{2k-1} \{\omega - \star\omega : \omega \in \Omega^i(M)\} \oplus \Omega^{2k}_+(M).$$

Thus we can divide the operator $d + \delta$ into a sum of operators

$$(d+\delta)^+ : \Omega^*_+ \to \Omega^*_- \quad \text{and} \quad (d+\delta)^- : \Omega^*_- \to \Omega^* + .$$

If we let $\mathcal{H}^{2k}_+(M)$ denote the space self-dual harmonic $(2k)$-forms, $\mathcal{H}^{2k}_+(M)$ the space of anti-self-dual harmonic $(2k)$-forms, it follows from Hodge theory that

$$\text{index of } ((d+\delta)^+) = \dim \mathcal{H}^2_+(M) - \dim \mathcal{H}^2_-(M).$$

If M is a compact oriented $(4k)$-dimensional manifold, the cup product defines a nondegenerate symmetric bilinear form

$$I : H^{2k}(M;\mathbb{R}) \times H^{2k}(M;\mathbb{R}) \to H^4(M:\mathbb{R}) \cong \mathbb{R}$$

by the formula

$$I([\alpha],[\beta]) = \int_M \alpha \wedge \beta.$$

Any such symmetric bilinear form can be represented by a matrix

$$\begin{pmatrix} I_{p\times p} & 0 \\ 0 & -I_{q\times q} \end{pmatrix}, \quad \text{where} \quad p+q = \dim H^2k(M;\mathbb{R}).$$

The difference $p-q$ is called the *signature* of M and is one of the basic topological invariants of M. For our choice of $\mathbb{Z}_2$-grading of $\Lambda^*TM \otimes \mathbb{C}$,

$$(\text{the index of the signature operator } (d+\delta)^+) = \text{signature of } M.$$

In this case, the Atiyah-Singer Index Theorem specializes to yield the Hirzebruch Signature Theorem, which gives a formula for the signature in terms of Pontrjagin classes of TM. More precisely, the signature is expressed in terms of a sequence of polynomials called the L-polynomials, which start with

$$L_1(p_1) = \frac{1}{3}p_1, \quad L_2(p_1,p_2) = \frac{1}{45}(7p_2 - p_1^2), \quad \ldots . \tag{5.19}$$

The L polynomials are generated in much the same way as the $\hat{A}$-polynomials. Namely, one imagines that the Pontrjagin classes are written as elementary symmetric functions in the indeterminates x_1^2, x_2^2, ..., and writes the formal power series

$$L = \Pi_{j=1}^{\infty} \frac{x_j}{\tanh x_j} = 1 + c_1 p_1 + (c_{11}p_1^2 + c_2 p_2) + \cdots .$$

The homogeneous term of degree $4k$ then constitutes the polynomial $L_k(p_1, \ldots, p_k)$.

Hirzebruch Signature Theorem. *If $(M, \langle\cdot,\cdot\rangle)$ is a compact oriented manifold of dimension $4k$, then the index of the signature operator $(d+\delta)^+ : \Omega_+^* \to \Omega_-^*$ is given by*

$$\textit{Signature of } M = \int_M L_k(p_1(TM), \ldots, p_k(TM)).$$

In particular, if M is an oriented four-dimensional manifold,

$$\text{Signature of } M = \frac{1}{3}\int_M p_1(TM).$$

5.10.4 Topological invariants of four-manifolds*

Suppose that M is a compact oriented four-dimensional Riemannian manifold. Then the Gauss-Bonnet Theorem and the Hirzebruch Signature Theorem give quite similar relating integrals of curvature to topological invariants. Indeed, if $b_i = \dim H^i(M : \mathbb{R})$, $b_2^+ = \dim \mathcal{H}_+^2(M)$ and $b_2^- = \dim \mathcal{H}_-^2(M)$, then

$$\chi(M) = 1 - b_1 + b_2 - b_3 + 1 = \frac{1}{4\pi^2}\int_M \mathrm{Pf}(\Omega),$$

while

$$\text{Signature of } M = b_2^+ - b_2^- = \frac{1}{3}\int_M p_1(\Omega) = \frac{1}{24\pi^2}\int_M \sum_{i,j=1}^4 \Omega_{ij}\wedge\Omega_{ji},$$

the middle equality following from the Hirzebruch signature theorem.

In the case where the compact oriented four-dimensional has a spin structure, the characteristic classes of the spin bundles S_+ and S_- are determined by the Euler and Pontrjagin classes of M. Indeed, in this case one could prove

$$\int_M c_2(S_+) = \frac{1}{2}\int_M e(TM) - \frac{1}{4}\int_M p_1(TM),$$
$$\int_M c_2(S_-) = -\frac{1}{2}\int_M e(TM) - \frac{1}{4}\int_M p_1(TM).$$

For simply connected four-dimensional compact spin manifolds, $b_1 = b_3 = 0$ and b_2^+ and b_2^- are completely determined by the Euler and Pontrjagin classes.

The Euler characteristic and signature are among the most important invariants of smooth four-dimensional manifolds. In fact, using results of Freedman, it can be proven that two smooth simply connected four-dimensional manifolds are homeomorphic if and only if they have the same Euler characteristic and signature, and are both either spin or non-spin.

However, two smooth four-manifolds can be homeomorphic without being diffeomorphic. This is one of the striking consequences of the invariants for four-manifolds discovered by Donaldson, and the closely related invariants of Seiberg and Witten.

5.11 Exotic spheres*

One of the striking applications of the Hirzebruch signature theorem is Milnor's proof [24] of the existence of smooth manifolds which are homeomorphic but not diffeomorphic to S^7. These are called *exotic spheres*. A nice modern presentation of which exotic spheres might admit metrics of positive sectional, Ricci or scalar curvature can be found in [17].

These exotic spheres are the total spaces of 3-sphere bundles over S^4. To construct them, we regard S^4 as "quaternionic projective space." Thus if $\mathbb{H}$ denotes the space of quaternions, we define an equivalence relation $\sim$ on $\mathbb{H}^2 - \{0\}$ by

$$(Q_1.Q_2) \sim (Q_1'.Q_2') \quad \Leftrightarrow \quad \text{there exists } \Lambda \in \mathbb{H} - \{0\} \text{ such that } Q_i = \Lambda Q_i'.$$

We let $[Q_1, Q_2]$ denote the equivalence class of (Q_1, Q_2). The set of all such equivalence classes is a manifold which is diffeomorphic to S^4.

We call $[0, 1]$ the north pole, $[1, 0]$ the south pole and let $U = S^4 - \{[1, 0]\}$

and $V = S^4 - \{[0,1]\}$. We then have two stereographic projection charts

$$\phi : U \to \mathbb{H} \quad \text{defined by} \quad \phi([Q_1, Q_2]) = Q_2^{-1} Q_1$$
$$\text{and} \quad \psi : V \to \mathbb{H} \quad \text{defined by} \quad \phi([Q_1, Q_2]) = Q_1^{-1} Q_2.$$

Note that $\phi \circ \psi^{-1}(Q) = Q^{-1}$.

We now define a family of oriented four-dimensional real vector bundles over S^4. These are defined by giving a transition function

$$g_{VU} : U \cap V \to GL^+(4, \mathbb{R}).$$

In fact, it would suffice to define the transition function on the equator

$$S^3 = \{[Q_1, Q_2] \in S^4 : |Q_1| = |Q_2|\}.$$

Indeed, the four-dimensional oriented real vector bundles are in one-to-one correspondence with the homotopy classes of maps $S^3 \to GL^+(4, \mathbb{R})$.

For fixed $r, s \in \mathbb{Z}$, we define

$$g_{VU}^{rs} : U \cup V \to GL^+(4, \mathbb{R}) \quad \text{by} \quad g_{VU}^{rs} \circ \phi^{-1}(Q)X = \frac{Q^r X Q^s}{|Q|^{r+s}}. \tag{5.20}$$

This defines an oriented four-dimensional real vector bundle (E_{rs}, π, S^4). Note that E_{rs} is obtained from $(\mathbb{H} - \{0\}) \times \mathbb{H}$ and $(\mathbb{H} - \{0\}) \times \mathbb{H}$ by the identification

$$(Q, X) \quad \leftrightarrow \quad (Q^{-1}, Q^r X Q^s).$$

For $Q \in S^3$ we find that $g_{VU}^{rs}(Q) \in SO(4)$ so the bundle E_{rs} is endowed with a fiber metric $\langle \cdot, \cdot \rangle$, and we set

$$B_{rs}^8 = \{e \in E_{rs} : \langle e, e \rangle \leq 1\}, \qquad M_{rs}^7 = \partial B_{rs}^8.$$

We will see that for certain choices of r and s, M_{rs}^7 is homeomorphic but not diffeomorphic to S^7.

We need to calculate the Pontrjagin class p_1 of the bundles E_{rs}. First, we verify that

$$p_1(E_{r+r', s+s'}) = p_1(E_{r,s}) + p_1(E_{r,'s'}),$$

from which we can conclude that

$$p_1(E_{r,s}) = (ar + bs)[\alpha],$$

where $a, b \in \mathbb{Z}$ and $[\alpha]$ is the standard generator of $H^4(S^4)$ such that α integrates to one. To determine the integers a and b we need only calculate the Pontrjagin classes for certain examples. First, we verify that $E_{1,1}$ is just the tangent bundle to S^4, so

$$p_1(E_{1,1}) = p_1(TS^4) = 0,$$

as we saw in Exercise XIX. Thus $a + b = 0$ and

$$p_1(E_{r,s}) = a(r - s)[\alpha].$$

On the other hand, if $r = 0$ and $s = -1$, then we get the canonical quaternionic line bundle over S^4. It is straightforward to verify that in this case

$$p_1(E_{0,-1}) = \pm 2[\alpha], \quad \text{so} \quad p_1(E_{r,s}) = \pm 2(r-s)[\alpha].$$

We can now present the key theorems which allow us to prove the existence of exotic spheres:

Theorem 1. *If $r + s = 1$, then M^7_{rs} is homeomorphic to S^7.*

Theorem 2. *Suppose that $r + s = 1$. If $(r - s)^2 \neq \pm 1$ modulo 7, then M^7_{rs} is not diffeomorphic to S^7.*

Thus for example, suppose that $r = 2$ and $s = -1$. Then $r + s = 1$ so M^7_{rs} is homeomorphic to S^7, but $(r - s)^2 = 9$ which is not ± 1 modulo 7, so M^7_{rs} is not diffeomorphic to S^7.

To prove Theorem 1, we construct a function $f : M^7_{rs} \to \mathbb{R}$ which has only two critical points, both nondegenerate, and apply a theorem of Reeb (see [25], §4, Theorem 4.1).

From the transition function (5.20) we see that we can regard M^7_{rs} as obtained by indentifying

$$(Q, X) \in \mathbb{H} \times S^3 \quad \text{with} \quad (Q', X') \in \mathbb{H} \times S^3,$$
$$\text{where} \quad Q' = Q^{-1} \quad \text{and} \quad X' = \frac{Q^r X Q^{1-r}}{|Q|},$$

where we have used the hypothesis that $r + s = 1$. We can also write the latter relation as

$$X' = Q^r \frac{XQ}{|Q|} Q^{-r}, \quad X = Q^{-r} \frac{X'Q'}{|Q'|} Q^r$$

If Q and X are any two quaternions, $\mathrm{Re}(QXQ^{-1}) = \mathrm{Re}(X)$, so

$$\mathrm{Re}(X) = \frac{\mathrm{Re}(X'Q')}{|Q'|} \quad \text{or} \quad \frac{\mathrm{Re}(X)}{\sqrt{1+|Q|^2}} = \frac{\mathrm{Re}(X'Q')}{\sqrt{1+|X'Q'|^2}}.$$

Thus we can define a smooth function $f : M^7_{rs} \to \mathbb{R}$ by setting

$$f \circ \tilde{\phi}^{-1}(Q, X) = \frac{\mathrm{Re}(X)}{\sqrt{1+|Q|^2}} = \frac{\mathrm{Re}(X'Q')}{\sqrt{1+|X'Q'|^2}} = f \circ \tilde{\psi}^{-1}(Q', X'),$$

where

$$\tilde{\phi} : M^7_{rs}|U \to \mathbb{H} \times S^3, \quad \tilde{\psi} : M^7_{rs}|V \to \mathbb{H} \times S^3$$

are the local trivializations over U and V. It is now straightforward to verify that f has two critical points, both nondegenerate, at the points where $Q = 0$ and $X = \pm 1$. Thus it follows from Reeb's Theorem that M^7_{rs} is homeomorphic to S^7.

To prove Theorem 2, we suppose that M^7_{rs} is diffeomorphic to S^7 via a diffeomorphism $F : M^7_{rs} \to S^7$. Note that $S^7 = \partial D^8$, where D^8 is the usual eight-dimensional disk. We can therefore construct a smooth compact oriented eight-dimensional manifold C^8_{rs} by using F to identify B^8_{rs} and $-D^8$ along their common boundary.

Since B^8_{rs} is homotopy equivalent to S^4, the signature of C^8_{rs} must be ± 1. Thus it follows from the Hirzebruch signature theorem that

$$\pm 45 = 7p_2(TC^8_{rs})[C^8_{rs}] - p_1^2(TC^8_{rs})[C^8_{rs}],$$

which implies that

$$p_1^2(TC^8_{rs})[C^8_{rs}] = \pm 3 \quad \text{modulo } 7. \tag{5.21}$$

So we are led to ask the question: What is $p_1^2(TC^8_{rs})[C^8_{rs}]$? To answer this, we note that the tangent bundle of B^8_{rs} is a direct sum $TB^8_{rs} = V \oplus H$, where H is a horizontal bundle and V is a vertical bundle. It is immediately verified that $H = \pi^*(TS^4)$ while $V = \pi^*(E_{rs})$. Hence

$$p_1(H) = 0, \qquad p_1(V) = \pm 2(r-s)[\alpha],$$

where $[\alpha]$ is the generator of $H^4(S^4)$ such that $\int_{S^4} \alpha = 1$. Since $H^4(C^8_{rs}) \to H^4(B^8_{rs})$ is an isomorphism, we find that

$$p_1(TC^8_{rs}) = \pm 2(r-s)[\hat{\alpha}],$$

where $[\hat{\alpha}]$ is a generator of $H^4(C^8_{rs})$. By the integer version of Poincaré duality, $[\hat{\alpha}] \cup [\hat{\alpha}]$ is a generator of $H^*(C^8_{rs})$ such that

$$([\hat{\alpha}] \cup [\hat{\alpha}])([C^8_{rs}]) = \pm 1.$$

Hence

$$p_1^2(TC^8_{rs})[C^8_{rs}] = 4(r-s)^2.$$

From (5.21) we conclude that $4(r-s)^2 = \pm 3$ modulo 7 or $(r-s)^2 = \pm 1$ modulo 7. This proves Theorem 2.

Bibliography

[1] C. B. Allendoerfer, *The Euler number of a Riemann manifold*, Amer. J. Math. **62** (1940), 243-248.

[2] C. B. Allendoerfer and A. Weil, *The Gauss-Bonnet theorem for Riemannian polyhedra*, Trans. Amer. Math. Soc. **53** (1943), 101-129.

[3] M. F. Atiyah and I. M. Singer, *The index of elliptic operators on compact manifolds*, Bull. Amer. Math. Soc. **69** (1963), 422-433.

[4] N. Berline, E. Getzler and M. Vergne, *Heat kernels and Dirac operators*, Springer, New York, 1992.

[5] William M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, Academic Press, New York, 2003.

[6] C. Böhm and B. Wilking, *Manifolds with positive curvature operators are space forms*, Annals of Math. **187** (2008), 1079-1097.

[7] R. Bott and L. Tu, *Differential forms in algebraic topology*, Springer Verlag, New York, 1982.

[8] S. S. Chern, *A simple intrinsic proof of the Gauss-Bonnet formula for closed Riemannian manifolds*, Annals of Math. **45** (1944), 747-752.

[9] M. do Carmo, *Riemannian geometry*, Birkhäuser, Boston, 1992.

[10] L. C. Evans, *Partial differential equations*, Amer. Math. Soc., Providence RI, 1998.

[11] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Jon Wiley, New York, 1978.

[12] D. Gromoll and W. Meyer, *Periodic geodesics on compact Riemannian manifolds*, J. Differential Geometry **3** (1969), 493-510.

[13] M. Gromov, *Homotopical effects of dilatation*, J. Differential Geometry **13** (1978), 303-310.

[14] A. Hatcher, *Algebraic topology*, Cambridge Univ. Press, Cambridge, UK, 2002.

[15] M. Hirsch, *Differential topology*, Springer, New York, 1997.

[16] J. Jost, *Riemannian geometry and global analysis, fifth edition*, Springer, New York, 2008.

[17] M. Joachim and D. Writh, *Exotic spheres and curvature*, Bull. Amer. Math. Soc. **45** (2009), 595-616.

[18] S. Lang, *Differential and Riemannian manifolds*, Springer, New York, 1995.

[19] H. B. Lawson and M. L. Michelsohn, *Spin geometry*, Princeton Univ. Press, Princeton, NJ, 1989.

[20] J. M. Lee, *Riemannian manifolds*, Springer, New York, 1997.

[21] J. Lohkamp, *Metrics of negative Ricci curvature*, Annals of Math. **140** (1994), 655-683.

[22] V. Mathai and D. Quillen, *Superconnections, Thom classes and equivariant differential forms*, Topology **25** (1986), 85-110.

[23] J. P. May, *A concise course in algebraic topology*, University of Chicago Press, Chicago, 1999.

[24] J. Milnor, *On manifolds homeomorphic to the* 7*-sphere* Annals of Math. **64** (1956), 399-405.

[25] J. Milnor, *Morse theory*, Annals of Math. Studies **51**, Princeton Univ. Press, Princeton, NJ, 1963.

[26] J. Milnor and J. Stasheff, *Characteristic classes*, Princeton Univ. Press, Princeton NJ, 1974.

[27] C.W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*, W. H. Freeman, New York, 1973.

[28] J. Nash, *The imbedding problem for Riemannian manifolds*. Annals of Math. **63** (1956), 20-63.

[29] A. Pressley, *Elementary differential geometry*, Springer, New York, 2001.

[30] D. Quillen, *Superconnections and the Chern character*, Topology **24** (1985), 89-95.

[31] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, New York, 1964.

[32] S. Stolz, *Simply connected manifolds of positive scalar curvature*, Annals of Math. **136** (1992), 511-540.

[33] M. Taylor, *Partial differential equations; basic theory*, Springer, New York, 1996.

[34] W. Thurston, *Three-dimensional geometry and topology*, Princeton Univ. Press, Princeton, NJ, 1997.

[35] R. M. Wald, *General relativity*, University of Chicago Press, Chicago, 1984.

[36] F. W. Warner, *Foundations of differentiable manifolds and Lie groups*, Scott-Foresman, Glenview, Ill, 1971.

[37] J. A. Wolf, *Spaces of constant curvature*, McGraw-Hill, New York, 1967.