DIFFERENTIAL GEOMETRY AND PHYSICS

Jouko Mickelsson

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

CHAPTER 1 : DIFFERENTIABLE MANIFOLDS

1.1 The definition of a differentiable manifold

Let M be a topological space. This means that we have a family Ω of *open sets* defined on M. These satisfy

(1) $\emptyset, M \in \Omega$
(2) the union of any family of open sets is open
(3) the intersection of a finite family of open sets is open

We normally assume also the *Hausdorff property:* For any pair x, y of distinct points there is a pair of nonoverlapping open sets U, V such that $x \in U$ and $y \in V$. In addition, we shall also assume that our manifolds are *paracompact.* This implies the existence of locally finite partitions of unity which is needed in the integration theory in Section 2.5. For finite dimensional manifolds the paracompactness is not considered to be a serious restriction; however, there are lots of infinite dimensional manifolds which are not paracompact.

In any topological space one can define the notion of convergence. A sequence $x_1, x_2, x_3, \dots$ converges towards $x \in M$ if any open set U such that $x \in U$ contains all the points x_n except a finite set.

The basic example of a topological space is $\mathbb{R}^n$ equipped with the Euclidean norm $||x||^2 = x_1^2 + x_2^2 + \cdots + x_n^2$ for $x = (x_1, \dots, x_n)$. A set $U \subset \mathbb{R}^n$ is open if for any $x \in U$ there is a positive number $\epsilon = \epsilon(x)$ such that $y \in U$ if $||x - y|| < \epsilon$. The convergence is then the usual one: $x^{(n)} \to x$ if for any $\epsilon > 0$ there is an integer n_ϵ

such that $||x - x^{(n)}|| < \epsilon$ for $n > n_\epsilon$.

Actually, all the spaces we study in (finite dimensional) differential geometry are *locally homeomorphic* to $\mathbb{R}^n$.

Definition. *A topological space M is called a smooth manifold* of dimension n if 1) there is a family of open sets U_α (with $\alpha \in \Lambda$) such that the union of all U_α's is equal to M, 2) for each α there is a homeomorphism $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ such that 3) the *coordinate transformations* $\phi_\alpha \circ \phi_\beta^{-1}$ on their domains of definition are smooth functions in $\mathbb{R}^n$.

Example 1 $\mathbb{R}^n$ is a smooth manifold. We need only one *coordinate chart* $U = M$ with $\phi : U \to \mathbb{R}^n$ the identity mapping.

Example 2 The same as above, but take $M \subset \mathbb{R}^n$ any open set.

Example 3 Take $M = S^1$, the unit circle. Set U equal to the subset parametrized by the polar angle $-0.1 < \phi < \pi + 0.1$ and V equal to the set $\pi < \phi < 2\pi$. Then $U \cap V$ consists of two intervals $\pi < \phi < \pi + 0.1$ and $-0.1 < \phi < 0 \sim 2\pi - 0.1 < \phi < 2\pi$. The coordinate transformation is the identity map $\phi \mapsto \phi$ on the former and the translation $\phi \mapsto \phi + 2\pi$ on the latter interval.

Exercise Define a manifold structure on the unit sphere S^2.

Example 4 The group $GL(n, \mathbb{R})$ of invertible real $n \times n$ matrices is a smooth manifold as an open subset of $\mathbb{R}^{n^2}$. It is an open subset since it is a complement of the closed surface determined by the polynomial equation $\det A = 0$.

1.2 Differentiable maps

Let M, N be a pair of smooth manifolds (of dimensions m, n) and $f : M \to N$ a continuous map. If (U, ϕ) is a local coordinate chart on M and (V, ψ) a coordinate chart on N then we have a map $\psi \circ f \circ \phi^{-1}$ from some open subset of $\mathbb{R}^m$ to an open subset of $\mathbb{R}^n$. If the composite map is smooth for any pair of coordinate charts we say that f is smooth. The reader should convince himself that the condition of smoothness for f does not depend on the choice of coordinate charts. From elementary results in differential calculus it follows that if $g : N \to P$ is another smooth map then also $g \circ f : M \to P$ is smooth.

Note that we can write the map $\psi \circ f \circ \phi^{-1}$ as $y = (y_1, \ldots, y_n) = (y_1(x_1, \ldots, x_m), \ldots$

$\ldots, y_n(x_1, \ldots, x_m))$ in terms of the Cartesian coordinates. Smoothness of f simply means that the coordinate functions $y_i(x_1, x_2, \ldots, x_m)$ are smooth functions.

Remark In a given topological space M one can often construct different inequivalent smooth structures. That is, one might be able to construct atlases $\{(U_\alpha, \phi_\alpha)\}$ and $\{(V_\alpha, \psi_\alpha)\}$ such that both define a structure of smooth manifold, say M_U and M_V, but the manifolds M_U, M_V are not diffeomorphic (see the definition below). A famous example of this phenomen are the spheres S^7, S^{11} (John Milnor, 1956). On the sphere S^7 there are exactly 28 inequivalent differentiable structures! On the Euclidean space $\mathbb{R}^4$ there is an infinite number of differentiable structures (S.K. Donaldson, 1983).

A diffeomorphism is a one-to-one smooth map $f : M \to N$ such that its inverse $f^{-1} : N \to M$ is also smooth. The set of diffeomorphisms $M \to M$ forms a group Diff(M). A smooth map $f : M \to N$ is *an immersion* if at each point $p \in M$ the rank of the derivative $\frac{dh}{dx}$ is equal to the dimension of M. Here $h = \psi \circ f \circ \phi^{-1}$ with the notation as before. Finally $f : M \to N$ is *an embedding* if f is injective and it is an immersion; in that case $f(M) \subset N$ is *an embedded submanifold.*

A smooth curve on a manifold M is a smooth map γ from an open interval of the real axes to M. Let $p \in M$ and (U, ϕ) a coordinate chart with $p \in U$. Assume that curves γ_1, γ_2 go through p, let us say $p = \gamma_i(0)$. We say that the curves are equivalent at p, $\gamma_1 \sim \gamma_2$, if

$$\frac{d}{dt}\phi(\gamma_1(t))|_{t=0} = \frac{d}{dt}\phi(\gamma_2(t))|_{t=0}.$$

This relation does not depend on the choice of (U, ϕ) as is easily seen by the help of the chain rule:

$$\frac{d}{dt}\psi(\gamma_1(t)) - \frac{d}{dt}\psi(\gamma_2(t)) = (\psi \circ \phi^{-1})' \cdot \left(\frac{d}{dt}\phi(\gamma_1(t)) - \frac{d}{dt}\phi(\gamma_2(t))\right) = 0$$

at the point $t = 0$. Clearly if $\gamma_1 \sim \gamma_2$ and $\gamma_2 \sim \gamma_3$ at the point p then also $\gamma_1 \sim \gamma_3$ and $\gamma_2 \sim \gamma_1$. Trivially $\gamma \sim \gamma$ for any curve γ through p so that " $\sim$ " is an equivalence relation.

A tangent vector v at a point p is an equivalence class of smooth curves $[\gamma]$ through p. For a given chart (U, ϕ) at p the equivalence classes are parametrized by the vector

$$\frac{d}{dt}\phi(\gamma(t))|_{t=0} \in \mathbb{R}^n.$$

Thus the space T_pM of tangent vectors $v = [\gamma]$ inherits the natural linear structure of $\mathbb{R}^n$. Again, it is a simple exercise using the chain rule that the linear structure does not depend on the choice of the coordinate chart.

We denote by TM the disjoint union of all the tangent spaces T_pM. This is called the *tangent bundle* of M. We shall define a smooth structure on TM. Let $p \in M$ and (U, ϕ) a coordinate chart at p. Let $\pi : TM \to M$ the natural projection, $(p, v) \mapsto p$. Define $\tilde{\phi} : \pi^{-1}(U) \to \mathbb{R}^n \times \mathbb{R}^n$ as

$$\tilde{\phi}(p, [\gamma]) = (\phi(p), \frac{d}{dt}\phi(\gamma(t))|_{t=0}).$$

If now (V, ψ) is another coordinate chart at p then

$$(\tilde{\phi} \circ \tilde{\psi}^{-1})(x, v) = (\phi(\psi^{-1}(x)), (\phi \circ \psi^{-1})'(x)v),$$

by the chain rule. It follows that $\tilde{\phi} \circ \tilde{\psi}^{-1}$ is smooth in its domain of definition and thus the pairs $(\pi^{-1}(U), \tilde{\phi})$ form an atlas on TM, giving TM a smooth structure.

Example 1 If M is an open set in $\mathbb{R}^n$ then $TM = M \times \mathbb{R}^n$.

Example 2 Let $M = S^1$. Writing $z \in S^1$ as a complex number of unit modulus, consider curves through z written as $\gamma(t) = ze^{ivt}$ with $v \in \mathbb{R}$. This gives in fact a parametrization for the equivalence classes $[\gamma]$ as vectors in $\mathbb{R}$. The tangent spaces at different points z_1, z_2 are related by the phase shift $z_1 z_2^{-1}$ and it follows that TM is simply the product $S^1 \times \mathbb{R}$.

Example 3 In general, $TM \neq M \times \mathbb{R}^n$. The simplest example for this is the unit sphere $M = S^2$. Using the spherical coordinates, for example, one can identify the tangent space at a given point (θ, ϕ) as the plane $\mathbb{R}^2$. However, there is no natural way to identify the tangent spaces at different points on the sphere; the sphere is not parallelizable. This is the content of the famous hairy ball theorem. Any smooth vector field on the sphere has zeros. (If there were a globally nonzero vector field on S^2 we would obtain a basis in all the tangent spaces by taking a (oriented) unit normal vector field to the given vector field. Together they would form a basis in the tangent spaces and could be used for identifying the tangent spaces as a standard $\mathbb{R}^2$.)

Exercise The unit 3-sphere S^3 can be thought of complex unitary 2×2 matrices with determinant =1. Use this fact to show that the tangent bundle is trivial, $TS^3 = S^3 \times \mathbb{R}^3$.

Let $f : M \to N$ be a smooth map. We define a linear map

$$T_p f : T_p M \to T_{f(p)} N, \text{ as } T_p f \cdot [\gamma] = [f \circ \gamma],$$

where γ is a curve through the point p. This map is expressed in terms of local coordinates as follows. Let (U, ϕ) be a coordinate chart at p and (V, ψ) a chart at $f(p) \in N$. Then the coordinates for $[\gamma] \in T_p M$ are $v = \frac{d}{dt}\phi(\gamma(t))|_{t=0}$ and the coordinates for $[f \circ \gamma] \in T_{f(p)} N$ are $w = \frac{d}{dt}\psi(f(\gamma(t)))|_{t=0}$. But by the chain rule,

$$w = (\psi \circ f \circ \phi^{-1})'(x) \cdot \frac{d}{dt}\phi(\gamma(t))|_{t=0} = (\psi \circ f \circ \phi^{-1})'(x) \cdot v$$

with $x = \phi(p)$. Thus in local coordinates the linear map $T_p f$ is the derivative of $\psi \circ f \circ \phi^{-1}$ at the point x. Putting together all the maps $T_p f$ we obtain a map

$$Tf : TM \to TN.$$

Proposition. *The map $Tf : TM \to TN$ is smooth.*

Proof. Recall that the coordinate charts $(U, \phi), (V, \psi)$ on M, N, respectivly, lead to coordinate charts $(\pi^{-1}(U), \tilde{\phi})$ and $(\pi^{-1}(V), \tilde{\psi})$ on TM, TN. Now

$$(\tilde{\psi} \circ Tf \circ \tilde{\phi}^{-1})(x, v) = ((\psi \circ f \circ \phi^{-1})(x), (\psi \circ f \circ \phi^{-1})'(x)v)$$

for $(x, v) \in \tilde{\phi}(\pi^{-1}(U)) \in \mathbb{R}^m \times \mathbb{R}^m$. Both component functions are smooth and thus Tf is smooth by definition.

If $f : M \to N$ and $g : N \to P$ are smooth maps then $g \circ f : M \to P$ is smooth and

$$T(g \circ f) = Tg \circ Tf.$$

To see this, the curve γ through $p \in M$ is first mapped to $f \circ \gamma$ through $f(p) \in N$ and further, by Tg, to the curve $g \circ f \circ \gamma$ through $g(f(p)) \in P$.

In terms of local coordinates x_i at p, y_i at $f(p)$ and z_i at $g(f(p))$ the chain rule becomes the standard formula,

$$\frac{\partial z_i}{\partial x_j} = \sum_k \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j}.$$

1.3 Vector fields

We denote by $C^\infty(M)$ the algebra of smooth real valued functions on M. A *derivation* of the algebra $C^\infty(M)$ is a linear map $d : C^\infty(M) \to C^\infty(M)$ such that

$$d(fg) = d(f)g + fd(g)$$

for all f, g. Let $v \in T_pM$ and $f \in C^\infty(M)$. Choose a curve γ through p representing v. Set

$$v \cdot f = \frac{d}{dt} f(\gamma(t))|_{t=0}.$$

Clearly $v : C^\infty(M) \to \mathbb{R}$ is linear. Furthermore,

$$v \cdot (fg) = \frac{d}{dt} f(\gamma(t))|_{t=0} g(\gamma(0)) + f(\gamma(0)) \frac{d}{dt} g(\gamma(t))|_{t=0} = (v \cdot f)g(p) + f(p)(v \cdot g).$$

A vector field on a manifold M is a smooth distribution of tangent vectors on M, that is, a smooth map $X : M \to TM$ such that $X(p) \in T_pM$. From the previous formula follows that a vector field defines a derivation of $C^\infty(M)$; take above $v = X(p)$ at each point $p \in M$ and the right- hand-side defines a smooth function on M and the operation satisfies the Leibnitz' rule.

We denote by $D^1(M)$ the space of vector fields on M. As we have seen, a vector field gives a linear map $X : C^\infty(M) \to C^\infty(M)$ obeying the Leibnitz' rule. Conversely, one can prove that any derivation of the algebra $C^\infty(M)$ is represented by a vector field.

One can develope an algebraic approach to manifold theory. In that the commutative algebra $\mathcal{A} = C^\infty(M)$ plays a central role. Points in M correspond to *maximal ideals* in the algebra $\mathcal{A}$. Namely, any point p defines the ideal $I_p \subset \mathcal{A}$ consisting of all functions which vanish at the point p.

The action of a vector field on functions is given in terms of local coordinates $x_1, \ldots, x_n$ as follows. If $v = X(p)$ is represented by a curve γ then

$$(X \cdot f)(p) = \frac{d}{dt} f(\gamma(t))_{t=0} = \sum_k \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}(t=0) \equiv \sum_k X_k(x) \frac{\partial f}{\partial x_k}.$$

Thus a vector field is locally represented by the vector valued function $(X_1(x), \ldots, X_n(x))$.

In addition of being a real vector space, $D^1(M)$ is a *left module* for the algebra $C^\infty(M)$. This means that we have a linear left multiplication $(f, X) \mapsto fX$. The value of fX at a point p is simply the vector $f(p)X(p) \in T_pM$.

As we have seen, in a coordinate system x_i a vector field defines a derivation with local representation $X = \sum_k X_k \frac{\partial}{\partial x_k}$. In a second coordinate system x'_k we have a representation $X = \sum X'_k \frac{\partial}{\partial x'_k}$. Using the chain rule for differentiation we obtain the coordinate transformation rule

$$X'_k(x') = \sum_j \frac{\partial x'_k}{\partial x_j} X_j(x),$$

for $x'_k = x'_k(x_1, \ldots, x_n)$.

We shall denote $\partial^k = \frac{\partial}{\partial x_k}$ and we use Einstein's summation convention over repeated indices,

Let $X, Y \in D^1(M)$. We define a new derivation of $C^\infty(M)$, the commutator $[X, Y] \in D^1(M)$, by

$$[X, Y]f = X(Yf) - Y(Xf).$$

We prove that this is indeed a derivation of $C^\infty(M)$.

$$\begin{aligned}
[X,Y](fg) &= X(Y(fg)) - Y(X(fg)) = X(fYg + gYf) - Y(fXg + gXf) \\
&= (Xf)(Yg) + fX(Yg) + (Xg)(Yf) + gX(Yf) - (Yf)(Xg) - fY(Xg) \\
&- (Yg)(Xf) - gY(Xf) = f[X,Y]g + g[X,Y]f.
\end{aligned}$$

Writing $X = X_k \partial^k$ and $Y = Y_k \partial^k$ we obtain the coordinate expression

$$[X, Y]_k = X_j \partial^j Y_k - Y_j \partial^j X_k.$$

Thus we may view $D^1(M)$ simply as the space of first order linear partial differential operators on M with the ordinary commutator of differential operators. The commutator $[X, Y]$ is also called the *Lie bracket* on $D^1(M)$. It has the basic properties

(1) $[X, Y]$ is linear in both arguments
(2) $[X, Y] = -[Y, X]$
(3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$

The last property is called the Jacobi identity. A vector space equipped with a Lie bracket satisfying the properties above is called a *Lie algebra.*

Other examples of Lie algebras:

Example 1 Any vector space with the bracket $[X, Y] = 0$.

Example 2 The angular momentum Lie algebra with basis L_1, L_2, L_3 and nonzero commutators $[L_1, L_2] = L_3$ + cyclically permuted relations.

Example 3 The space of $n \times n$ matrices with the usual commutator of matrices, $[X, Y] = XY - YX$.

Exercise Check the relations

$$[X, fY] = f[X, Y] + (Xf)Y, \text{ and } [fX, Y] = f[X, Y] - (Yf)X$$

for $X, Y \in D^1(M)$ and $f \in C^\infty(M)$.

Let $f : M \to N$ be a diffeomorphism and $X \in D^1(M)$. We can define a vector field $Y = f_* X$ on N by setting $Y(q) = T_p f \cdot X(p)$ for $q = f(p)$. In terms of local coordinates,

$$Y = Y_k \frac{\partial}{\partial y_k} = X_j \frac{\partial y_k}{\partial x_j} \frac{\partial}{\partial y_k}.$$

In the case $M = N$ this gives back the coordinate transformation rule for vector fields.

Let $X \in D^1(M)$. Consider the differential equation

$$X(\gamma(t)) = \frac{d}{dt}\gamma(t)$$

for a smooth curve γ. In terms of local coordinates this equation is written as

$$X_k(x(t)) = \frac{d}{dt} x_k(t), k = 1, 2, \ldots, n.$$

By the theory of ordinary differential equations this system has locally, at a neighborhood of an initial point $p = \gamma(0)$, a unique solution. However, in general the solution does not need to extend to $-\infty < t < +\infty$ except in the case when M is a compact manifold. The (local) solution γ is called *an integral curve* of X through the point p.

The integral curves for a vector field X define a (local) *flow* on the manifold M. This is a (local) map

$$f : \mathbb{R} \times M \to M$$

given by $f(t, p) = \gamma(t)$ where γ is the integral curve through p. We have the identity

$$f(t + s, p) = f(t, f(s, p)), \tag{1}$$

which follows from the uniqueness of the local solution to the first order ordinary differential equation. In coordinates,

$$\frac{d}{dt} f_k(t, f(s,x)) = X_k(f(t, f(s,x)))$$

and

$$\frac{d}{dt} f_k(t+s, x) = X_k(f(t+s, x)).$$

Thus both sides of (1) obey the same differential equation. Since the initial conditions are the same, at $t = 0$ both sides are equal to $f(0, f(s,x)) = f(s,x)$, the solutions must agree.

Denoting $f_t(p) = f(t,p)$, observe that the map $\mathbb{R} \to \mathrm{Diff}_{loc}(M)$, $t \mapsto f_t$, is a homomorphism,

$$f_t \circ f_s = f_{t+s}.$$

Thus we have a *one parameter group of (local) transformations* f_t on M. In the case when M is compact we actually have globally globally defined transformations on M.

Example Let $X(r,\phi) = (-r\sin\phi, r\cos\phi)$ be a vector field on $M = \mathbb{R}^2$. The integral curves are solutions of the equations

$$x'(t) = -r(t)\sin\phi(t)$$
$$y'(t) = r(t)\cos\phi(t)$$

and the solutions are easily seen to be given by $(x(t), y(t)) = (r_0\cos(\phi+\phi_0), r_0\sin(\phi+\phi_0))$, where the initial condition is specified by the constants ϕ_0, r_0. The one parameter group of tranformations generated by the vector field X is then the group of rotations in the plane.

Further reading: M. Nakahara: *Geometry, Topology and Physics,* Institute of Physics Publ. (1990), sections 5.1 - 5.3 . S. S. Chern, W.H. Chen, K.S. Lam: *Lectures on Differential Geometry,* World Scientific Publ. (1999), Chapter 1.

CHAPTER 2: DIFFERENTIAL FORMS

2.1 Multilinear forms

Let V be a vector space, $\dim V = n < \infty$, over the field $K = \mathbb{R}$ or $K = \mathbb{C}$. The *dual space* V^* consists of all linear functions $f : V \to K$ and it is a vector space under the usual addition and scalar multiplication of functions. If $\{e_1, \dots, e_n\}$ is a basis of V then we can define a basis $\{f_1, \dots, f_n\}$ of V^* by $f_i(e_j) = \delta_{ij}$. We denote $\Omega^1(V) = V^*$.

Next we define $\Omega^2(V) = V^* \wedge V^*$ as the space of antisymmetric functions $f : V \times V \to K$ which are linear in both arguments. $\Omega^2(V)$ is a vector space of dimension $n(n-1)/2$. A basis is given by the functions f_{ij} defined by

$$f_{ij}(e_k, e_l) = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$$

with $1 \le i < j \le n$. We set $f_{ji} = -f_{ij}$. A general element of $\Omega^2(V)$ is then a linear combination $f = a_{ij} f_{ij}$ with $a_{ij} = -a_{ji}$, that is, elements in $\Omega^2(V)$ are antisymmetric tensors on V^*.

If $f, g \in \Omega^1(V)$ then $f \wedge g \in \Omega^2(V)$ with $(f \wedge g)(x, y) = f(x)g(y) - f(y)g(x)$. In particular, $f_{ij} = f_i \wedge f_j$. The wedge product is antisymmetric, $f \wedge g = -g \wedge f$.

Example When $V = \mathbb{R}^3$ the wedge product is simply the cross product of vectors. We can identify $\Omega^2(\mathbb{R}^3)$ as the space $\mathbb{R}^3$ by using the standard basis: The elements in an antisymmetric tensor (a_{ij}) are parametrized by a vector (a_{23}, a_{31}, a_{12}) and then $x \wedge y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$.

In general, $\Omega^k(V)$ denotes the space of alternating multilinear forms $f : V \times V \times \cdots \times V \to K$ (k arguments). Alternating means that the sign of the function is reversed when a pair of arguments is transposed. In other words, a permutation σ of the arguments can be compensated by a multiplication by $\epsilon(\sigma)$ where $\epsilon(\sigma) = \pm 1$ is the parity of the permutation,

$$f(x_{\sigma(1)}, ..., x_{\sigma(n)}) = \epsilon(\sigma) f(x_1, ..., x_n).$$

A basis in $\Omega^k(V)$ is given by the multilinear forms $f_{i_1 i_2 \dots i_k}$ defined by

$$f_{i_1 \dots i_k}(x^{(1)}, \dots, x^{(k)}) = \det\,(x^{(j)}_{i_m}).$$

By the symmetry properties of the determinant the right-hand-side indeed defines an alternating form. The dimension of $\Omega^k(V)$ is equal to the binomial factor $\binom{n}{k}$. In particular, dim $\Omega^n(V) = 1$ and $\Omega^k(V) = 0$ for $k > n$. We set $\Omega^0(V) = K$ and

$$\Omega(V) = \Omega^0(V) \oplus \Omega^1(V) \oplus \cdots \oplus \Omega^n(V).$$

The dimension of the direct sum is

$$\dim \Omega(V) = \sum_k \binom{n}{k} = 2^n.$$

We generalize the wedge product to a product

$$\Omega^j(V) \times \Omega^k(V) \to \Omega^{j+k}(V)$$

by the formula

$$(f \wedge g)(x^{(1)}, \ldots, x^{(j+k)}) = \frac{1}{j!}\frac{1}{k!} \sum_{\sigma \in S_{j+k}} \epsilon(\sigma) f(x^{(\sigma(1))}, \ldots, x^{(\sigma(j))}) g(x^{(\sigma(j+1))}, \ldots, x^{(\sigma(j+k))}),$$

where S_n is the group of permutations of integers $1, 2, \ldots, n$.

Exercise 1 Show that $f \wedge g$ is alternating.

Exercise 2 Prove that $f \wedge g = (-1)^{jk} g \wedge f$.

Exercise 3 Prove that $f \wedge (g \wedge h) = (f \wedge g) \wedge h$.

Note that the basis $f_{i_1 i_2 \ldots i_k}$ defined above is obtained from the f_i's,

$$f_{i_1 i_2 \ldots i_k} = f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_k}.$$

2.2 Differential forms

Let M be a smooth manifold of dimension n. A *differential form* of degree k on M is a smooth distribution $\omega_x \in \Omega^k(T_x M)$ of alternating forms in the tangent spaces. We denote by $\Omega^k(M)$ the set of differential forms of degree k. Smoothness of the distribution $x \mapsto \omega_x$ is defined in terms of local coordinates $x_1, \ldots, x_n$. Recall that each coordinate x_i defines a local vector field $\partial^i = \frac{\partial}{\partial x_i}$, interpreted as a derivation of the algebra $C^\infty(M)$. A tangent vector at a point x is uniquely written as $v = v_i \partial^i$. For this reason ω is given in terms of the coordinate functions

$$\omega^{i_1 \ldots i_k}(x) = \omega_x(\partial^{i_1}, \ldots, \partial^{i_k}).$$

Smoothness of ω means that the coordinate functions $\omega^{i_1\cdots i_k}(x)$ are smooth functions of the coordinates x_i.

Locally, a basis for $\Omega^1(M)$ is given by the differential 1-forms dx_i defined by

$$dx_i(\partial^j) = \delta_{ij}.$$

A basis for k-forms is given by

$$dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k} \text{ with } 1 \leq i_1 < i_2 < \cdots < i_k \leq n.$$

In the given coordinate chart we have then

$$\omega = \frac{1}{k!}\omega^{i_1 i_2 \ldots i_k}(x) dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}.$$

The wedge pruduct of forms $\omega \in \Omega^j(M)$ and $\theta \in \Omega^k(M)$ is a form in $\Omega^{j+k}(M)$ defined pointwise as $(\omega \wedge \theta)_x = \omega_x \wedge \theta_x$. The product is associative and

$$\omega \wedge \theta = (-1)^{jk}\theta \wedge \omega.$$

The *exterior derivative* of $\omega \in \Omega^k(M)$ is defined in terms of local coordinates as an element $d\omega$ of $\Omega^{k+1}(M)$,

$$d\left(\omega^{i_1\ldots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}\right) = \partial^j \omega^{i_1 \ldots i_k} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \tag{1}$$

We define $\Omega^0(M) = C^\infty(M)$ and then

$$df = \partial^j f dx_j$$

for a smooth function f. We must also check that the definition of $d\omega$ in terms of local coordinates is compatible with coordinate tranformations. Since $\partial'^k = \frac{\partial}{\partial x'_k} = \frac{\partial x_j}{\partial x'_k}\partial^j$ by the chain rule, we obtain

$$\omega'^{i_1\ldots i_k} = \omega(\partial'^{i_1}, \ldots, \partial'^{i_k}) = \omega(\partial^{j_1}, \ldots, \partial^{j_k})\frac{\partial x_{j_1}}{\partial x'_{i_1}} \cdots \frac{\partial x_{j_k}}{\partial x'_{i_k}}.$$

In other words,

$$\omega^{i_1\ldots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k} = \omega'^{j_1\ldots j_k} dx'_{j_1} \wedge \cdots \wedge dx'_{j_k},$$

for

$$\omega'^{i_1\ldots i_k} = \omega^{j_1\cdots j_k}\frac{\partial x_{j_1}}{\partial x'_{i_1}} \cdots \frac{\partial x_{j_k}}{\partial x'_{i_k}}.$$

When the exterior differentiation is applied to the right-hand-side we obtain an expression similar to (1) but the coordinates x_i replaced by x'_i; But the exterior derivative of the right-hand-side is equal to

$$\partial'^j \omega'^{i_1 \dots i_k} dx'_j \wedge dx'_{i_1} \wedge \dots \wedge dx'_{i_k} = \frac{\partial x_l}{\partial x'_j} \partial^l \left(\frac{\partial x_{j_1}}{\partial x'_{i_1}} \dots \frac{\partial x_{j_k}}{\partial x'_{i_k}} \omega^{j_1 \dots j_k} \right) dx'_j \wedge dx'_{i_1} \wedge \dots \wedge dx'_{i_k}$$
$$= \partial^l \omega^{i_1 \dots i_k} dx_l \wedge dx_{i_1} \dots \wedge dx_{i_k} + \omega^{j_1 \dots j_k} \frac{\partial^2 x_{j_1}}{\partial x'_j x'_{i_1}} \frac{\partial x_{j_2}}{\partial x'_{i_2}} \dots \frac{\partial x_{j_k}}{\partial x'_{i_k}} dx'_j \wedge dx'_{i_1} \dots \wedge dx'_{i_k} + \dots.$$

Using the antisymmetry of the wedge products $dx_j \wedge dx_{i_p}$ and the symmetry of the second derivatives we observe that all the terms involving second derivatives are identically zero and therefore only the first term remains, giving the exterior derivative of ω in the x_i coordinates.

To remember the transformation rule for differential forms it is sufficient to keep in mind the transformation for 1-forms,

$$dx'_i = \frac{\partial x'_i}{\partial x_j} dx_j,$$

since the higher order forms are exterior products of the basic 1-forms and smooth functions.

Theorem. $d^2 = 0$.

Proof.

$$d^2(\omega^{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}) = d\left(\partial^j \omega^{i_1 \dots i_k} dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}\right)$$
$$= \partial^l \partial^j \omega^{i_1 \dots i_k} dx_l \wedge dx_j \wedge \dots dx_{i_k}.$$

Again, using the symmetry of second derivatives and antisymmetry of the wedge product $dx_l \wedge dx_j$ we see that all terms on the right vanish and thus $d^2\omega = 0$.

Note that $d\omega = 0$ for $\omega \in \Omega^0(M)$ implies that ω is a constant function in each connected component of M. Set $\Omega(M) = \Omega^0(M) \oplus \Omega^1(M) \oplus \dots \Omega^n(M)$ with $n = \dim M$.

Theorem. *Let $\omega \in \Omega^p(M)$ and $\theta \in \Omega^q(M)$. Then $d(\omega \wedge \theta) = d\omega \wedge \theta + (-1)^p \omega \wedge d\theta$.*

Proof. Set $\omega = \omega^{i_1\dots i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p}$ and $\phi = \phi^{j_1\dots j_q} dx_{j_1} \wedge \dots \wedge dx_{j_q}$. Then

$$\begin{aligned}
\omega \wedge \phi &= \omega^{i_1\dots i_p}\phi^{j_1\dots j_q} dx_{i_1} \wedge \dots dx_{i_p} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_q} \\
d(\omega \wedge \phi) &= \phi^{j_1\dots j_q}\partial^k \omega^{i_1\dots i_p} dx_k \wedge dx_{i_1} \wedge \dots dx_{i_p} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_q} \\
&\quad + \omega^{i_1\dots i_p}\partial^k \phi^{j_1\dots j_q} dx_k \wedge dx_{i_1} \wedge \dots dx_{j_q} \\
&= d\omega \wedge \phi + (-1)^p \omega^{i_1\dots i_p}\partial^k \phi^{j_1\dots j_q} dx_{i_1} \dots dx_{i_p} \wedge dx_k \wedge dx_{j_1} \wedge \dots dx_{j_q} \\
&= d\omega \wedge \phi + (-1)^p \omega \wedge d\phi,
\end{aligned}$$

where we have used the alternating property of the wedge product, $dx_{i_1} \wedge \dots dx_{i_p} \wedge dx_k = (-1)^p dx_k \wedge dx_{i_1} \wedge \dots dx_{i_p}$.

There is alternative way to think about differential forms. Let $X \in D^1(M)$ and $\omega \in \Omega^1(M)$. We can define a smooth function on M by $f(x) = \omega_x(X(x))$ by the natural pairing of tangent vectors $X(x)$ and the elements $\omega_x \in T^*_x M$ in the dual. Thus a 1-form is a map from $D^1(M)$ to $C^\infty(M)$. This map is linear, moreover $\omega(gX) = g\omega(X)$ for any smooth function g.

In a similar way, any $\omega \in \Omega^k(M)$ can be thought of as a multilinear function $\omega : D^1(M) \times D^1(M) \times \dots D^1(M) \to C^\infty(M)$ by

$$\omega(X_1, X_2, \dots, X_k)(x) = \omega_x(X_1(x), \dots, X_k(x)).$$

By the definition of a differential form, this map is alternating.

There is converse result which we state without proof: Any alternating map $D^1(M) \times \dots \times D^1(M) \to C^\infty(M)$ which is $C^\infty(M)$ linear in each variable, is uniquely represented by a differential form.

Let $f \in \Omega^0(M)$ and $X \in D^1(M)$. Then

$$df(X) = (\partial^k f dx_k)(X_j \partial^j) = X_j \partial^k f dx_k(\partial^j) = X_j \partial^j f = X \cdot f.$$

Next let $\omega \in \Omega^1(M)$ and $X, Y \in D^1(M)$. Now

$$\begin{aligned}
(d\omega)(X,Y) &= (\partial^j \omega^i dx_j \wedge dx_i)(X,Y) = (\partial^j \omega^i)(dx_j(X)dx_i(Y) - dx_j(Y)dx_i(X)) \\
&= (\partial^j \omega^i)(X_j Y_i - Y_j X_i) = X \cdot \omega(Y) - Y \cdot \omega(X) - \omega^i(X_j \partial^j Y_i - Y_j \partial^j X_i) \\
&= X \cdot \omega(Y) - Y \cdot \omega(X) - \omega([X,Y]).
\end{aligned}$$

Exercise Generalize the formula above to differential forms of degree k,

$$(d\omega)(X_1,\dots,X_{k+1}) = \sum_i (-1)^{i-1} X_i\cdot\omega(X_1,\dots,\hat{X}_i,\dots,X_{k+1}) + \sum_{i<j}(-1)^{i+j}\omega([X_i,X_j],X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_{k+1}),$$

where the hat over X_i means that this variable is deleted from the sequence.

Next we define the *interior product* of a vector field X and a k form ω.

$$(i_X\omega)(X_1,\dots,X_{k-1}) = \omega(X,X_1,\dots,X_{k-1}).$$

Note that $i_X\omega$ is a form of degree $k-1$. For example, when $\omega\in\Omega^1(M)$, then $i_X\omega$ is simply the function $\omega(X)$. If $\omega = \frac{1}{2}\omega^{ij}dx_i\wedge dx_j$ then

$$(i_X\omega)(Y) = \omega(X,Y) = \frac{1}{2}\omega^{ij}(X_iY_j - X_jY_i) = Y_i\omega^{ji}X_j,$$

by $\omega^{ij} = -\omega^{ji}$. Thus $i_X\omega = \theta$ with $\theta^i = X_j\omega^{ji}$. In general, for $\omega = \frac{1}{k!}\omega^{i_1\dots i_k}dx_{i_1}\wedge\dots\wedge dx_{i_k}$ we have

$$(i_X\omega)^{i_1\dots i_{k-1}} = X_j\omega^{ji_1\dots i_{k-1}}.$$

For any smooth map $f: M\to N$ we define *the pull-back* operator $f^*:\Omega^k(N)\to\Omega^k(M)$ by

$$(f^*\omega)_x(v_1,\dots,v_k) = \omega_{f(x)}(T_xf\cdot v_1,\dots,T_xf\cdot v_k)$$

for $v_1,\dots,v_k\in T_xM$. In terms of local coordinates x_i on M and y_i on N we have

$$(f^*\omega)^{i_1\dots i_k}(x) = \frac{\partial y_{j_1}}{\partial x_{i_1}}\dots\frac{\partial y_{j_k}}{\partial x_{i_k}}\omega^{j_1\dots j_k}(y).$$

In the case when $M=N$ this gives us again the coordinate transformation rule of differential forms.

Exercise Show that $f^*(d\omega) = d(f^*\omega)$ and $f^*(\omega\wedge\theta) = (f^*\omega)\wedge(f^*\theta)$ for all differential forms ω,θ and a smooth map f.

The pull-back of a form $h\in C^\infty(M) = \Omega^0(M)$ is simply the composed function $f^*h = h\circ f$.

Finally we define *the Lie derivative* of a $k-$ form ω in the direction of a vector field X as the $k-$form $\mathcal{L}_X\omega$,

$$(\mathcal{L}_X\omega)(X_1,\dots,X_k) = X\cdot\omega(X_1,\dots,X_k) - \sum_{i=1}^k \omega(X_1,\dots,[X,X_i],\dots,X_k).$$

In terms of local coordinates,

$$(\mathcal{L}_X\omega)^{i_1\dots i_k} = X\cdot\omega^{i_1\dots i_k} + \sum_{\alpha=1}^{k}(\partial^{i_\alpha}X_j)\omega^{i_1\dots i_{\alpha-1}ji_{\alpha+1}\dots i_k}.$$

Exercise Prove the relation $\mathcal{L}_X = d\circ i_X + i_X\circ d$.

2.3 Maxwell's equations and differential forms

We arrange the Cartesian coordinates of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ as an antisymmetric 4×4 matrix,

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & +cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix}.$$

We label the rows and columns by $\mu,\nu = 0,1,2,3$ and we set $F = \frac{1}{2}F^{\mu\nu}dx_\mu\wedge dx_\nu$.

Let ϕ be an electric scalar potential and $\mathbf{A}$ a magnetic vector potential. Then

$$\mathbf{E} = -\nabla\phi - \partial^0\mathbf{A} \text{ and } \mathbf{B} = \nabla\times\mathbf{A},$$

where $\partial^0 = \frac{1}{c}\frac{\partial}{\partial t}$ but we shall work in units with speed of light $c=1$. Define the 1-form $A = A^\mu dx_\mu$ with $A^0 = \phi$ and $A^i = c\mathbf{A}^i$. Thus we may write

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

that is, $F = dA$.

Since $d^2 = 0$ we have automatically $dF = 0$. Written in electric and magnetic field components this gives the second set of Maxwell's equations,

$$\nabla\cdot\mathbf{B} = 0$$
$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}.$$

In order to obtain a differential form expression for the first set of Maxwell's equations,

$$\nabla\cdot\mathbf{E} = \rho/\epsilon_0$$
$$\nabla\times\mathbf{B} = \mu_0\left(\epsilon_0\frac{\partial\mathbf{E}}{\partial t} + \mathbf{j}\right),$$

with $\mu_0\epsilon_0 = 1/c^2$, we must first fix a metric tensor $(g_{\mu\nu})$ in space-time; this could be just the Minkowski metric $\mathrm{diag}(1,-1,-1,-1)$ but we may take any (pseudo) Riemannian metric. Note that the second set of Maxwell's equations is intrinsic to any smooth manifold, it does not depend on the choice of metric.

We shall denote $g^{ij} = g(\partial^i, \partial^j)$ for a (pseudo) Riemannian metric $g_x : T_xM \times T_xM \to \mathbb{R}$. Recall from the relativity course that by definition the matrix (g^{ij}) is symmetric and nondegenerate. The matrix elements of the inverse matrix are denoted by (g_{ij}), so $g_{ij}g^{jk} = g^{ij}g_{jk} = \delta_{ik}$.

We define an *orientation* on a manifold M of dimension n. The manifold is oriented if we have a complete system of local coordinates such that all coordinate transformations $x'_i = x'_i(x_1, \dots, x_n)$ satisfy the condition $\det(\frac{\partial x'_i}{\partial x_j}) > 0$.

Not every manifold can be oriented. The standard spheres S^n inherit an orientation from $\mathbb{R}^{n+1}$. The orientation on $\mathbb{R}^n$ is given by the ordered set of Cartesian coordinates $(x_1, x_2, \dots, x_n)$. A coordinate system $(y_1, \dots, y_n)$ on the embedded unit sphere in $\mathbb{R}^{n+1}$ is then oriented if the vectors $(\partial_y^1, \dots, \partial_y^n, v)$ are compatible with the orientation of $\mathbb{R}^{n+1}$. Here v is the outward unit normal vector field on the sphere and compatibility means that the matrix relating the given tangent vectors to the standard basis has positive determinant. On the other hand, *the real projective plane* $P\mathbb{R}^2 = S^2/\mathbb{Z}_2 = (\mathbb{R}^3 - \{0\})/\mathbb{R}_+$, consisting of lines through the origin in $\mathbb{R}^3$, has no orientation.

A metric defines a preferred n-form on an oriented manifold, called the volume form. In terms of local oriented coordinates it is defined as

$$vol_M = |\det(g^{ij})|^{1/2} dx_1 \wedge dx_2 \wedge \dots dx_n.$$

Let x'_i be another set of oriented coordinates. Then

$$dx'_1 \wedge dx'_2 \dots \wedge dx'_n = \det\left(\frac{\partial x'_i}{\partial x_j}\right) dx_1 \wedge dx_2 \dots \wedge dx_n.$$

On the other hand, $g'^{ij} = \frac{\partial x_k}{\partial x'_i}\frac{\partial x_l}{\partial x'_j} g^{kl}$, and this gives

$$\det(g'^{ij}) = \det(g^{ij}) \left(\det\left(\frac{\partial x_i}{\partial x'_j}\right)\right)^2.$$

This implies that

$$|\det(g'^{ij})|^{1/2} dx'_1 \wedge dx'_2 \dots \wedge dx'_n = |\det(g^{ij})|^{1/2} dx_1 \wedge dx_2 \dots \wedge dx_n$$

and thus the definition of vol_M is compatible with change of oriented coordinates.

Note that the orientation is really important: If the determinant of the coordinate transformation is negative then the volume would change the sign.

A metric defines also *a duality operation* $* : \Omega^k(M) \to \Omega^{n-k}(M)$ on differential forms. In local coordinates,

$$*\omega = \theta^{i_1 i_2 \dots i_{n-k}} dx_{i_1} \wedge dx_{i_2} \dots dx_{i_{n-k}} \text{ with}$$
$$\theta^{i_1 \dots i_{n-k}} = |\det(g^{ij})|^{-1/2} \frac{1}{k!} \epsilon^{i_1 \dots i_{n-k}}{}_{j_1 \dots j_k} \omega^{j_1 \dots j_k},$$

where $\epsilon_{i_1 \dots i_n}$ is the totally antisymmetric tensor with $\epsilon_{12 \dots n} = +1$ and the raising of indices is done with the help of the metric tensor as in general relativity.

Example Let $M = \mathbb{R}^4$ and g^{ij} the Minkowski metric. Then $vol_M = dx_0 \wedge dx_1 \wedge dx_2 \wedge dx_3$. The dual of the Maxwell 2-form $F = \frac{1}{2} F^{\mu\nu} dx_\mu \wedge dx_\nu$ is given by

$$(*F)^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu}{}_{\alpha\beta} F^{\alpha\beta},$$

so $(*F)^{12} = F^{03}$, and cyclic permutations of 123, and $(*F)^{01} = -F^{23}$, and cyclic permutations of 123. That is, the magnetic components of the dual are equal to $(-1)\times$ the electric components of the original and the electric components of the dual are equal to the magnetic components of the original field.

The complete set of Maxwell's equations can now be written as

$$d * F = J$$
$$dF = 0,$$

where the 3-form J is defined as $\frac{1}{3!} \epsilon^{\mu\alpha\beta\gamma} J_\mu dx_\alpha \wedge dx_\beta \wedge dx_\gamma$ with $J_0 = \rho/\epsilon_0$ and $J^k = c\mu_0 j^k$. Here ρ is the charge density and $\mathbf{j}$ is the electric current density.

2.4 de Rham cohomology

Recall that $d : \Omega^k(M) \to \Omega^{k+1}(M)$ is a linear map with $d^2 = 0$. We set $B^k(M) = d(\Omega^{k-1}(M)) \subset \Omega^k(M)$ and $Z^k(M) = \ker d = \{\omega \in \Omega^k | d\omega = 0\} \subset \Omega^k(M)$. These are linear subspaces with the property $B^k(M) \subset Z^k(M)$, because of $d^2 = 0$. Elements of Z^k are called *closed forms* and elements of B^k are *exact forms.* We set

$$H^k(M) = Z^k(M)/B^k(M), \text{ with } k = 0, 1, 2, \dots$$

where $H^0(M) \equiv Z^0(M)$. Note that $H^k(M) = 0$ for $k > n$ since $\Omega^k(M) = 0$ for $k > n$. The vector spaces $H^k(M)$ are called the *de Rham cohomology groups* of M. In case when M is compact, one can prove that $\dim H^k(M) < \infty$ for all k.

Example $M = \mathbb{R}^3$. Since $df = 0$ for $f \in C^\infty(M) = \Omega^0(M)$ means that f is a constant function, we get $H^0(\mathbb{R}^3) = \mathbb{R}$. If $\omega = \omega^i dx_i$ satisfies $d\omega = 0$ then the vector field $(\omega^1, \omega^2, \omega^3)$ has zero curl, and we know vector analysis that there is a scalar potential f such that $\nabla f = \omega$, in other words, $df = \omega$. Thus $B^1 = Z^1$ and so $H^1(\mathbb{R}) = 0$. If $\omega = \frac{1}{2}\omega^{ij} dx_i \wedge dx_j$ is a 2-form with $d\omega = 0$ then $\operatorname{div} \omega = 0$ with $\omega = (\omega_{23}, \omega_{31}, \omega_{12})$. This implies that there is a vector potential $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \omega$, or in other words, $dA = \omega$, $A = A^i dx_i$. Again, $Z^2 = B^2$ and $H^2(\mathbb{R}^3) = 0$. In the same vein one can show that $H^3(\mathbb{R}^3) = 0$.

Poincare's lemma. *Let $M \subset \mathbb{R}^n$ be a star shaped open set. This means that there is a point $z \in M$ such that the line $tx + (1-t)z, 0 \le t \le 1$, belongs to M for any $x \in M$. Let ω be a closed $k-$form on M, $k > 0$. Then there exists a $(k-1)$-form θ such that $d\theta = \omega$.*

Proof. Define

$$\theta^{i_1 \dots i_{k-1}}(x) = k \int_0^1 t^{k-1}(x_j - z_j)\omega^{j i_1 i_2 \dots i_{k-1}}(tx + (1-t)z)dt.$$

We claim that $d\theta = \omega$. Now

$$\begin{aligned} d\theta = k &\int_0^1 t^{k-1} \omega^{j i_1 \dots i_{k-1}}(tx + (1-t)z) dx_j \wedge dx_{i_1} \wedge \dots dx_{i_{k-1}} dt \\ + k &\int t^k (x_j - z_j) \partial^l \omega^{j i_1 \dots i_{k-1}}(tx + (1-t)z) dx_l \wedge dx_{i_1} \wedge \dots dx_{i_{k-1}} dt. \end{aligned} \tag{1}$$

The equation $d\omega = 0$ gives

$$\partial^l \omega^{j i_1 \dots i_{k-1}} \pm \text{ cyclic permutations of } l j i_1 \dots i_{k-1} \ = 0,$$

where the signs are given by the parity of the cyclic permutation. From this equation one can reduce, by setting the contraction $i_{\partial^j} d\omega$ equal to zero,

$$k \partial^l \omega^{j i_1 \dots i_{k-1}} dx_l \wedge dx_{i_1} \wedge \dots \wedge dx_{i_{k-1}} = \partial^j \omega^{l i_1 \dots i_{k-1}} dx_l \wedge \dots \wedge dx_{i_{k-1}}.$$

Note that in local coordinates

$$i_{\partial^j} d\omega^* + d i_{\partial^j} \omega^* = \mathcal{L}_{\partial^j} \omega^* = \partial^j \omega^*.$$

Inserting this to the second term I_2 on the right-hand-side of (1) we obtain

$$
\begin{aligned}
I_2 &= \int_0^1 (x_j - z_j) t^k \partial^j \omega^{li_1\dots i_{k-1}}(tx + (1-t)z) dx_l \wedge dx_{i_1} \cdots \wedge dx_{i_{k-1}} dt \\
&= \int_0^1 t^k \frac{d}{dt} \omega^{li_1\dots i_{k-1}}(tx + (1-t)z) dx_l \wedge dx_{i_1} \dots dx_{i_{k-1}} dt \\
&= -k \int_0^1 t^{k-1} \omega^{li_1\dots i_{k-1}}(tx + (1-t)z) dx_l \wedge dx_{i_1} \wedge \dots dx_{i_{k-1}} dt \\
&+ \omega^{li_1\dots i_{k-1}} dx_l \wedge dx_{i_1} \cdots \wedge dx_{i_{k-1}}.
\end{aligned}
$$

Insertion to (1) completes the proof of $d\theta = \omega$.

The above result extends (by a use of coordinates) to the case when M is a contractible subset of a smooth manifold: contractibility means that the identity map on M can be smoothly deformed to a constant map $x \mapsto X_0$ on M. Let $f_t : M \to M$ be such a contraction, $f_0(x) = x_0$ and $f_1(x) = x$, $0 \le t \le 1$. Then one can repeat the proof but with the straight lines $t \mapsto tx + (1-t)z$ replaced by $t \mapsto f_t(x), z = x_0$, see Nakahara, section 6.3, for details.

Example 1 Let $M = S^1$. The 1-form $d\phi$ is closed but $d\phi \ne df$ for any smooth function f on S^1. Note that the polar angle ϕ is not a function on S^1 since it is nonperiodic. Any 1-form on S^1 is given as $f(\phi)d\phi$ for some periodic function f of ϕ. The integral of f over the interval $[0, 2\pi]$ gives a real number λ_f. If $\lambda_f = \lambda_g$ for any two functions f, g then we can write $f - g = h'$ for a periodic function h, that is, $fd\phi - gd\phi = dh$. It follows that the cohomology classes $[f] \in H^1(S^1)$ are parametrized by the integral λ_f and so $H^1(S^1) = \mathbb{R}$.

Example 2 On the unit sphere S^2 the area form is given as $\omega = \sin\theta d\theta \wedge d\phi$ in spherical coordinates. Locally, $\omega = d(-\cos\theta d\phi) = d(-\phi \sin\theta d\theta)$. Note that the first expression becomes singular at the poles $\theta = 0, \pi$ whereas the second is nonperiodic in the coordinate ϕ. One can prove that $H^2(S^2) = \mathbb{R}$ and that the cohomology classes are parametrized by the integral of the 2-form over S^2. In general, it is known that $H^k(S^n) = 0$ for $1 \le k \le n-1$ and that $H^0(S^n) = \mathbb{R} = H^n(S^n)$.

Example 3 $H^1(S^1 \times S^1) = \mathbb{R}^2$ (basis of 1-forms $d\phi_1, d\phi_2$) and $H^2(S^1 \times S^1) = \mathbb{R}$, basis $d\phi_1 \wedge d\phi_2$.

2.5 Integration of differential forms

Let M be a smooth oriented manifold of dimension n. We fix an atlas of coordinate neighborhoods compatible with the given orientation. Let $x_1, \dots, x_n$ be local coordinates on an open set $U \subset M$. Asssume that $f \in C^\infty(M)$ is such that $f(x) = 0$ when x is outside of a compact subset K of U. Then $\omega = f(x) dx_1 \wedge dx_2 \cdots \wedge dx_n$ is a $n-$ form on M. We **define** the integral

$$\int \omega = \int f(x) dx_1 dx_2 \dots dx_n,$$

as the ordinary Riemann integral in $\mathbb{R}^n$.

Let us assume that we have a *locally finite* atlas (U_α, ϕ_α). This means that for any $x \in M$ there is an open neighborhood V of x such that V intersects only a finite number of the sets U_α. A space which has a locally finite cover is said to be *paracompact.* In fact, any finite-dimensional manifold is paracompact according to our definition, page 1 in these notes. A locally finite atlas has a subordinate *partition of unity.* That is, there is a family of smooth nonnegative functions $\rho_\alpha : M \to \mathbb{R}$ such that

(1) $\mathrm{supp}\rho_\alpha \subset U_\alpha$
(2) $\sum_\alpha \rho_\alpha(x) = 1$ for all $x \in M$.

The support $\mathrm{supp}f$ of a function f is defined as a closure of the set of points x for which $f(x) \neq 0$.

Let $\omega \in \Omega^n(M)$. we define

$$\int_M \omega = \sum_\alpha \int \rho_\alpha \omega,$$

and we apply the previous definition to each term on the right-hand-side. The integral converges always when M is compact.

Exercise Show that the above definition does not depend on the choice of the partition of unity or of the locally finite atlas.

Next we want to define the integral of a form $\omega \in \Omega^k(M)$ over a parametrized $k-$ surface for arbitrary $0 \leq k \leq n$.

A *standard k-simplex* in $\mathbb{R}^k$ is the subset

$$\sigma_k = \{(x_1, \dots, x_k) \in \mathbb{R}^k | \sum x_i \leq 1, x_j \geq 0\}.$$

So σ_0 is just a point, σ_1 is the unit interval, σ_2 is a triangle, etc. *A singular k-simplex* is any smooth map $s_k : \sigma_k \to M$. A *k-chain* is a formal linear combination $\sum a_\alpha s_{k,\alpha}$, with $a_\alpha \in \mathbb{R}$ and each $s_{k,\alpha}$ is a singular $k-$ simplex. We define an affine map $F_k^i : \sigma_{k-1} \to \sigma_k$ where $i = 0, 1, \ldots, k$. Note that the subset of points in σ_k with the coordinate $x_i = 0$ can be naturally identified as a $k-1$ simplex σ_{k-1} for $1 \leq i \leq k$. This defines the map (as an identity map) for $i = 1, 2 \ldots, k$. The remaining map F_k^0 sends the $(k-1)$-simplex σ_{k-1} to the face of the k-simplex which is not parallel to any of the coordinate axes. The map is completely fixed by requiring it to be affine and *compatible with the orientations,* and such that the origin of σ_{k-1} is mapped to the vertex of σ_k lying on the first coordinate axes, and the vertex of σ_{k-1} lying on the i :th coordinate axes is mapped to the vertex of σ_k on the $(i+1)$:th coordinate axes, for $i = 1, 2, \ldots, k-1$. See the picture below.

macro:-¿simplex.pdf

The *boundary* of a singular k-simplex $s_k : \sigma_k \to M$ is the singular k-chain defined as

$$\partial s_k = \sum_{i=0}^{k} (-1)^i s_k \circ F_k^i.$$

we extend the definition, by linearity, to the space C_k of singular k-chains, $\partial : C_k \to C_{k-1}$.

Theorem. $\partial^2 = 0$.

Proof. We first observe that

$$F_k^i \circ F_{k-1}^j = F_k^j F_{k-1}^{i-1}, \text{ for } j < i.$$

Let $s = \sum_\alpha a_\alpha s_{k,\alpha} \in C_k$. Then

$$\begin{aligned}
\partial^2 s &= \partial \sum_\alpha a_\alpha \sum_{i=0}^{k} (-1)^i s_{k,\alpha} \circ F_k^i \\
&= \sum_\alpha a_\alpha \sum_{i=0}^{k} (-1)^i \sum_{j=0}^{k-1} s_{k,\alpha} \circ F_k^i \circ F_{k-1}^j (-1)^j \\
&= \sum_\alpha a_\alpha \left(\sum_{0 \le i \le j \le k-1} (-1)^{i+j} s_{k,\alpha} F_k^i \circ F_{k-1}^j \right. \\
&\left. + \sum_{0 \le j < i \le k} (-1)^{i+j} s_{k,\alpha} F_k^i \circ F_{k-1}^j \right) \\
&= \sum_\alpha a_\alpha \left(\sum_{0 \le i \le j \le k-1} (-1)^{i+j} s_{k,\alpha} F_k^i \circ F_{k-1}^j \right. \\
&\left. + \sum_{0 \le j < i \le k} (-1)^{i+j} s_{k,\alpha} F_k^j \circ F_{k-1}^{i-1} \right).
\end{aligned}$$

Relabel $i \mapsto j, j \mapsto i-1$ in the first term of right-hand-side of the last equality; then the terms cancel.

A *cycle* is a singular chain s such that $\partial s = 0$. A *boundary* is a singular chain b such that $b = \partial s$ for some singular chain s. Denote by Z_k the space of k-cycles and by B_k the space of k-boundaries. Finally, the *singular k-homology group* is the space

$$H_k(M) = H_k(M, \mathbb{R}) = Z_k(M)/B_k(M).$$

Sometimes one considers also the homology group $H_k(M, \mathbb{Z})$ which is defined as the real homology group but one restricts to integral linear combinations of the singular k-simplexes.

Exercise Show that $H_0(M)$ is isomorphic with $\mathbb{R}^k$, where k is the number of path connected components of M.

The homology groups H_k of contractible manifolds vanish for $k > 0$, so in particular $H_k(\mathbb{R}^n) = 0$ for $k > 0$. On the other hand, $H_n(S^n) = \mathbb{R}$ but $H_k(S^n) = 0$ for $0 < k < n$.

We define the integral of a k-form over a singular k-chain $s = \sum_\alpha a_\alpha s_{k,\alpha}$,

$$\int_s \omega = \sum_\alpha a_\alpha \int_{\sigma_k} s_{k,\alpha}^* \omega.$$

Each of the integral on the right is an ordinary Riemann integral of a smooth function defined in the standard simplex $\sigma_k \subset \mathbb{R}^k$, after writing each of the pull-back forms as $f(x)dx_1 \wedge \ldots dx_k$.

Theorem. *(Stokes' theorem)*

$$\int_s d\omega = \int_{\partial s} \omega$$

for any $\omega \in \Omega^{k-1}(M)$ and for any singular k-chain s.

Proof. By linearity, it is sufficient to give the proof for a single singular k-simplex s_k. But in this case a typical term in $s_k^*\omega$ can be written as

$$s_k^*\omega = \sum_{j=1}^{k} b_j(x)dx_1 \wedge \ldots d\hat{x}_j \wedge \ldots dx_k(-1)^{j-1}$$

for some smooth functions b_j. Then

$$d(s_k^*\omega) = s_k^*(d\omega) = \sum (\partial^j b_j)dx_1 \wedge \cdots \wedge dx_k = f(x)dx_1 \wedge \cdots \wedge dx_k.$$

We can now apply to familiar Gauss' theorem for vector fields in $\mathbb{R}^k$,

$$\int_{\sigma_k} \partial^j b_j dx_1 \ldots dx_k = \int_{\partial\sigma_k} \mathbf{b} \cdot \mathbf{n} dS,$$

where $\mathbf{n}$ is the outward normal vector field on σ_k and dS is the Euclidean area measure on the surface $\partial\sigma_k$ of the k-simplex. But the right-hand-side of the equation is equal to the integral $\int_{\partial\sigma_k} s_k^*\omega$, which proves the theorem.

We have a pairing $H_k(M) \times H^k(M) \to \mathbb{R}$ which is given as

$$< [s], [\omega] >= \int_s \omega.$$

Because of Stokes' theorem the right-hand-side does not depend on particular representatives of the (co)homology classes, i.e., if $s - s'$ is a boundary and $\omega - \omega'$ is a coboundary then

$$\int_s \omega = \int_{s'} \omega'.$$

For compact oriented manifolds one can prove that the pairing is nondegenerate, i.e., if $< [s], [\omega] >= 0$ for all $[\omega]$ (resp. for all $[s]$) then $[s] = 0$ (resp. $[\omega] = 0$).

There is a more refined version of Stokes' theorem (which we are not going to prove). This uses the idea of a *closed submanifold with boundary.* A manifold M with boundary is defined using the half space $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n | x_n \geq 0\}$ as a model instead of the vector space $\mathbb{R}^n$. That is, M should be equipped with a cover by open sets U and coordinate maps $\phi : U \to \mathbb{R}^n_+$ which are homeomorphism to open subsets of the half space. The coordinate transformations $\phi \circ \psi^{-1}$ are again required to be smooth in their domain of definition. Note that the derivative in the x_n direction at the boundary points $x_n = 0$ is only defined to the positive direction.

Example The closed unit ball $B^n = \{x \in \mathbb{R}^n | ||x|| \leq 1\}$ is a manifold with boundary. The set of boundary points is the manifold S^{n-1}.

Let $N \subset M$ be an oriented manifold with boundary (dimension n) embedded in M. Its boundary ∂N is a manifold of dimension $n-1$. Let $\omega \in \Omega^{n-1}(M)$. Then one can prove

$$\int_N d\omega = \int_{\partial N} \omega.$$

Note that the integral on the left is an integral of a n-form over a manifold of dimension n (and this we have already defined) and on the right we have an integral of a $(n-1)$-form over a manifold of dimension $n-1$.

Additional reading: Nakahara: 5.4, 5.5, and Chapter 6
Chern, Chen, and Lam: Chapters 2 and 3

CHAPTER 3: RIEMANN GEOMETRY

3.1 Affine connection

According to the definition, a vector field $X \in D^1(M)$ determines a derivation of the algebra of smooth real valued functions on M. This action is linear in X such that $(fX)g = f(Xg)$ for any pair f, g of smooth functions. Next, we want to define an action of X on $D^1(M)$ itself, which has similar properties. Let $\nabla_X : D^1(M) \to D^1(M)$ for any $X \in D^1(M)$ be an operator satisfying the following conditions:

(1) The map $Y \mapsto \nabla_X Y$ is real linear in Y for any fixed X,
(2) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$ for any vector fields X, Y, Z and any smooth real valued functions f, g, and
(3) $\nabla_X(fY) = f\nabla_X Y + Y(X \cdot f)$ for any vector fields X, Y and any smooth real valued function f.

An operator ∇ satisfying these conditions is called an *affine connection* on the manifold M.

Example 1 Let $M = \mathbb{R}^n$ and define

$$\nabla_X Y = (X \cdot Y^j)\frac{\partial}{\partial x^j}.$$

Then, ∇ is an affine connection.

Warning! The above example needs a modification when applied to an arbitrary manifold M. The difficulty is that the right-hand side depends on the choice of local coordinates and it does not transform like a true vector. If we transform to coordinates $y^j = y^j(x^1, \ldots, x^n)$, then in the new coordinates

$$Y'^j(y) = \frac{\partial y^j}{\partial x^i} Y^i(x)$$

and therefore

$$(X \cdot Y'^j)\partial'_j = (X \cdot Y^i)\frac{\partial y^j}{\partial x^i}\partial'_j + Y_i\left(X \cdot \frac{\partial yj}{\partial x^i}\right)\partial'_j.$$

The coordinates of the first term on the right-hand side are equal to $\frac{\partial y^j}{\partial x^i}(\nabla_X Y)^i$, but for any non-linear coordinate transformation we also have a second inhomogeneous term.

Choosing local coordinates, the difference

$$H^i(X,Y) = (\nabla_X Y)^i - X \cdot Y^i$$

is linear in both arguments in the extended sense

$$H^i(fX, gY) = fgH^i(X,Y),$$

for any smooth functions f and g. For this reason, we can write

$$H^i(X,Y) = \Gamma^i_{jk} X^j Y^k.$$

Here $\Gamma^i_{jk} = \Gamma^i_{jk}(x)$ are smooth (local) functions on M. Once again,

$$(\nabla_X Y)^i = X \cdot Y^i + \Gamma^i_{jk} X^j Y^k.$$

The functions Γ^i_{jk} are called the *Christoffel symbols* of the affine connection ∇. Let us look what happens to the Christoffel symbols in a coordinate transformation $y = y(x)$. Let us denote by ∇_i the *covariant derivative* $\nabla_{\frac{\partial}{\partial x^i}}$. Then,

$$\nabla_i \partial_j = \Gamma^k_{ij} \partial_k.$$

Denoting $\partial'_i = \frac{\partial}{\partial y^i}$, we get

$$\begin{aligned}
\nabla'_i \partial'_j = \Gamma'^k_{ij} \partial'_k &= \frac{\partial x^a}{\partial y^i} \nabla_a \left(\frac{\partial x^b}{\partial y^j} \partial_b \right) \\
&= \frac{\partial x^a}{\partial y^i} \left[\frac{\partial x^b}{\partial y^j} \nabla_a \partial_b + \partial_a \left(\frac{\partial x^b}{\partial y^j} \right) \partial_b \right] \\
&= \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^j} \Gamma^c_{ab} \partial_c + \frac{\partial^2 x^b}{\partial y^i \partial y^j} \partial_b.
\end{aligned}$$

Transforming back to the x coordinates on the left-hand side, we finally get

$$\Gamma'^k_{ij}(y) = \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^j} \frac{\partial y^k}{\partial x^c} \Gamma^c_{ab}(x) + \frac{\partial y^k}{\partial x^c} \frac{\partial^2 x^c}{\partial y^i \partial y^j}. \tag{3.1.1}$$

Note that in linear coordinate transformations, the inhomogeneous term containing second derivatives vanishes and the Christoffel symbols transform like components of a third rank tensor.

Exercise 1 We define the Christoffel symbols on the unit sphere, using spherical coordinates (θ, ϕ). When $\theta \neq 0, \pi$, we set

$$\Gamma^\theta_{\phi\phi} = -\frac{1}{2} \sin 2\theta, \ \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta,$$

and all the other Γ's are equal to zero. Show that the apparent singularity at $\theta = 0, \pi$ can be removed by a better choice of coordinates at the poles of the sphere. Thus, the above affine connection extends to the whole S^2.

3.2 Parallel Transport

The tangent vectors at a point $p \in M$ form a vector space T_pM. Thus, tangent vectors at the *same point* can be added. However, at different points p and q, there is in general no way to compare the tangent vectors $u \in T_pM$ and $v \in T_qM$. In particular, the sum $u + v$ is ill-defined. An affine connection gives a method to relate tangent vectors at p to tangent vectors at q, provided that we have fixed some smooth curve $\gamma(t)$ starting from p and ending at q.

A curve γ defines a distribution of tangent vectors along the curve by

$$X(s) = \dot{x}^i(s)\partial_i.$$

We have chosen a local coordinate system x^i. Thus, $X(s) \in T_{\gamma(s)}M$. Consider the system of first order ordinary differential equations given by

$$\dot{Y}^i(s) + \Gamma^i_{kj}(x(s))\dot{x}^k(s)Y^j(s) = 0, \;\; i = 1, 2, \dots, n. \tag{3.2.1}$$

Here $Y(s)$ is an unknown vector field along the curve $x(s)$.

Exercise 2 Show that the set of equations above is coordinate independent in the sense that if the equations are valid in one coordinate system, then they are also valid in any other coordinate system.

A vector field Y along the curve $x(s)$ satisfying the differential equation is called a *parallel vector field*. The existence and uniqueness theorem in the theory of first order differential equations gives the following fundamental theorem in geometry:

Theorem 3.2.2. *Given a tangent vector $v \in T_pM$ at the initial point $p = \gamma(s_0)$ of a smooth curve $\gamma(s)$ then there is a unique parallel vector field $Y(s)$ along $\gamma(s)$ satisfying the initial condition $Y(s_0) = v$.*

Definition 3.2.3. *A curve $\gamma(s)$ is a geodesic if its tangent vectors $\dot{\gamma}(s)$ at each point are parallel.*

Thus, the statement $\gamma(s)$ is a geodesic means that the coordinate functions $x_i(s)$ satisfy

$$\ddot{x}^i(s) + \Gamma^i_{jk}(x(s))\dot{x}^j(s)\dot{x}^k(s) = 0.$$

This condition is a second order ordinary differential equation for the coordinate functions. We can use existence and uniqueness results from the theory of differential equations:

Theorem 3.2.4. *For given point $p \in M$ and a tangent vector $u \in T_pM$ there is, in some open neighborhood of p, a unique geodesic $\gamma(s)$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = u$.*

Example 2 Let $M = S^2$ and let Γ be the affine connection in Exercise 1. Then, the coordinates $\theta(s)$ and $\phi(s)$ of a geodesic satisfy

$$\begin{aligned} \ddot{\theta}(s) - \frac{1}{2}\sin 2\theta(s)\,\dot{\phi}(s)\dot{\phi}(s) &= 0, \\ \ddot{\phi}(s) + 2\cot\theta(s)\,\dot{\phi}(s)\dot{\theta}(s) &= 0. \end{aligned} \tag{3.2.5}$$

Find the general solution to the geodesic equations. The solutions are great circles on the sphere. For example, $\theta = \alpha s + \beta$ and $\phi = \text{const.}$

Let ∇ be a connection on M and $\gamma(s)$ a curve connecting points $p = \gamma(s_1)$ and $q = \gamma(s_2)$. We define the *parallel transport* from the point p to the point q along the curve γ as a linear map

$$\hat{\gamma} : T_pM \to T_qM.$$

The map is given as follows: Let $u \in T_pM$ and let $X(s)$ be a parallel vector field along γ such that $X(s_1) = u$. We set $\hat{\gamma}(u) = X(s_2)$. The map is linear, because the differential equation

$$\dot{X}^i(s) + \Gamma^i_{kj}\dot{x}_k(s)X^j(s) = 0 \tag{3.2.6}$$

is linear in X^i and therefore the solution depends linearly on the initial condition u.

Example 3 If $M = \mathbb{R}^n$ and $\Gamma^i_{jk} = 0$, then the parallel transport $\hat{\gamma}$ is the identity map $u \mapsto u$ for any curve γ.

Example 4 Let M and Γ be as in Example 2. Let $(\theta,\phi)=(\alpha s+\beta,\phi_0)$. Now, the parallel transport is determined by the equations

$$\dot{X}_\theta = 0,$$

$$\dot{X}_\phi + \cot\theta\cdot\dot{\theta}X_\phi = \dot{X}_\phi + X_\phi\alpha\cot(\alpha s+\beta) = 0. \tag{3.2.7}$$

This set has the solution $X_\theta=$ const. and $X_\phi = const.\cdot(\sin(\alpha s+\beta))^{-1}$. If u is the tangent vector $(1,1)$ at the point $(\theta,\phi)=(\pi/4,0)$, then the parallel transported vector v at $(\theta,\phi)=(\pi/2,0)$ is $(1,1/\sqrt{2})$.

3.3 Torsion and Curvature

Given an affine connection ∇ on a manifold M we can define a third rank tensor field $T=(T^k_{ij})$ as follows. Any pair of vector fields X and Y gives another vector field

$$T(X,Y)=\nabla_X Y-\nabla_Y X-[X,Y]. \tag{3.3.1}$$

The dependence on X and Y is linear, after choosing local coordinates, we may write

$$T(X,Y)^i = X^jY^kT^i_{jk},$$

which defines the components T^i_{jk} of the tensor. Note that $T(X,Y)$ is linear in the extended sense,

$$T(fX,Y)=T(X,fY)=fT(X,Y),\ T(X,Y+Z)=T(X,Y)+T(X,Z)$$

for any real function f. Note further that $T(X,Y)=-T(Y,X)$. Since

$$T(\partial_i,\partial_j)^k=\Gamma^k_{ij}-\Gamma^k_{ji}, \tag{3.3.2}$$

we see that T is precisely the antisymmetric part (in the lower indices) of the Christoffel symbols.

From the above equation and the transformation formula for the Christoffel symbols follows that the components of the *torsion* T really transform like tensor components in coordinate transformations,

$$T'^i_{jk}(y)=\frac{\partial y^i}{\partial x^p}\frac{\partial x^\ell}{\partial y^j}\frac{\partial x^m}{\partial y^k}T^p_{\ell m}(x).$$

Next, we define the *curvature tensor* R. For a triple X, Y, Z of vector fields, we can define a vector field

$$R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z. \tag{3.3.3}$$

In local coordinates,

$$R(\partial_i, \partial_j)\partial_k = R^m_{kij}\partial_m.$$

From the definition of the curvature we get

$$\begin{aligned} R^m_{kij}\partial_m &= \nabla_i\nabla_j\partial_k - \nabla_j\nabla_i\partial_k \\ &= \nabla_i(\Gamma^m_{jk}\partial_m) - \nabla_j(\Gamma^m_{ik}\partial_m) \\ &= \partial_i\Gamma^m_{jk}\partial_m + \Gamma^m_{jk}\Gamma^p_{im}\partial_p - \partial_j\Gamma^m_{ik}\partial_m - \Gamma^m_{ik}\Gamma^p_{jm}\partial_p, \end{aligned}$$

$$R^m_{kij} = \partial_i\Gamma^m_{jk} - \partial_j\Gamma^m_{ik} + \Gamma^p_{jk}\Gamma^m_{ip} - \Gamma^p_{ik}\Gamma^m_{jp}. \tag{3.3.4}$$

For fixed i and j, we may think of $R^{\bullet}_{\bullet ij}$ as a real $n \times n$ matrix. With this notation,

$$R^{\bullet}_{\bullet ij} = \partial_i\Gamma^{\bullet}_{j\bullet} - \partial_j\Gamma^{\bullet}_{i\bullet} + [\Gamma^{\bullet}_{i\bullet}, \Gamma^{\bullet}_{j\bullet}] = \left[\partial_i + \Gamma^{\bullet}_{i\bullet}, \partial_j + \Gamma^{\bullet}_{j\bullet}\right]. \tag{3.3.5}$$

The curvature is antisymmetric in i and j,

$$R^m_{kij} = -R^m_{kji}.$$

One checks by a direct computation that in a coordinate transformation $y = y(x)$,

$$R'^m_{kij}(y) = \frac{\partial y^m}{\partial x^q}\frac{\partial x^r}{\partial y^k}\frac{\partial x^s}{\partial y^i}\frac{\partial x^p}{\partial y^j}R^q_{rsp}(x).$$

Thus, R^m_{kij} is really a 4th rank tensor in contrast to the Christoffel symbols Γ^k_{ij}, which transform inhomogeneously in coordinate transformations.

Exercise 3 Check that

$$T'^k_{ij}(y) = \frac{\partial y^k}{\partial x^m}\frac{\partial x^r}{\partial y^i}\frac{\partial x^s}{\partial y^j}T^m_{rs}(x)$$

in a coordinate transformation $y = y(x)$.

Assume that the torsion T vanishes, we deduce from (3.3.4) the *first Bianchi identity*

$$R^m_{kij} + R^m_{jki} + R^m_{ijk} = 0 \tag{3.3.6}$$

for all indices. This can also be written as

$$R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0 \tag{3.3.7}$$

for all vector fields X, Y, Z. This identity is in general not true when $T \neq 0$.

Another important tensor in general relativity is the *Ricci tensor*

$$R_{ij} = R^k_{ikj}.$$

Exercise 4 Show that R_{ij} transforms like a second rank tensor in coordinate transformations.

The curvature is related to the parallel transport in the following way. Consider a *very small* parallelogram with edges at $x, x+\delta x, x+\delta x+\delta y, x+\delta y$. According to the differential equation determining a parallel transport, a tangent vector Y at x when parallel transported to the point $x+\delta x$ becomes approximately (in given local coordinates)

$$Y^i(x+\delta x) = Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta x^j.$$

At the next point $x+\delta x+\delta y$, we get

$$\begin{aligned} Y^i(x+\delta x+\delta y) &= Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta xj \\ &- \Gamma^i_{jk}(x+\delta x)[Y^k(x) - \Gamma^k_{\ell m}(x)Y^m(x)\delta x^j]\delta y^\ell \\ &= Y^i(x) - \Gamma^i_{jk}(x)Y^k(x)\delta x^j - \Gamma^i_{jk}(x)Y^k(x)\delta y^j \\ &- \partial_m\Gamma^i_{jk}(x)\delta x^m\delta y^j Y^k(x) \\ &+ \Gamma^i_{jk}(x)\Gamma^k_{\ell m}(x)Y^m(x)\delta x^\ell\delta y^j. \end{aligned}$$

In the same way, we can compute the parallel transport of Y from x to $x+\delta y$ and further to $x+\delta y+\delta x$. The parallel transport around the parallelogram is then obtained as a combination of the right-hand side of the above formula and the latter transport (note the direction of motion!); the result is

$$\begin{aligned} \delta Y^i &= R^i_{kmj}(x)Y^k(x)\delta x^m\delta y^j \\ &= \frac{1}{2}R^i_{kmj}(x)Y^k(x)(\delta x^m\delta y^j - \delta x^j\delta y^m). \end{aligned} \tag{3.3.8}$$

Thus, the parallel transport around the small parallelogram is proportional to the curvature at x and the area of the parallelogram.

Example 5

We compute the curvature tensor of the unit sphere S^2. Since there are only two independent coordinates, all the non-zero components of R are given by the tensor $R^i_j = R^i_{j\theta\phi} = -R^i_{j\phi\theta}$, where $i, j = \theta, \phi$. Looking at the table (Exercise 1) of the Christoffel symbols, we get

$$R^\theta_\phi = \sin^2\theta,\ R^\phi_\theta = -1,$$

and the other components $= 0$.

The *second Bianchi identity*

$$\partial_i R^\bullet_{\bullet jk} + [\Gamma^\bullet_{i\bullet}, R^\bullet_{\bullet jk}] + \partial_j R^\bullet_{\bullet ki} + [\Gamma^\bullet_{j\bullet}, R^\bullet_{\bullet ki}] + \partial_k R^\bullet_{\bullet ij} + [\Gamma^\bullet_{k\bullet}, R^\bullet_{\bullet ij}] = 0 \tag{3.3.9}$$

follows from the formula (3.3.5) and the Jacobi identity for matrices,

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

3.4 Metric and Pseudo-Metric

In order to define distances and inner products between tangent vectors on a manifold M, we have to define a *metric.* A *Riemannian metric* is an inner product defined in each of the tangent spaces. That is, for each $p \in M$, we have a non-degenerate bilinear mapping

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

which is symmetric, $g_p(u,v) = g_p(v,u)$ for all tangent vectors $u, v \in T_pM$, and $g_p(u,u) > 0$ for all $u \neq 0$, and it depends smoothly on the coordinates of the point p. Choosing local coordinates x_i and writing the tangent vectors in the coordinate basis, $u = u^i\partial_i$, we can write a symmetric bilinear mapping as a second rank symmetric tensor,

$$g_p(u,v) = g_{ij}u^iv^j.$$

Non-degenerate means that $\det(g_{ij}) \neq 0$. Since (g_{ij}) is symmetric, it can be diagonalized. Positivity of the inner product means then that all eigenvalues of g are positive.

In relativity, we need a generalization of the Riemann metric to a *pseudo-Riemannian metric*. In this generalization, we shall drop the requirement that the inner product is positive. In particular, we want to include the Minkowski space metric $(\eta_{\mu\nu})$, which has *signature* $(1,3)$, it has one positive eigenvalue $(=1)$ and three negative eigenvalues $(=-1)$.

A metric (or a pseudo-metric) can be used to define distances. If $\gamma(s)$ is a parametrized curve such that its tangent vector at each point on the curve has non-negative length, then we define the length of the curve (between the parameter values a and b) as

$$\ell(\gamma) = \int_a^b \sqrt{g_{\gamma(s)}(\dot\gamma(s), \dot\gamma(s))}\, ds.$$

The extremal curves $\gamma(t)$ for the functional $\ell(\gamma)$ are the geodetic curves for a certain connection (the Levi-Civita connection, see the discussion below and Theorem 3.4.4). Recall the *Euler-Lagrange* variational equations: Let $x(t) = (x^1(t), x^2(t), \ldots, x^n(t))$ be a vector valued function of a real variable t and

$$S(x(\cdot)) = \int_a^b L(x(t), x'(t), x''(t), \ldots) dt$$

where L is some (differentiable) function of the derivatives $x, x', x'', \ldots$. Then the derivative of S in the direction $\delta x(t)$ of a variation of the curve $x(t)$ is

$$\delta S = \sum_i \int_a^b \delta x^i(t) \left\{ \frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial x^{i\prime}} + (\frac{d}{dt})^2 \frac{\partial L}{\partial x^{i\prime\prime}} - \ldots \right\} dt,$$

where we have used partial integration in the variable t in order to factor out δx under the integral sign. The requirement that the variation δS vanishes in arbitrary directions δx in the path space is then equivalent to the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial x^{i\prime}} + (\frac{d}{dt})^2 \frac{\partial L}{\partial x^{i\prime\prime}} - \cdots = 0,$$

where $i = 1, 2, \ldots n$.

Example 6 If $M = \mathbb{R}^n$, then we can define a constant metric $g_{ij} = \delta_{ij}$. This is the standard Euclidean metric. In general in $\mathbb{R}^n$, a Riemannian metric is given by smooth real functions $g_{ij}(x) = g_{ji}(x)$ such that the matrix $(g_{ij}(x))$ is strictly positive for all $x \in \mathbb{R}^n$.

Example 7 If $M \subset \mathbb{R}^n$ is any smooth surface in the Euclidean space, then we can define a metric g as follows. Let $u, v \in T_pM$ be a pair of tangent vectors to the surface at the point p. The tangent vectors are also vectors in $\mathbb{R}^n$, thus we may compute the scalar product $u \cdot v$. We set $g_p(u,v) = u \cdot v$. From the fact that the Euclidean metric is positive definite follows at once that g is a positive symmetric form.

Example 8 Let $M = S^2 \subset \mathbb{R}^3$. We compute the metric g on M, as defined in Example 7, in terms of the spherical coordinates θ and ϕ. The spherical coordinates are related to the standard coordinates by

$$\partial_\theta = \cos\theta\,\cos\phi \frac{\partial}{\partial x} + \cos\theta\,\sin\phi \frac{\partial}{\partial y} - \sin\theta \frac{\partial}{\partial z},$$
$$\partial_\phi = -\sin\theta\,\sin\phi \frac{\partial}{\partial x} + \sin\theta\,\cos\phi \frac{\partial}{\partial y}.$$

From this we obtain the inner products

$$g_{\theta\theta} = g(\partial_\theta, \partial_\theta) = 1,$$
$$g_{\phi\phi} = g(\partial_\phi, \partial_\phi) = \sin^2\theta,$$
$$g_{\theta\phi} = g_{\phi\theta} = 0.$$

For example, the inner product of vectors $(1,2)$ and $(2,-1)$ (in the θ and ϕ coordinates) is $1 \cdot 2 \cdot g_{\theta\theta} + 2 \cdot (-1) \cdot g_{\phi\phi} = 2 - 2\sin^2\theta$, at the point (θ, ϕ). Note that the spherical coordinates are *orthogonal*, the off-diagonal matrix elements of g are equal to zero.

According to the last example, the distance between to points on a sphere along a curve $\gamma(t) = (\theta(t), \phi(t))$ is given by

$$\ell(\gamma) = \int_a^b [\theta'(t)^2 + \sin^2\theta(t)\phi'(t)^2]^{1/2} dt.$$

The Euler-Lagrange equations give then (check this!)

$$\theta''(t) - \frac{1}{2}\phi'^2 \sin(2\theta(t)) = 0$$
$$\frac{d}{dt}[\phi'(t)\sin^2\theta(t)] = 0$$

which agrees with the equations in example 2.

Suppose a (pseudo) metric g is given on a manifold M. From the metric, we can construct a preferred affine connection, called the *Levi-Civita connection.* Its Christoffel symbols (in given local coordinates) are given by the formula

$$\Gamma^k_{ij} = \frac{1}{2} g^{k\ell}(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}), \tag{3.4.1}$$

where g^{ij} are the matrix elements of the inverse matrix g^{-1}.

One should always be extremely careful when trying to define something with the help of local coordinates. It is not *a priori* clear that the locally defined Christoffel symbols in various coordinate systems match together to define a connection on whole manifold M. To investigate the patching problem, we compute what happens in a coordinate transformation $y = y(x)$. Since

$$\frac{\partial}{\partial y^i} = \frac{\partial x^k}{\partial y^i} \frac{\partial}{\partial x^k},$$

we get

$$\begin{aligned} g'_{ij}(y) &= g_y\left(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right) = \frac{\partial x^k}{\partial y^i} \frac{\partial x^\ell}{\partial y^j} g_x\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^\ell}\right) \\ &= g_{k\ell}(x) \frac{\partial x^k}{\partial y^i} \frac{\partial x^\ell}{\partial y^j}. \end{aligned}$$

Inserting this transformation law into the definition of Christoffel symbols, we get

$$\Gamma'^k_{ij}(y) = \frac{\partial y^k}{\partial x^c} \frac{\partial x^a}{\partial y^i} \frac{\partial x^b}{\partial y^j} \Gamma^c_{ab} + \frac{\partial y^k}{\partial x^c} \frac{\partial^2 x^c}{\partial y^i \partial y^j},$$

as expected. Thus, the Christoffel symbols defined in different coordinate systems are compatible and define indeed an affine connection.

Example 9 Since the standard Euclidean metric is constant in the standard coordinates, the Christoffel symbols of the Levi-Civita connection vanish.

Example 10 The Christoffel symbols computed from the metric defined in Example

The Levi-Civita connection has two characteristic properties. The first property is that its torsion $T = 0$, since $\Gamma^k_{ij} - \Gamma^k_{ji} = 0$. The second property is that the parallel transport defined by the Levi-Civita connection is *metric compatible* in the following sense: Let $X(s)$ and $Y(s)$ be a pair of parallel vector fields along a curve $\gamma(s)$. Then,

$$\frac{d}{ds} g_{\gamma(s)}(X(s), Y(s)) = 0,$$

so *the inner products of parallel vector fields are constant along the curve.* This means that the parallel transport $\hat{\gamma} : T_pM \to T_qM$ between the end points of the curve is an *isometry.*

Theorem 3.4.2. *An affine connection ∇ is compatible with a metric g if and only if*

$$Z \cdot g(X,Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

for all vector fields X, Y, Z.

A word about the notation: We write $g(X,Y)$ for the real valued smooth function $p \mapsto g_p(X(p), Y(p))$. Remember that a vector field acts on functions as derivations, so the left-hand side is a well-defined smooth function, too.

Proof. 1) Assume that the condition for g in the theorem is satisfied. Let $X(s)$ and $Y(s)$ be a pair of parallel vector fields along a curve $\gamma(s)$. We shall extend X and Y to vector fields defined in an open neighborhood of the curve. Let Z be some vector field defined in a neighborhood of the curve such that along the curve $Z(\gamma(s)) = \dot\gamma(s)$. Since X and Y are parallel along γ, we have

$$\nabla_Z X = \nabla_Z Y = 0 \text{ on the curve } \gamma.$$

Thus,

$$\frac{d}{ds} g_{\gamma(s)}(X(s), Y(s)) = Z \cdot g(X,Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) = 0 \text{ on } \gamma.$$

2) Assume that ∇ is compatible with g. Let X, Y, Z be a triple of vector fields. Let $p \in M$ and γ any curve through p such that at p, $\dot\gamma(s_1) = Z(p)$. Define vector fields along γ by $X(s) = X(\gamma(s))$ and $Y(s) = Y(\gamma(s))$.

Let $X_1, \dots, X_n$ be an orthonormal basis of tangent vectors at p. We define a set of parallel vector fields $X_i(s)$ along γ such that at $p = \gamma(s_1)$, we have $X_i(s_1) = X_i$. Any pair of vector fields along γ can then be written as

$$X(s) = \alpha^i(s) X_i(s),\ Y(s) = \beta^i(s) X_i(s).$$

Now, we have

$$\begin{aligned} \frac{d}{ds} g_{\gamma(s)}(X(s), Y(s)) &= \frac{d}{ds} \alpha^i(s) \beta^j(s) g_{\gamma(s)}(X_i(s), X_j(s)) \\ &= \frac{d}{ds} \alpha^i(s) \beta^i(s) = \dot\alpha^i(s)\beta^i(s) + \alpha^i(s)\dot\beta^i(s) \\ &= g_{\gamma(s)}(\dot\alpha^i(s) X_i(s), \beta^j(s) X_j(s)) \\ &+ g_{\gamma(s)}(\alpha^i(s) X_i(s), \dot\beta^j(s) X_j(s)) \\ &= g_{\gamma(s)}(\nabla_{\dot\gamma} X(s), Y(s)) + g_{\gamma(s)}(X(s), \nabla_{\dot\gamma} Y(s)). \end{aligned}$$

Applying this formula to the vector field Z at p, $Z(p) = \dot\gamma(s_1)$, we get the condition of the theorem at (the arbitrary point) p.

Theorem 3.4.3. *For a given metric, the Levi-Civita connection is the unique torsion free metric compatible connection.*

Proof.

Use the equation on the previous page for coordinate vector fields $X, Y, Z = \partial_i, \partial_j, \partial_k$ and the symmetry $\Gamma^k_{ij} = \Gamma^k_{ji}$ of a torsion free connection.

Theorem 3.4.4. *A geodesic of the Levi-Civita connection gives an extremal for the path length between two points. If the points are close enough, then the extremal gives the minimum length.*

Proof. Compare the differential equations obtained from the Euler-Lagrange variational principle, applied to curve length, with the differential equations of a geodesic, for the Levi-Civita connection. Note that the Euler-Lagrange equations obtained from the variation of the curve length are the same as obtained from variation of the integral (without square root!)

$$\int_a^b g_{x(t)}(\dot{x}(t), \dot{x}(t))dt.$$

APPENDIX: The Einstein Field Equations

The *Einstein tensor* is defined as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R,$$

where $R = g^{\mu\nu} R_{\mu\nu}$ is the *Ricci scalar*. We assume that the metric $g_{\mu\nu}$ is pseudo-Riemannian of signature $(1, 3)$ (one positive direction and three negative directions). The connection is the Levi-Civita connection computed from the metric and $R_{\mu\nu} = R^{\lambda}_{\mu\lambda\nu}$ is the Ricci tensor.

Exercise 1 Writing $R_{\alpha\beta\mu\nu} = g_{\alpha\lambda} R^{\lambda}_{\beta\mu\nu}$, show that

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}.$$

Show that this implies that $R_{\mu\nu}$ is symmetric.

The Einstein tensor is symmetric. Furthermore, its *covariant divergence* vanishes,

$$\nabla_\mu G^{\mu\nu} = \partial_\mu G^{\mu\nu} + \Gamma^\mu_{\mu\alpha} G^{\alpha\nu} + \Gamma^\nu_{\mu\alpha} G^{\mu\alpha} = 0. \tag{A1}$$

This is seen as follows. First, taking $Z = \partial_\alpha, X = \partial_\mu, Y = \partial_\nu$ in Theorem 3.4.2, we obtain

$$\partial_\alpha g_{\mu\nu} = \Gamma^\beta_{\alpha\mu} g_{\beta\nu} + \Gamma^\beta_{\alpha\nu} g_{\mu\beta} = \Gamma_{\alpha\nu\mu} + \Gamma_{\alpha\mu\nu}. \tag{A2}$$

This can be also written as

$$(\nabla_\alpha g)_{\mu\nu} = 0.$$

For the inverse tensor $g^{\mu\nu} = (g^{-1})_{\mu\nu}$, one gets

$$\partial_\alpha g^{\mu\nu} + \Gamma^\nu_{\alpha\beta} g^{\mu\beta} + \Gamma^\mu_{\alpha\beta} g^{\beta\nu} = 0. \tag{A3}$$

Note the difference in sign for the covariant derivative of the metric tensor and its inverse.

Exercise 2 For any vector field $X = X^\mu \partial_\mu$ the components of the covariant derivatives are $(\nabla_\nu X)^\mu = \partial_\nu X^\mu + \Gamma^\mu_{\nu\alpha} X^\alpha$. Show that the *covariant divergence* is given by

$$(\nabla_\mu X)^\mu = (-\det g)^{-1/2} \partial_\mu ((-\det g)^{1/2} X^\mu).$$

In relativity theory literature, it is a custom to use the abbreviation $X_{\mu;\nu} = (\nabla_\nu X)_\mu$ for the covariant differentiation of vector (and higher order tensor) indices. With this notation, we can write the second Bianchi identity as

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu} = 0. \tag{A4}$$

Contracting the α and μ indices in this identity with the metric tensor, we get

$$g^{\alpha\mu}(R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\nu\lambda;\mu} + R_{\alpha\beta\lambda\mu;\nu}) = 0.$$

By the definition of the Ricci tensor, this can be written as

$$R_{\beta\nu;\lambda} + R^\mu_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu} = 0, \tag{A5}$$

where we have taken into account that the covariant derivative of $g^{\mu\nu}$ vanishes, implying that the multiplication with the components of the metric tensor commutes with covariant differentiation; in particular, index raising and lowering commutes with covariant derivatives. Contracting Eq. (A5) once again with $g^{\beta\nu}$, we get

$$g^{\beta\nu}(R_{\beta\nu;\lambda} + R^{\mu}_{\beta\nu\lambda;\mu} - R_{\beta\lambda;\nu}) = 0. \tag{6-A}$$

Using the results of Exercise 1, we get

$$g^{\beta\nu}R_{\beta\nu\lambda;\mu} = -g^{\beta\nu}g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = -g^{\beta\nu}g^{\alpha\mu}R_{\beta\alpha\nu\lambda;\mu} = -g^{\mu\alpha}R_{\alpha\lambda;\mu} = -R^{\mu}_{\lambda;\mu} \tag{A7}$$

Inserting this into the second term in (A6) we obtain

$$R_{;\lambda} - R^{\mu}_{\lambda;\mu} - R^{\nu}_{\lambda;\nu} = 0. \tag{A8}$$

Note that since R is a scalar, $R_{;\mu} = \partial_\mu R$. An equivalent form of the previous equation is

$$(2R^{\mu}_{\lambda} - \delta^{\mu}_{\lambda}R)_{;\mu} = 0.$$

Raising the index λ and dividing by 2 finally leads to

$$\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)_{;\mu} = 0. \tag{A9}$$

Einstein's gravitational field equations are written simply as

$$G^{\mu\nu} = 8\pi\frac{G}{c^4}T^{\mu\nu}, \tag{A10}$$

where G on the right-hand side (not to be confused with Einstein's tensor!) is Newton's gravitational constant and $T^{\mu\nu}$ is the *stress-energy (energy-momentum) tensor*. It describes the distribution of matter and energy in space-time. For example, the electromagnetic field gives a contribution to $T_{\mu\nu}$ defined by $T^{EM}_{\mu\nu} = \epsilon_0 {F_\mu}^{\lambda}F_{\lambda\nu} + \frac{\epsilon_0}{4}g_{\mu\nu}F^{\lambda\omega}F_{\lambda\omega}$.

Another example is the energy-momentum tensor of a *perfect fluid* .A perfect fluid is characterized by a 4-velocity field u, a scalar density field ρ_0 and a scalar pressure field p. The energy-momentum tensor is defines as

$$T_{\mu\nu} = (\rho_0 + p)u_\mu u_\nu - pg_{\mu\nu}.$$

A special case of this is $p = 0$ which can be viewed as the energy momentum tensor of a flow of noninteracting dust particles. Normally p and ρ_0 are not independent but they are related by the equation of state of the form $p = p(\rho_0, T)$, where T is the temperature. The requirement that the covariant divergence of the energy-momentum tensor vanishes leads to equations of motion for the perfect fluid. In fact, in case of Minkowski space-time and in a certain limit one gets the classical *Navier-Stokes equations* (from $\partial^\mu T_{\mu k} = 0$ for $k = 1, 2, 3$),

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} \right] = -\nabla p$$

and the continuity equation (from $\partial^\mu T_{\mu 0} = 0$),

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0.$$

Here $\rho = \rho_0(1 - \mathbf{u}^2)$.

Let S be some space-like surface with a time-like unit normal vector field n^μ, $n_0 > 0$. Then,

$$\int_S (-\det g)^{1/2} T^{\mu\nu} n_\nu \, d^3x$$

gives the energy and momentum contained in S. Equation (

In order to avoid convergence problems with the infinite integrals, we assume that all energy and momentum are contained in a compact region K in space-time. Consider a surface S, consisting of two space-like components S_1 and S_2 and some surface S_3 'far away' such that T vanishes on S_3. Using Gauss' law and the current conservation, we conclude that the surface integral of $(-\det g)^{1/2} T^\nu_\mu n_\nu$ over S vanishes. In other words,

$$\int_{S_1} (-\det g)^{1/2} T^\nu_\mu n_\nu \, d^3x = \int_{S_2} (-\det g)^{1/2} T^\nu_\mu n_\nu \, d^3x.$$

We have taken into account that, since n is future pointing, one of the normal vector fields on S_1 and S_2 is outward directed and the second inward directed. Equation ( ) tells us that the stress-energy, in the μ-direction, on S_1 is the same as the corresponding quantity on S_2; one could think of S_i as a fixed time slice at time t_i and one obtains the usual law of conservation of energy or momentum.

Often one uses units in which $G = 1$ and $c = 1$ so that one does not need to write explicitly the coefficient G/c^4 in Einstein's equations.

The Newtonian Limit

It is known that the Newtonian gravitational theory is valid for fields, which can produce only velocities much smaller than the velocity of light. Since the components T^{0i} and T^{ij} are related to spatial momenta and T^{00} is related to energy, this condition says that $|T^{00}|$ is much larger than the other components. Because of Einstein's equations, the same is true for the components of the Einstein tensor. Furthermore, we expect that for weak gravitational fields the metric $g^{\mu\nu}$ differs slightly from the Minkowski metric $\eta^{\mu\nu}$,

$$g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}$$

for a small perturbation $h^{\mu\nu}$. Next, we compute the connection, curvature, and finally the Ricci tensor to first order in the perturbation $h^{\mu\nu}$. A straight-forward computation, starting from the definitions of the various tensors, gives $G^{\mu\nu} = -\frac{1}{2}\Box(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h)$, where $h = \eta_{\mu\nu}h^{\mu\nu}$. Thus, Einstein's equations, in this approximation, are linear,

$$-\frac{1}{2}\Box\left(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h\right) = 8\pi\frac{G}{c^4}T^{\mu\nu}.$$

Taking into account the remark in the beginning of this section, only the 00-component is relevant,

$$\Box\left(h^{00} - \frac{1}{2}h\right) = -16\pi\frac{G}{c^2}\rho,$$

where $\rho = T^{00}/c^2$ is the matter density in the rest system of the source. We can also drop the time derivatives (in the system of coordinates, where the source is slowly moving, because $\partial_0 = \frac{1}{c}\partial_t$) and so the only relevant equation becomes

$$\nabla^2\left(h^{00} - \frac{1}{2}h\right) = 16\pi\frac{G}{c^2}\rho.$$

This means that,

$$h^{00} - \frac{1}{2}h = \frac{4}{c^2}\phi,$$

where ϕ is the gravitational potential for the matter distribution ρ. (Compare Eq. (62) with the Newtonian equation $\nabla^2\phi = 4\pi G\rho$, where $\phi = -GM/r$!)

Since all the other components of $h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h$ vanish at this order of approximation, we finally get

$$h^{\mu\mu} = \frac{2}{c^2}\phi = -\frac{2GM}{c^2 r} \text{ (no summation!)}$$

for all $\mu = 0, 1, 2, 3$.

Next, we shall compute the geodesics for the metric $g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}$ in the linear approximation (we neglect higher order terms in $h^{\mu\nu}$). For small velocities, the time component $\dot{x}_0(s)$ of the 4-velocity is much larger than the spatial components. For this reason, we can approximate the geodesic equations of motion as

$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{00}\left(\frac{dx^0}{ds}\right)^2 = 0.$$

In the linear approximation,

$$\Gamma^0_{00} = \partial_0\phi, \;\; \Gamma^i_{00} = \partial_i\phi.$$

Thus, the geodesic equations become

$$\ddot{x}_0 + \partial_0\phi(\dot{x}_0)^2 = 0, \;\; \ddot{x}_i + \partial_i\phi(\dot{x}_0)^2 = 0.$$

In the coordinate system, where the source is at rest, the first equation says that we can choose the time t as the geodesic parameter, $x_0(s) = s = ct$, and then the second equation becomes

$$\ddot{x}_i = -\partial_i\phi.$$

The right-hand side (after multiplication by the mass m of the test particle) is the gravitational force of the source on m, so this equation is just Newton's second law, $m\mathbf{a} = \mathbf{F}$, where $\mathbf{F} = -\nabla\Phi$ and $\Phi = m\phi$.

The Schwarzschild Metric

The basic problem in Newtonian celestial mechanics is to solve the equations of motions outside of a spherically symmetric mass distribution (orbits of the planets around the Sun, orbits of satellites around the Earth). In general relativity the first natural problem is to search for spherically symmetric solutions of Einstein's equations.

Actually, there is a unique 1-parameter family of spherically symmetric solutions, which are *asymptotically flat*, meaning that at large distances from the source the metric tends to the flat Minkowski metric $ds^2 = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2$. This is the content of Birkhoff's theorem (which we are not going to prove). The line element of the metric is given as

$$ds^2 = \left(1 - \frac{2GM}{c^2r}\right)dx_0^2 - \left(1 - \frac{2GM}{c^2r}\right)^{-1}dr^2 - r^2d\Omega^2, \tag{A11}$$

where $d\Omega^2$ is the angular part of the Euclidean metric in $\mathbb{R}^3$, $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$. It is clear from ( ) that for large distances r the metric approaches the Minkowski metric. The line element is called the *Schwarzschild metric.*

When $r > 2GM/c^2$ the Schwarzschild metric is supposed to describe the gravitational field outside of a spherically symmetric star. The other disconnect region $r < 2GM/c^2$ is the *Schwarzschild black hole.* The singularity at $r = 2GM/c^2$ is actually due to a bad choice of coordinates. There is a way to glue the inside solution in a smooth way to the outside solution by a suitable choice of coordinates; the complete discussion of this was first given by Kruskal and Szekeres in 1960. The Kruskal–Szekeres metric is given as follows. The coordinates are denoted by (u, v, θ, ϕ). The latter two are the ordinary spherical coordinates on a unit sphere. The coordinates (u, v) are restricted to the region $L \subset \mathbb{R}^2$ defined by

$$uv < \frac{2GM}{c^2 e}.$$

The metric is then

$$ds^2 = \frac{16\mu^2}{r} e^{(2\mu - r)/2\mu} du dv - r^2 d\Omega^2,$$

where $\mu = MG/c^2$ and r is a function of u, v defined by

$$uv = (2\mu - r)e^{(r-2\mu)/2\mu}.$$

Note that $f(x) = xe^{x/a}$ is monotonically increasing when $x > -a$ (and $f(x) > -a/e$) and therefore $y = f(x)$ has a unique solution x for any $y > -a/e$. We treat u as a kind of universal time; a time-like vector is future directed if its projection to ∂_u is positive. The orientation (needed in integration!) is defined by the ordering (v, u, θ, ϕ) of coordinates. Note that the radial null lines (radial light rays) are given by $du = 0$ or $dv = 0$.

The Kruskal–Szekeres space-time can be divided into four regions: K_1 consists of points $v > 0$, $u < 0$, region K_2 of points $u, v > 0$, in region K_4 we have $u, v < 0$, and finally region K_3 is characterized by $u > 0$, $v < 0$. The boundaries between these regions are non-singular points for the metric. The only singularities are at the boundary $uv = 2\mu/e$.

The region K_1 is equivalent with the outer region of a Schwarzschild space-time. This is seen by performing the coordinate transformation $(v, u, \theta, \phi) \mapsto (t, r, \theta, \phi)$, where $r = r(u,v)$ as above and the Schwarzschild time is $t = 2\mu \ln(-v/u)$. With a similar coordinate transformation the region K_3 is seen to be equivalent with the outer Schwarzschild solution. The region K_2 is equivalent with the Schwarzschild black hole. The equivalence is obtained through the coordinate transformation $(v, u, \theta, \phi) \mapsto (t, r, \theta, \phi)$, where $r = r(u,v)$ is the same as before but now $t = 2\mu \ln(v/u)$.

It is easy to construct smooth time-like curves which go from either K_1 or K_3 to the black hole K_2. However, we shall prove that once an observer falls to the black hole there is no way to go back to the 'normal' regions K_1 and K_3.

Let $x(t)$ be the time-like path of the observer. Then along the path

$$\frac{dr}{dt} = \frac{\partial r}{\partial u}\frac{du}{dt} + \frac{\partial r}{\partial v}\frac{dv}{dt} = \frac{r}{8\mu^2} e^{(r-2\mu)/2\mu} \left[\frac{\partial r}{\partial u} g(\partial_v, x'(t)) + \frac{\partial r}{\partial v} g(\partial_u, x'(t)) \right] < 0,$$

since $x(t)$ is time-like and in K_2 holds $r\frac{\partial r}{\partial u} = -2\mu v e^{(2\mu - r)/2\mu} < 0$ and similarly for the v-coordinate.

The boundary between K_2 and the normal regions is $r = 2\mu$ (*i.e.*, $u = 0$ or $v = 0$). The function $r(x(t))$ was seen to be decreasing, and therefore the path $x(t)$ can never hit the boundary $r = 2\mu$. But the observer entering K_2 has a deplorable future, since it will eventually hit the true singularity $r = 0$, again using the monotonicity of the function $r(x(t))$.

There is also another singularity, the outer boundary of region K_3. But this is of no great concern because it is in the past; no future directed time-like curve can enter that singularity.

CHAPTER 4: PRINCIPAL BUNDLES

4.1 Lie groups

A *Lie group* is a group G which is also a smooth manifold such that the multiplication map $G \times G \to G$, $(a, b) \mapsto ab$, and the inverse $G \to G, a \mapsto a^{-1}$, are smooth.

Actually, one can prove (but this is not easy) that it is sufficient to assume continuety, smoothness comes free. (This was one of the famous problems listed by David Hilbert in his address to the international congress of mathematicians in 1900. The result was proven by A. Gleason, D. Montgomery and L. Zippin in 1952.)

Examples The vector space $\mathbb{R}^n$ is a Lie group. The group multiplication is just the addition of vectors. The set $GL(n, \mathbb{R})$ of invertible real $n \times n$ matrices is a Lie group with respect to the usual matrix multiplication.

Theorem 4.1.1. *Any closed subgroup of a Lie group is a Lie group.*

The proof is complicated. See for example S. Helgason: Differential Geometry, Lie Groups and Symmetric Spaces, section II.2.

The theorem gives an additional set of examples of Lie groups: The group of real orthogonal matrices, the group of complex unitary matrices, the group of invertible upper triangular matrices ....

For a fixed $a \in G$ in a Lie group we can define a pair of smooth maps, the *left translation* $l_a : G \to G, l_a(g) = ag$, and the right translation $r_a : G \to G, r_a(g) = ga$. We say that a vector field $X \in D^1(G)$ is *left (resp. right) invariant* if $(l_a)_* X = X$ (resp. $(r_a)_* X = X$) for all $a \in G$.

Since the left (right) translation is bijective, a left (right) invariant vector field is uniquely determined by giving its value at a single point, at the identity, say. Thus as a vector space, the space of left invariant vector fields can be identified as the tangent space $T_e G$ at the neutral element $e \in G$.

Theorem 4.1.2. *Let X, Y be a pair of left (right) invariant vector fields. Then $[X, Y]$ is left (right) invariant.*

Proof. Denote $f = l_a : G \to G$. Recall that

$$X'^i = (f_* X)^i = \frac{\partial y^i}{\partial x^j} X_j,$$

where we have written the map f in terms of local coordinates as $y = y(x)$. Then

$$\begin{aligned}
[X', Y']^i &= X'^j \partial'_j Y'^i - Y'^j \partial'_j X'^i = X^j \partial_j Y'^i - Y^j \partial_j X'^i \\
&= X^j \partial_j \left(\frac{\partial y^i}{\partial x^k} Y^k\right) - Y^j \partial_j \left(\frac{\partial y^i}{\partial x^k} X^k\right) \\
&= \frac{\partial y^i}{\partial x^k}(X^j \partial_j Y^k - Y^j \partial_j X^k) + \frac{\partial^2 y^i}{\partial x^j \partial x^k}(X^j Y^k - Y^j X^k).
\end{aligned}$$

The second term on the right vanishes since the second derivative is symmetric. Thus we have $[X', Y'] = [X, Y]'$, i.e., $[(l_a)_* X, (l_a)_* Y] = (l_a)_* [X, Y]$. Thus the commutator is left invariant.

It follows that the left invariant vector field form a *Lie algebra.* This Lie algebra is denoted by $Lie(G)$ and it is called the Lie algebra of the Lie group G. Observe that $\dim Lie(G) = \dim T_g G = \dim G$.

Example 1 Let $G = \mathbb{R}^n$. The property that a vector field $X = X_i \partial^i$ is left (right) invariant means simply that the coefficient functions $X_i(x)$ are constants. Thus left invariant vector fields can be immediately identified as vectors $(X_1, \dots, X_n)$ in $\mathbb{R}^n$. Constant vector fields commute, thus $Lie(\mathbb{R}^n)$ is a commutative Lie algebra.

Example 2 Let $G = GL(n, \mathbb{R})$. Let X be a left invariant vector field and $z = X(1) = \frac{d}{dt} e^{tz}|_{t=0}$. Then

$$X(g) = \frac{d}{dt} g e^{tz}|_{t=0} = gz.$$

This implies that

$$(X \cdot f)(g) = \frac{d}{dt} f(g e^{tz})|_{t=0} = \left(\frac{d}{dt} g e^{tz}\right)_{ij}|_{t=0} \frac{\partial}{\partial x_{ij}} f(x)|_{x=g} = g_{ik} z_{kj} \frac{\partial}{\partial g_{ij}} f.$$

When Y is another left invariant vector field with $w = Y(1)$, then

$$\begin{aligned}
[X, Y] &= [g_{ik} z_{kj} \frac{\partial}{\partial g_{ij}}, g_{lm} w_{mp} \frac{\partial}{\partial g_{lp}}] \\
&= g_{ik} z_{kj} w_{jp} \frac{\partial}{\partial g_{ip}} - g_{lm} w_{mp} z_{pj} \frac{\partial}{\partial g_{lp}} \\
&= g_{ik} [z, w]_{kp} \frac{\partial}{\partial g_{ip}}.
\end{aligned}$$

That is, the commutator of the vector fields X, Y is simply given by the commutator $[z, w]$ of the parameter matrices.

Example 3 The group $SO(n) \subset GL(n, \mathbb{R})$ of rotations in $\mathbb{R}^n$. Each antisymmetric matrix L defines a 1-parameter group of rotations by $R(t) = e^{tL}$. The tangent vector of this curve at $t = 0$ is L. We can define a left invariant vector field as above as $X(g) = gL$. The commutator of a pair of antisymmetric matrices is again antisymmetric. The condition that L is antisymmetric is necessary in order that it is tangential to the orthogonal group at the identity: Take a derivative of $R(t)^t R(t) = 1$ at $t = 0$! Thus the Lie algebra of $SO(n)$ consists precisely of all antisymmetric real $n \times n$ matrices. When $n = 2$ we recover the 1-dimensional group of rotations in the plane (the Lie algebra is commutative) and when $n = 3$ we get the 3-dimensional group of rotations in $\mathbb{R}^3$ and its Lie algebra is the angular momentum algebra.

The complex unitary group $U(n)$ has as its Lie algebra the algebra of antihermitean matrices. This is shown by differentiating $R(t)^* R(t) = 1$ at $t = 0$ for $R(t) = e^{tL}$. The Lie algebra of $SU(n)$ is given by antihermitean traceless matrices. Here $SU(n) \subset U(n)$ is the subgroup consisting of matrices of unit determinant.

In the case of a matrix Lie group we have an *exponential mapping* $\exp : Lie(G) \to G$ from the Lie algebra to the corresponding Lie group, which is given through the usual power series expansion $e^X = 1 + X + \frac{1}{2!}X^2 \dots$. This is because the left invariant vector fields are parametrized by the value of the tangent vector at the identity which is equal to the derivative of a 1-parameter group of matrices at the identity. The exponential mapping, which has a central role in Lie group theory, can be generalized to arbitrary Lie groups. If X is a left invariant vector field on the group then, at least locally, there is a unique curve $g(t)$ with $g(0) = 1$ and $g'(t) = X(g(t))$ by the theory of first order differential equations. In fact, it is easy to see that this solution is actually globally defined by a use of group multiplication. Since X is left invariant we have $g'(t) = \ell_{g(t)} \cdot X(1)$, that is, $X(1) = \ell_{g(t)}^{-1} \cdot g'(t)$. The exponential mapping is then defined as

$$e^X = g(1).$$

See S. Helgason, Chapter II, for more details.

Exercise Prove with the help of the chain rule of differentiation that $e^{tX}e^{sX} = e^{(t+s)X}$ in every Lie group, for real t, s and for any left invariant vector field X.

Let $X_1, \dots, X_n$ be a basis of $Lie(G)$. Then

$$[X_i, X_j] = c_{ij}^k X_k$$

for some numerical constants c_{ij}^k, the so-called *structure constants.* Since the Lie bracket is antisymmetric we have $c_{ij}^k = -c_{ji}^k$ and by the Jacobi identity we have

$$c_{ij}^k c_{kl}^m + c_{li}^k c_{kj}^m + c_{jl}^k c_{ki}^m = 0$$

for all i, j, l, m. In terms of the left invariant vector fields X_i, any tangent vector v at $g \in G$ can be written as $v = v^i X_i(g)$. Let us define $\theta^i \in \Omega^1(G)$ as $\theta^i(g)v = v^i$. We compute the exterior derivative $d\theta^i$:

$$\begin{aligned} d\theta^i(g)(X_j, X_k) &= X_j \theta^i(X_k) - X_k \theta^i(X_j) - \theta^i([X_j, X_k]) \\ &= X_j \delta_{ik} - X_k \delta_{ij} - \theta^i(c_{jk}^l X_l) = -c_{jk}^i. \end{aligned}$$

On the other hand,

$$(\theta^i \wedge \theta^j)(X_k, X_l) = \theta^i(X_k)\theta^j(X_l) - \theta^i(X_l)\theta^j(X_k) = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}.$$

Thus we obtain *Cartan's structural equations,*

$$d\theta^i = -\frac{1}{2} c_{kl}^i \theta^k \wedge \theta^l.$$

Denote $X_i \theta^i = g^{-1}dg$. This is a $Lie(G)$-valued 1-form on G. It is tautological at the identity: $(g^{-1}dg)(v) = v$ for $v \in T_1G$. For $\theta = g^{-1}dg$ the structural equations can be written as

$$d\theta + \frac{1}{2}[\theta \wedge \theta] = 0,$$

where $[\theta \wedge \theta] = [X_i, X_j]\theta^i \wedge \theta^j$.

A *left action* of a Lie group G on a manifold M is a smooth map $G \times M \to M$, $(g, x) \mapsto gx$, such that $g_1(g_2)x = (g_1g_2)x$ for all g_i, x and $1 \cdot x = x$ when 1 is the neutral element. Similarly, one defines the *right action* as a map $M \times G \to M$, $(x, g) \mapsto xg$, such that $x(g_1g_2) = (xg_1)g_2$ and $x \cdot 1 = x$.

The (left) action is *transitive* if for any $x, y \in M$ there is an element $g \in G$ such that $y = gx$ and the action is *free* if for any x, $gx = x$ only when $g = 1$. The action is *faithful* if $gx = x$ for all $x \in M$ only when $g = 1$. The *isotropy group* at $x \in M$ is the group $G_x \subset G$ of elements g such that $gx = x$.

Example 1 Let $H \subset G$ be a closed subgroup of a Lie group G. Then the left (right) multiplication on G defines a left (right) action of H on G.

Example 2 Let $H \subset G$ be a closed subgroup in a Lie group. Then the space $M = G/H$ of left cosets gH is a smooth manifold, see S. Helgason, section II.4. There is a natural left action given by $g' \cdot (gH) = g'gH$. In general, when G acts transitivly (from the left) on a manifold M, we can write $M = G/H$ with $H = G_x$ for any fixed element $x \in M$. The bijection $\phi : G/H \to M$ is given by $\phi(gH) = gx$.

For example, when $G = SO(3)$ and $H = SO(2)$ the quotient $M = SO(3)/SO(2)$ can be identified as the unit sphere S^2. Similarly, $SU(3)/SU(2)$ can be identified as the sphere S^5. The sphere S^5 is equal to the set of points $(z_1, z_2, z_3) \in \mathbb{C}^3$ such that $|z_1|^2 + |z_2|^2 + |z_3|^2 = 1$. The point $(1, 0, 0)$ is left invariant exactly by the elements in the subgroup $SU(2) \subset SU(3)$ operating in the z_2z_3-plane. On the other hand, $SU(3)$ acts transitivly on S^5 and so $S^5 = SU(3)/SU(2)$.

Let a left action of a Lie group G on a manifold M be given. Then for each $X \in Lie(G)$ there is a canonical vector field $\hat{X}$ on M defined by $\hat{X}(x) = \frac{d}{dt} e^{tX} \cdot x|_{t=0}$. Similarly, a right action gives a canonical vector field by differentiating $x \cdot e^{tX}$. In the case when $M = G$ and the left action is given by the left multiplication on the group, we have simply $\hat{X} = X$.

A Lie group G acts on itself also through the formula $g \mapsto g_0 g g_0^{-1}$ for $g_0 \in G$. This is called the adjoint action and is denoted by $Ad_{g_0}(g) = g_0 g g_0^{-1}$. Note that the adjoint action is a left action. Because of $Ad_{g_0}(1) = 1$, the derivative of the adjoint action at $g = 1$ gives a linear map, denoted by ad_{g_0}, from T_1G to T_1G, that is, we may view ad_g as a linear map

$$ad_g : Lie(G) \to Lie(G).$$

In the case of a matrix Lie group we have simply $ad_g(X) = gXg^{-1}$, matrix multiplication. Thus we have also

$$ad_g([X, Y]) = [ad_g(X), ad_g(Y)]$$

for all $X, Y \in Lie(G)$. This holds also in the case of an arbitrary Lie group. This means that ad_g is *an automorphism* of the Lie algebra of G.

Exercise Prove the above statement for an arbitrary Lie group.

We also observe, by the the chain rule for differentiation, that $ad_{gg'} = ad_g \circ ad_{g'}$ for all $g, g \in G$. This means that the map $g \mapsto ad_g$ is a representation of the Lie group in the vector space $Lie(G)$.

4.2. Definition of a principal bundle and examples

Let G be a Lie group and M a smooth manifold. *A principal G bundle over M* is a manifold which locally looks like $M \times G$.

Definition 4.2.1. *A smooth manifold P is a principal G bundle over the manifold M, if a smooth right action of G on P is given, i. e., a map $P \times G \to P$, $(p, g) \mapsto pg$, such that $p(gg') = (pg)g' \forall p \in P$ and g, g' in G, and if there is given a smooth map $\pi : P \to M$ such that*

1. *$\pi(pg) = \pi(p)$ for all g in G.*
2. *$\forall x \in M$ there exists an open neighborhood U of x and a diffeomorphism (local trivialization) $f : \pi^{-1}(U) \to U \times G$ of the form $f(p) = (\pi(p), \phi(p))$ such that $\phi(pg) = \phi(p)g \, \forall p \in \pi^{-1}(U), g \in G$.*

The manifold P is the *total space* of the bundle, M is the *the base space,* and π is the bundle projection. The *trivial bundle* $P = M \times G$ is defined by the projection $\pi(x, g) = x$ and by the natural right action of G on itself.

Consider two bundles $P_i = (P_i, \pi_i, M_i; G)$ with the same structure group G. A smooth map $\phi : P_1 \to P_2$ is a *G bundle map* , if $\phi(pg) = \phi(p)g$ for all p and g. Two bundles P_1 and P_2 are *isomorphic* if there is a bijective bundle map $P_1 \to P_2$. An isomorphism of a bundle onto itself is an *automorphism* .

If $H \subset G$ is a closed subgroup then G is a principal H bundle over the homogeneous space G/H. The right action of H on G is just the right multiplication in G and the projection is the canonical projection on the quotient.

Example 4.2.2. Take $G = SU(2)$ and $H = U(1)$

$$H : \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}, \varphi \in \mathbb{R}.$$

A general element g of G is

$$g = \begin{pmatrix} z_1 & -\overline{z}_2 \\ z_2 & \overline{z}_1 \end{pmatrix},$$

with $|z_1|^2 + |z_2|^2 = 1$. Writing z_1 and z_2 in terms of their real and imaginary parts we see that the group G can be identified with the unit sphere S^3 in $\mathbb{R}^4$. We can define a map $\pi : G \to S^2$ by $\pi(g) = g\sigma_3 g^{-1}$, where σ_3 is the matrix $diag(1, -1)$; elements of $\mathbb{R}^3$ are represented by Hermitian traceless 2×2 matrices. The Euclidean metric is given by $\|x\|^2 = -\det x$. The kernel of the map π is precisely $U(1)$; thus we have a $U(1)$ fibration over $S^2 = SU(2)/U(1)$ in S^3.

Exercise 4.2.3. Let $S_+ = \{x \in S^2 | x_3 \neq -1\}$ and $S_- = \{x \in S^2 | x_3 \neq +1\}$. Construct local trivializations $f_\pm : \pi^{-1}(S_\pm) \to S_\pm \times U(1)$.

The bundle $S^3 \to S^2$ is nontrivial; it is not isomorphic to $S^2 \times S^1$ for topological reasons. Namely, S^3 is a simply connected manifold whereas the fundamental group of $S^2 \times S^1$ is equal to $\pi_1(S^1) = \mathbb{Z}$ [M. Greenberg: Lectures on Algebraic Topology].

Let $\{U_\alpha\}_{\alpha \in \Lambda}$ be an open cover of the base space M of a principal bundle P and let $p \mapsto (\pi(p), \phi_\alpha(p)) \in U_\alpha \times G$ be a set of local trivializations. If $p \in \pi^{-1}(U_\alpha \cap U_\beta)$, we can write

$$\phi_\alpha(p) = \xi_{\alpha\beta}(p)\phi_\beta(p),$$

where $\xi_{\alpha\beta}(p) \in G$. Now $\phi_\alpha(pg) = \phi_\alpha(p)g$ and $\phi_\beta(pg) = \phi_\beta(p)g$ from which follows that $\xi_{\alpha\beta}(pg) = \xi_{\alpha\beta}(p)$ and thus $\xi_{\alpha\beta}$ can be thought of as a function on the base space $U_\alpha \cap U_\beta$. If $p \in \pi^{-1}(U_\alpha \cap U_\beta \cap U_\gamma)$ and $x = \pi(p)$, then $\phi_\alpha(p) = \xi_{\alpha\beta}(x)\phi_\beta(p) = \xi_{\alpha\beta}(x)\xi_{\beta\gamma}(x)\phi_\gamma(p)$ so that

$$\xi_{\alpha\beta}(x)\xi_{\beta\gamma}(x) = \xi_{\alpha\gamma}(x).$$

In general, a collection of G-valued functions $\{\xi_{\alpha\beta}\}$ for the covering $\{U_\alpha\}$ is a *one-cocycle (with values in G)* if the above equation holds for all x in $U_\alpha \cap U_\beta \cap U_\gamma$ and for all triples of indices.

If we make the transformations $\phi'_\alpha = \eta_\alpha \phi_\alpha$ for some functions $\eta_\alpha : U_\alpha \to G$, then

$$\xi_{\alpha\beta} \mapsto \xi'_{\alpha\beta} = \eta_\alpha^{-1} \xi_{\alpha\beta} \eta_\beta.$$

If we can find the maps η_α such that $\xi'_{\alpha\beta} = 1 \forall \alpha, \beta$, then $\xi_{\alpha\beta} = \eta_\alpha \eta_\beta^{-1}$ and we say that the one-cocycle ξ is a *coboundary*.

Let $(P, \pi, M), (P', \pi', M')$ be a pair of principal G bundles and let $f : P \to P'$ be a bundle map. We define the induced map $\hat{f} : M \to M'$ by $\hat{f}(x) = \pi'(f(p))$, where p is an arbitrary element in the fiber $\pi^{-1}(x)$.

Theorem 4.2.4. *Let P and P' be a pair of principal G bundles over M. Let $\{U_\alpha, \phi_\alpha\}_{\alpha\in\Lambda}$ (respectively, $\{U_\alpha, \phi'_\alpha\}_{\alpha\in\Lambda}$) be a system of local trivializations for P (respectively, for P'). Let $\xi_{\alpha\beta}$ and $\xi'_{\alpha\beta}$ be the corresponding transition functions. Then there exists an isomorphism $f : P \to P'$ such that $\hat{f} = id_M$ if and only if the transition functions differ by a coboundary, that is, $\xi'_{\alpha\beta}(x) = \eta_\alpha(x)^{-1}\xi_{\alpha\beta}(x)\eta_\beta(x)$ in $U_\alpha \cap U_\beta$ for some functions $\eta_\alpha : U_\alpha \to G$.*

Proof. 1) Suppose first that $\xi'_{\alpha\beta} = \eta_\alpha^{-1}\xi_{\alpha\beta}\eta_\beta$ for all $\alpha, \beta \in \Lambda$. Define $f : P \to P'$ as follows. Let $p \in P$ and $x = \pi(p)$. Choose $\alpha \in \Lambda$ such that $x \in U_\alpha$. Using a local trivialization (U_α, ϕ'_α) at x we set $f(p) = (x, f_\alpha(p))$, where $f_\alpha(p) = \eta_\alpha(x)^{-1}\phi_\alpha(p)$. We have to show that the map is well-defined: If $x \in U_\alpha \cap U_\beta$ then $\phi_\beta(p) = \xi_{\beta\alpha}(x)\phi_\alpha(p)$ and thus

$$\begin{aligned} f_\beta(p) &= \eta_\beta(x)^{-1}\phi_\beta(p) = \eta_\beta(x)^{-1}\xi_{\beta\alpha}(x)\phi_\alpha(p) \\ &= \xi'_{\beta\alpha}(x)[\eta_\alpha(x)^{-1}\phi_\alpha(p)] = \xi'_{\beta\alpha}(x)f_\alpha(p). \end{aligned}$$

We conclude that $(x, f_\alpha(p))$ and $(x, f_\beta(p))$ represent the same element in P'. The equation $f(pg) = f(p)g$ follows from $\phi_\alpha(pg) = \phi_\alpha(p)g$.

2) Let $f : P \to P'$ be an isomorphism. We can define

$$\eta_\alpha(x) = \phi_\alpha(p)\phi'_\alpha(f(p))^{-1},$$

where $p \in \pi^{-1}(x)$ is arbitrary. It follows at once from the definition of the transition functions that the collection $\{\eta_\alpha\}_{\alpha\in\Lambda}$ satisfies the requirements.

Let $\{\xi_{\alpha\beta}\}_{\alpha,\beta\in\Lambda}$ be a one-cocycle with values in G, subordinate to an open cover $\{U_\alpha\}$ on a manifold M. We can construct a principal G bundle P from this data. Let $C = \amalg(\alpha, U_\alpha \times G)$ be the disjoint union of all the sets $U_\alpha \times G$. Define an equivalence relation in C by $(\alpha, x, g) \sim (\alpha', x', g')$ if and only if $x = x'$ and $g' = \xi_{\alpha'\alpha}(x)g$. Set $P = C/\sim$. The action of G in P is given by $(\alpha, x, g)g_0 = (\alpha, x, gg_0)$. The smooth structure on P is defined such that the sets $U_\alpha \times G$ are smooth coordinate charts for P.

Exercise 4.2.5. Complete the construction of P above.

Let (P, π, M) be a principal G bundle. A *(global) section* of P is a map $\psi : M \to P$ such that $\pi \circ \psi = id_M$.

Exercise 4.2.6. Show that a principal bundle is trivial if and only if it has a global section.

A local section consists of an open set $U \subset M$ and a map $\psi : U \to P$ such that $\pi \circ \psi = id_U$. If $f : \pi^{-1}(U) \to U \times G$ is a local trivialization we can define a local section by $\psi(x) = f^{-1}(x, h(x))$, where $h : U \to G$ is an arbitrary (smooth) function.

Let $H \subset G$ be a closed subgroup. We say that the bundle P has been *reduced* to a principal H *subbundle* Q, if $Q \subset P$ is a submanifold such that $qh \in Q$ for all $q \in Q, h \in H, \pi(Q) = M$ and H acts transitively in each fiber $Q_x = \pi^{-1}(x) \cap Q$.

Any manifold M of dimension n carries a natural principal $GL(n, \mathbb{R})$ bundle, namely, the bundle FM of linear frames. The fiber F_xM at a point $x \in M$ consists of all frames (ordered basis) of the tangent space T_xM. The group $GL(n, \mathbb{R})$ acts in F_xM by $(f_1, f_2, ..., f_n)A = (\sum_{i=1}^n A_{i1}f_i, \sum_{i=1}^n A_{i2}f_i, ..., \sum_{i=1}^n A_{in}f_i)$, where the f_i's are tangent
vectors at x and $A = (A_{ij}) \in GL(n, \mathbb{R})$. One can construct a local trivialization by choosing a local coordinate system $(x_1, x_2, ..., x_n)$ in M. In local coordinates the vectors of a frame f can be written as $f_i = \sum f_{ij}\partial_j$. This defines a mapping $f \mapsto (f_{ij}) \in GL(n, \mathbb{R})$. The collection $(\partial_1, ..., \partial_n)$ of vector fields defines a local section of FM.

If the manifold M has some additional structure the bundle FM can generally be reduced to a subbundle. For example, if M is a Riemannian manifold with metric g, then we can define the subbundle $OFM \subset FM$ consisting of *orthonormal* frames with respect to the metric g. If in addition M is oriented, then it makes sense to speak of the bundle *SOFM of oriented orthonormal* frames: A frame $(f_1, \ldots, f_n)$ at a point x is oriented if $\omega(x; f_1, \ldots, f_n)$ is positive, where ω is a n form defining the orientation. The structure group of OFM is the orthogonal group $O(n)$ and of $SOFM$ the special orthogonal group $SO(n)$ consisting of orthogonal matrices with determinant=1.

Let $\mathbf{g}$ be the Lie algebra of the Lie group G. To any $A \in \mathbf{g}$ there corresponds canonically a one-parameter subgroup $h_A(t) = \exp tA$. We define a vector field $\hat{A}$ on the G bundle P such that the tangent vector $\hat{A}(p)$ at $p \in P$ is equal to $\frac{d}{dt}[p \cdot h_A(t)]|_{t=0}$. Let $g \in G$ be any fixed element. The right translation $r_g(p) = pg$ on P determines canonically a transformation $X \mapsto (r_g)_*X$ on vector fields: The tangent vector of the transformed field at a point p is simply obtained by applying the derivative of the mapping r_g to the tangent vector $X(pg^{-1})$.

Proposition 4.2.7. *For any $A \in \mathbf{g}$ the vector field $\hat{A}$ is equivariant, that is, $(r_g)_*\hat{A} = \widehat{ad_g^{-1}A} \forall g \in G$.*

Proof. Using a local trivialization,

$$\hat{A}(p) = \frac{d}{dt}(\pi(p), \phi(pe^{tA}))\Big|_{t=0}$$

and therefore

$$\begin{aligned} ((r_g)_*\hat{A})(p) &= T_{pg^{-1}}r_g \cdot \frac{d}{dt}(\pi(pg^{-1}), \phi(pg^{-1}e^{tA}))\Big|_{t=0} \\ &= \frac{d}{dt}(\pi(pg^{-1}), \phi(pg^{-1}e^{tA}g))\Big|_{t=0} \\ &= \frac{d}{dt}(0, \phi(pe^{tad_g^{-1}A}))\Big|_{t=0} \\ &= \widehat{ad_g^{-1}A}(p). \end{aligned}$$

4.3. Connection and curvature in a principal bundle

Let E and M be a pair of manifolds, V a vector space and $\pi : E \to M$ a smooth surjective map.

Definition 4.3.1. *The manifold E is a vector bundle over M with fiber V, if*

1. *$E_x = \pi^{-1}(x)$ is isomorphic with the vector space V for each $x \in M$*
2. *$\pi : E \to M$ is locally trivial: Any $x \in M$ has an open neighborhood U with a diffeomorphism $\phi : \pi^{-1}(U) \to U \times V$, $\phi(z) = (\pi(z), \xi(z))$, where the restriction of ξ to a fiber E_x is a linear isomorphism onto V.*

The product $M \times V$ is the *trivial vector bundle* over M, with fiber V. In this case the projection map $M \times V \to M$ is simply the projection onto the first factor.

A direct sum of two vector bundles E and F over the same manifold M is the bundle $E \oplus F$ with fiber $E_x \oplus F_x$ at a point $x \in M$. The *tensor product bundle* $E \otimes F$ is the vector bundle with fiber $E_x \otimes F_x$ at $x \in M$.

Example 4.3.2. The tangent bundle TM of a manifold M is a vector bundle over M with fiber $T_xM \simeq \mathbb{R}^n$, where $n = dimM$. The local trivializations are given by local coordinates: If $(x_1, x_2, ..., x_n)$ are local coordinates on $U \subset M$, then the value of ξ for a tangent vector $w \in T_x M$, $x \in M$, is obtained by expanding w in the basis defined by the vector fields $(\partial_1, ..., \partial_n)$.

A section of a vector bundle E is a map $\psi : M \to E$ such that $\pi \circ \psi = id_M$. The space $\Gamma(E)$ of sections of E is a linear vector space; the addition and multiplication by scalars is defined pointwise. A principal bundle may or may not have global sections but a vector bundle always has nonzero sections. A section can be multiplied by a smooth function $f \in C^\infty(M)$ pointwise, $(f\psi)(x) = f(x)\psi(x)$.

Let (P, π, M) be a principal G bundle. *The space V of vertical vectors* in the tangent bundle TP is the subbundle of TP with fiber $\{v \in T_pP | \pi(v) = 0\}$ at $p \in P$. If P is trivial, $P = M \times G$, then the vertical subspace at $p = (x, g)$ consists of vectors tangential to G at g. In general, the dimension of the fiber V_p is equal to $\dim G$.

Definition 4.3.3. *A connection in the principal bundle P is a smooth distribution $p \mapsto H_p$ of subspaces of T_p such that*

(1) *The tangent space T_p is a direct sum of V_p and H_p $\forall p \in P$*
(2) *The distribution is equivariant, i.e., $r_g H_p = H_{pg}$ $\forall p \in P, g \in G$.*

Smoothness means that the distribution can be locally spanned by smooth vector fields. We shall denote by pr_h (respectively, pr_v) the projection in T_p to the *horizontal* subspace H_p (respectively, vertical subspace V_p).

Let $A \in \mathbf{g}$ and let $\hat{A}$ be the corresponding equivariant vector field on P. The field $\hat{A}$ is vertical at each point. Since the group G acts freely and transitively on P, the mapping $A \mapsto \hat{A}(p)$ is a linear isomorphism onto V_p for all $p \in P$. Thus for each $X \in T_pP$ there is a uniquely defined element $\omega_p(X) \in \mathbf{g}$ such that

$$\widehat{\omega_p(X)} = pr_v X$$

at p. The mapping $\omega_p : T_pP \to \mathbf{g}$ is linear, thus defining a differential form of degree one on P, with values in the Lie algebra $\mathbf{g}$. The form ω is the *connection form* of the connection H.

Proposition 4.3.4. *The connection form satisfies*

(1) $\omega_p(\hat{A}(p)) = A \,\forall A \in \mathbf{g}$,
(2) $r_a^* \omega = ad_a \omega \,\forall a \in G$.

Furthermore, each $\mathbf{g}$*-valued differential form on* P *which satisfies the above conditions is a connection form of a uniquely defined connection in* P*.*

Proof. The first equation follows immediately from the definition of ω. To prove the second, we first note that

$$(\widehat{ad_a^{-1}A})(p) = \frac{d}{dt} p e^{t ad_a^{-1}A}|_{t=0} = \frac{d}{dt} pa^{-1}e^{tA}a|_{t=0} = r_a \hat{A}(pa^{-1}).$$

By the equivariantness property of the distribution H_p, the right translations r_a commute with the horizontal and vertical projection operators. Thus [we write $(A)\hat{}$ for $\hat{A}$ in case of long expressions]

$$\begin{aligned}(ad_a\omega_p(X))\hat{}(p) &= r_a^{-1} \cdot \widehat{\omega_p(X)}(pa) \\ &= r_a^{-1}(pr_v X)(pa) = pr_v(r_a^{-1}X)(pa) \\ &= (\omega_p(r_a^{-1}X))\hat{}(pa).\end{aligned}$$

Taking account that $(r_a^*\omega)_p(X) = \omega_{pa}(r_a X)$ we get the second relation.

Let then ω be any form satisfying both equations. We define the horizontal subspaces $H_p = \{X \in T_p | \omega_p(X) = 0\}$. If $X \in H_p \cap V_p$, then $X = \hat{A}(p)$ for some $A \in \mathbf{g}$ and $\omega_p(\hat{A}(p)) = A = \omega_p(X) = 0$, from which follows $H_p \cap V_p = 0$. By (1) and a simple dimensional argument we get $T_p = H_p + V_p$. For $X \in H_p$ and $a \in G$ we obtain

$$\omega_{pa}(r_a X) = (r_a^*\omega)_p(X) = ad_a\omega_p(X) = 0,$$

and therefore $r_a X \in H_{pa}$, which shows that the distribution H_p is equivariant and indeed defines a connection in P.

Let ω be a connection form in (P, π, M). Let $U \subset M$ be open and $\psi : U \to P$ a local section. The pull-back $A = \psi^*\omega$ is a one-form on U. Consider another local section $\phi : V \to P$ and set $A' = \phi^*\omega$. We can write $\psi(x) = \phi(x)g(x)$ for $g : U \cap V \to G$, where $g(x)$ is a smooth G valued function. We want to relate A to A'. Noting that

$$T_x\psi = r_{g(x)}T_x\phi + (g^{-1}T_x g)\hat{}(\psi(x))$$

by the Leibnitz rule, we get

$$\begin{aligned}A_x(u) &= \omega_{\psi(x)}(T_x\psi \cdot u) = \omega_{\psi(x)}(r_{g(x)}T_x\phi \cdot u + (g^{-1}T_x g \cdot u)\hat{}(\psi(x))) \\ &= ad_{g(x)}^{-1}\omega_{\phi(x)}(T_x\phi \cdot u) + g^{-1}T_x g \cdot u.\end{aligned}$$

For a matrix group G we can simply write

$$A = g^{-1}A'g + g^{-1}dg.$$

The transformation relating A to A' is called *a gauge transformation* . Next we define the two-form

$$F = dA + \frac{1}{2}[A, A] \tag{4.3.5}$$

on U. The commutator of Lie algebra valued one-forms is defined by

$$[A, B](u, v) = [A(u), B(v)] - [A(v), B(u)]$$

for a pair u, v of tangent vectors. We shall compute the effect of a gauge transformation $(U, \psi) \to (V, \phi)$ on F:

$$\begin{aligned} F &= dA + \tfrac{1}{2}[A, A] \\ &= g^{-1}dA'g - [g^{-1}dg, g^{-1}A'g] - \tfrac{1}{2}[g^{-1}dg, g^{-1}dg] \\ &\quad + \tfrac{1}{2}[g^{-1}A'g + g^{-1}dg, g^{-1}A'g + g^{-1}dg] \\ &= g^{-1}(dA' + \tfrac{1}{2}[A', A'])g = g^{-1}F'g. \end{aligned}$$

The *curvature form* F is a pull-back under ψ of a gobally defined two-form Ω on P. The latter is defined by

$$\Omega_p(u, v) = a^{-1}F_x(\pi u, \pi v)a,$$

where $p \in \pi^{-1}(x)$, u, v tangent vectors at p and $a \in G$ is an element such that $p = \psi(x)a$. The left-hand side does not depend on the local section. Writing $p = \phi(x)a' = \psi(x)g(x)a'$ we get

$$a'^{-1}F'_x(\pi u, \pi v)a' = a'^{-1}g(x)^{-1}F_x(\pi u, \pi v)g(x)a' = a^{-1}F_x(\pi u, \pi v)a.$$

Since A is the pull-back of ω and F is the pull-back of Ω we obtain from 4.3.5

$$\Omega = d\omega + \frac{1}{2}[\omega, \omega] \tag{4.3.6}$$

Exercise 4.3.7. Prove the *Bianchi identity* $dF + [A, F] = 0$. (The 3-form $[A, F]$ is defined by an antisymmetrization of $[A(u), F(v, w)]$ with respect to the triplet (u, v, w) of tangent vectors.)

Let (P, π, M) be a principal G bundle and $\rho : G \to AutV$ a linear representation of G in a vector space V. We define the manifold $P \times_G V$ to be the set of equivalence classes $P \times V/\sim$, where the equivalence relation is defined by $(p, v) \sim (pg^{-1}, \rho(g)v)$, for $g \in G$. There is a natural projection $\theta : P \times_G V \to M$, $[(p, v)] \mapsto \pi(p)$. The inverse image $\theta^{-1}(x) \cong V$, since G acts freely and transitively in the fibers of P. The linear structure in a fiber $\theta^{-1}(x)$ is defined by $[(p, v)] + [(p, w)] = [(p, v + w)]$, $\lambda[(p, v)] = [(p, \lambda v)]$. Local trivializations of $P \times_G V$ are obtained from local trivializations $p \mapsto (\pi(p), \phi(p)) \in M \times G$ of P by $[(p, v)] \mapsto (\pi(p), \rho(\phi(p))v)$. Thus $P \times_G V$ is a vector bundle over M, *the vector bundle associated to* P via the representation ρ of G.

Example 4.3.8. Let $P = SU(2), M = S^2 = SU(2)/U(1), G = U(1), V = \mathbb{C}$ and $\rho(\lambda) = \lambda^2$ for $\lambda \in U(1)$. The associated vector bundle $E = SU(2) \underset{U(1)}{\times} \mathbb{C}$ is in fact the tangent bundle of the sphere S^2. The isomorphism is obtained as follows. Fix a linear isomorphism of $\mathbb{C} \cong \mathbb{R}^2$ with the tangent space of S^2 at the point x, which has as its isotropy group the given $U(1)$. The map $E \to TS^2$ is defined by $(g, v) \mapsto D(g)v$, where $D(g)$ is the 2-1 representation of $SU(2)$ in $\mathbb{R}^3$. The tangent vectors of S^2 are represented by vectors in $\mathbb{R}^3$ by the natural embedding $S^2 \subset \mathbb{R}^3$.

4.4. Parallel transport

Let H be a connection in a principal G bundle (P, π, M). *A horizontal lift* of a smooth curve $\gamma(t)$ on the base manifold M is a smooth curve $\gamma^*(t)$ on P such that the tangent vector $\dot{\gamma}^*(t)$ is *horizontal* at each point on the curve and $\pi(\gamma^*(t)) = \gamma(t)$.

Lemma 4.4.1. *Let $X(t)$ be a smooth curve on the Lie algebra* $\mathbf{g}$ *of G, defined on an interval $[t_0, t_1]$. Then there exists a unique smooth curve $a(t)$ on G such that $\dot{a}(t)a(t)^{-1} = X(t) \forall t \in [t_0, t_1]$ and such that $a(t_0) = e$.*

Proof. See Kobayashi and Nomizu, vol. I, p. 69.

Proposition 4.4.2. *Let $\gamma(t)$ be a smooth curve on M and p an element in the fiber over $\gamma(t_0)$. Then there exists a unique horizontal lift $\gamma^*(t)$ of $\gamma(t)$ such that $\gamma^*(t_0) = p$.*

Proof. Choose first any (smooth) curve $\phi(t)$ on P such that $\pi(\phi) = \gamma$ and $\phi(t_0) = p$.

We are looking for the solution in the form $\gamma^*(t) = \phi(t)g(t)$, where $g(t)$ is a curve on G such that $g(t_0) = e$. Now $\gamma^*(t)$ is a solution if

$$\dot{\gamma}^*(t) = r_{g(t)} \cdot \dot{\phi}(t) + (g(t)^{-1}\dot{g}(t))\hat{}[\phi(t)g(t)]$$

is horizontal. Let ω be the connection form of the connection H. A tangent vector on P is horizontal if and only if it is in the kernel of ω. We get the differential equation

$$\begin{aligned} 0 = \omega(\dot{\gamma}^*(t)) &= \omega(r_{g(t)}\dot{\phi}(t)) + \omega([g(t)^{-1}\dot{g}(t)]\hat{}[\phi(t)g(t)]) \\ &= ad_{g(t)}^{-1}\omega(\dot{\phi}(t)) + g(t)^{-1}\dot{g}(t). \end{aligned}$$

Applying ad_g to this equation we get

$$\dot{g}(t)g(t)^{-1} = -\omega(\dot{\phi}(t)).$$

The solution $g(t)$ exists and is unique by the previous lemma.

Example 4.4.3. Let $P = M \times U(1)$, M simply connected. A connection form ω can be written as $\omega_{(x,g)}(u, a) = A_x(u) + g^{-1} \cdot a$, where u is a tangent vector at $x \in M$ and a is a tangent vector at $g \in U(1)$; the Lie algebra of $U(1)$ is identified by the set of purely imaginary complex numbers. Let $\gamma(t)$ be a curve on M. The horizontal lift of $\gamma(t)$ which goes through $(\gamma(t), g)$ at time $t = 0$ is $\gamma^*(t) = (\gamma(t), g(t))$ with

$$g(t) = g \cdot \exp\left(\int_0^t -A_{\gamma(s)}(\dot{\gamma}(s))ds\right).$$

In particular, for a *closed contractible curve*, $\gamma(0) = \gamma(1)$, we get by Stokes's theorem

$$g(1) = g \cdot \exp(-\int_S F),$$

where $F = dA$ is the curvature two-form and the integration is taken over any surface on M bounded by the closed curve γ.

We define *the parallel transport* along a curve $\gamma(t)$ on M as a mapping $\tau : \pi^{-1}(x_0) \to \pi^{-1}(x_1)$ ($x_0 = \gamma(t_0), x_1 = \gamma(t_1)$ points on the curve). The value $\tau(p_0)$ for $p_0 \in \pi^{-1}(x_0)$ is given as follows: Let $\gamma^*(t)$ be a horizontal lift of $\gamma(t)$ such that $\gamma^*(t_0) = p_0$. Then $\tau(p_0) = \gamma^*(t_1)$.

Exercise 4.4.4. Prove the following properties of the parallel transport.

(1) $\tau \circ r_g = r_g \circ \tau \forall g \in G$

(2) If γ_1 is a path from x_0 to x_1 and γ_2 is a path from x_1 to x_2 then the parallel transport along the composed path $\gamma_2 * \gamma_1$ is equal to the product of parallel transport along γ_1 followed by a parallel transport along γ_2.

(3) The parallel transport is a one-to-one mapping between the fibers $\pi^{-1}(x_0)$ and $\pi^{-1}(x_1)$.

4.5. Covariant differentiation in vector bundles

Let E be a vector bundle over a manifold M with fiber V, dim $V = n$. The vector space V is defined over the field $\mathbb{K} = \mathbb{R}$ or $K = \mathbb{C}$. A vector bundle can always be thought of as an associated bundle to a principal bundle. Namely, let P_x denote the space of all linear frames in the fiber E_x for $x \in M$. Using the local trivializations of E it is not difficult to see that the spaces P_x fit together and form naturally a smooth manifold P. Fix a basis $w = \{w_1, \dots, w_n\}$ in E_x. Then any other basis of E_x can be obtained from w by a linear tranformation $w_i' = \sum A_{ji} w_j$ and therefore P_x can be identified with the group $GL(n, \mathbb{K})$ of all linear transformations in $\mathbb{K}^n$; it should be stressed that this identification depends on the choice of w. However, we have a well-defined mapping $P \times GL(n, \mathbb{K}) \to P$ given by the basis transformations and this shows that P can be thought of as a principal $GL(n, \mathbb{K})$ bundle over M.

The vector bundle E is now isomorphic with the associated bundle $P \times_\rho \mathbb{K}^n$, where ρ is the natural representation of $GL(n, \mathbb{K})$ in $\mathbb{K}^n$. The isomorphism is defined as follows. Let $w \in P_x$ and $a \in \mathbb{K}^n$. We set $\phi(w, a) = \sum a_i w_i$. This gives a mapping from $P \times \mathbb{K}^n$ to E which is obviously linear in a. For a fixed w the mapping $a \mapsto \phi(w, a)$ gives an isomorphism between $\mathbb{K}^n$ and E_x. Let $w' = w \cdot g$ and $a' = \rho(g^{-1})a$ for some $g \in GL(n, \mathbb{K})$. We have to show that $\phi(w', a') = \phi(w, a)$; but this follows immediately from the definitions.

Often the bundle E can be thought of as an associated bundle to a principal bundle with a smaller structure group than the group $GL(n, \mathbb{K})$. This happens when there is some extra structure in E. For example, assume there is a *fiber metric* in E: This means that there is an inner product $< \cdot, \cdot >_x$ in each fiber E_x such that $x \mapsto < \psi(x), \psi(x) >_x$ is a smooth function for any (local) section ψ. We can then define the bundle of orthonormal frames in E

with structure group $U(n)$ in the complex case and $O(n)$ in the real case. The vector bundle E is now an associated bundle to the bundle of orthonormal frames.

We shall now assume that E is given as an associated vector bundle $P \times_\rho V$ to some principal bundle P, with a connection H, over M. Let G be the structure group of P. For each vector field X on M we can define a linear map ∇_X of the space $\Gamma(E)$ of sections into itself such that

(1) $\nabla_{X+Y} = \nabla_X + \nabla_Y$
(2) $\nabla_{fX} = f\nabla_X$
(3) $\nabla_X(f\psi) = (Xf)\psi + f\nabla_X\psi$

for all vector fields X, Y, smooth functions f and sections ψ. We shall give the definition in terms of a local trivialization $\xi : U \to P$, where $U \subset M$ is open. Locally, a section $\psi : M \to E$ can be written as

$$\psi(x) = (\xi(x), \phi(x)),$$

where $\phi : U \to V$ is some smooth function. Let A denote the pull-back $\xi^*\omega$ of the connection form ω in P. The representation ρ of G in V defines also an action of the Lie algebra $\mathbf{g}$ in V. We set

$$\nabla_X\psi = (\xi, X\phi + A(X)\phi),$$

where $A(X)$ is the Lie algebra valued function giving the value of the one-form A in the direction of the vector field X.

We have to check that our definition does not depend on the choice of the local trivialization. So let $\xi'(x) = \xi(x) \cdot g(x)$ be another local trivialization, where $g : U \to G$ is a smooth function. The vector potential with respect to the trivialization ξ' is $A' = g^{-1}Ag + g^{-1}dg$. Now $(\xi, \phi) \sim (\xi', \phi')$, where $\phi' = g^{-1}\phi$ (we simplify the notation by dropping ρ) and therefore $(\xi', X\phi' + A'(X)\phi')$ is equal to

$$\begin{aligned}(\xi', -g^{-1}(Xg)g^{-1}\phi + g^{-1}X\phi + (g^{-1}Ag + g^{-1}Xg)g^{-1}\phi)&\\ = (\xi', g^{-1}(X\phi + A(X)\phi)) \sim (\xi, X\phi + A(X)\phi)&\end{aligned}$$

which shows that ∇_X is well-defined.

Exercise 4.5.1. Prove that ∇_X defined above satisfies (1)-(3).

The commutator of the *covariant derivatives* ∇_X is related to the curvature of the connection in the following way:

$$
\begin{aligned}
[\nabla_X, \nabla_Y]\psi &= (\xi, [X + A(X), Y + A(Y)]\phi) \\
&= (\xi, ([X,Y] + X \cdot A(Y) - Y \cdot A(X) + [A(X), A(Y)])\phi) \\
&= (\xi, (F(X,Y) + [X,Y] + A([X,Y]))\phi)
\end{aligned}
$$

where $F = dA + \frac{1}{2}[A, A]$. Thus we can write

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

when acting on the functions ϕ.

A section ψ is *covariantly constant* if $\nabla_X \psi = 0$ for all vector fields. From the above commutator formula we conclude that one can find at each point in the base space a local basis of covariantly constant sections of the vector bundle if and only if the curvature vanishes.

4.6. An example: The monopole line bundle

Construction of the basic monopole bundle

Let G be a Lie group and $\mathbf{g}$ its Lie algebra. Let us denote by ℓ_g the left translation $\ell_g(a) = ga$ in G. The left invariant Maurer-Cartan form $\theta_L = g^{-1}dg$ is the $\mathbf{g}$-valued one form on G which sends a tangent vector X at $g \in G$ to the element $\ell_g^{-1}X \in T_eG$ in the Lie algebra. Similarly, we can define the right invariant Maurer-Cartan form $\theta_R = dgg^{-1}$, $\theta_R(g; X) = r_g^{-1}X$. By taking commutators, we can define higher order forms. For example, the form $[g^{-1}dg, g^{-1}dg]$ sends the pair (X, Y) of tangent vectors at g to $2[\ell_g^{-1}X, \ell_g^{-1}Y] \in \mathbf{g}$.

Taking projections to one dimensional subspaces of $\mathbf{g}$ we get real valued one-forms on G.

Let $< \cdot, \cdot >$ be a bilinear form on $\mathbf{g}$ and $\sigma \in \mathbf{g}$. Then $\alpha =< \sigma, g^{-1}dg >$ is a well-defined one form. Let us compute the exterior derivative of α. Let X, Y be a pair of left invariant vector fields on G. Now

$$
\begin{aligned}
d\alpha(g; X, Y) &= X \cdot \alpha(Y) - Y \cdot \alpha(X) - \alpha([X, Y]) \\
&= -\alpha([X, Y])
\end{aligned}
$$

since $\alpha(Y)(g) =< \sigma, \ell_g^{-1} Y >$ is a constant function on G and similarly for $\alpha(X)$. Since the left invariant vector fields on a Lie group span the tangent space at each point, we conclude

$$d\alpha = - < \sigma, \tfrac{1}{2}[g^{-1}dg, g^{-1}dg] > .$$

We have not yet defined the exterior derivative of a Lie algebra valued differential form, but motivated by the computation above we set

$$d(g^{-1}dg) = -\frac{1}{2}[g^{-1}dg, g^{-1}dg].$$

A bilinear form $< \cdot, \cdot >$ on $\mathbf{g}$ is *invariant* if

$$< [X, Y], Z >= - < Y, [X, Z] >$$

for all X, Y, and Z. Given an invariant bilinear form, the group G has a natural **closed** three-form c_3 which is defined by

$$c_3(g; X, Y, Z) =< \ell_g^{-1} X, [\ell_g^{-1} Y, \ell_g^{-1} Z] > .$$

Thus

$$c_3 =< g^{-1}dg, \tfrac{1}{2}[g^{-1}dg, g^{-1}dg] > .$$

Proposition 4.6.1. $dc_3 = 0$.

Proof.. Recall the definition of the exterior differentiation d: If ω is a n-form and $V_1, ..., V_{n+1}$ are vector fields, then

$$\begin{aligned} d\omega(V_1, ..., V_{n+1}) = &\sum_{i=1}^{n+1} (-1)^{i+1} V_i \cdot \omega(V_1, ..., \hat{V}_i, ..., V_{n+1}) \\ &+ \sum_{i<j} (-1)^{i+j} \omega([V_i, V_j], V_1, ..., \hat{V}_i, ..., \hat{V}_j, ..., V_{n+1}), \end{aligned}$$

where the caret means that the corresponding variable has been dropped. Let us compute dc_3 for left invariant vector fields $X_1, ..., X_4$. Taking account that $c_3(X_i, X_j, X_k)$ is a constant function we get

$$\begin{aligned} dc_3(X_1, ...X_4) = &- 2 < [X_1, X_2], [X_3, X_4] > +2 < [X_1, X_3], [X_2, X_4] > \\ &- 2 < [X_1, X_4], [X_2, X_3] > \\ =&2 < X_1, [[X_3, X_4], X_2] - [[X_2, X_4], X_3] + [[X_2, X_3], X_4] > \\ =&0 \end{aligned}$$

by Jacobi's identity.

If G is a group of matrices we can define an invariant form on $\mathbf{g}$ by $< X, Y >=$ $\mathrm{tr}\, XY$. Then the form c_3 can be written as

$$c_3 = \mathrm{tr}\,(g^{-1}dg)^3.$$

As an example we shall consider in detail the case $G = SU(2)$. Let $\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ and define the one-form $\alpha = -\frac{1}{2}\mathrm{tr}\,\sigma_3 g^{-1}dg$. Remember that $SU(2) \to SU(2)/U(1) = S^2$ is a principal $U(1)$ bundle. The form α is invariant with respect to right translations $g \mapsto gh$ by $h \in U(1)$. Thus α is a connection form in the bundle $SU(2)$ [the Lie algebra of the structure group $U(1)$ can be identified with $i\mathbb{R}$]. Let us compute the curvature. The exterior derivative of α is $\frac{1}{4}\mathrm{tr}\,\sigma_3[g^{-1}dg, g^{-1}dg]$. A tangent vector at $x \in S^2$ can be represented by a tangent vector $\ell_g X$ at $g \in \pi^{-1}(x)\,, X \in \mathbf{g}$, such that X is orthogonal to the $U(1)$ direction, $\mathrm{tr}\,\sigma_3 X = 0$. The curvature in the base space S^2 is then $\Omega(X,Y) = \frac{1}{2}\mathrm{tr}\,\sigma_3[X,Y]$. The form Ω is $\frac{1}{2}\times$ the volume form on S^2: If {X,Y} is an ortonormal system at $x \in S^2$, then $[X,Y] = \pm\frac{i}{2}\sigma_3$ (exercise), the sign depending on the orientation. We obtain $\Omega(X,Y) = \pm\frac{i}{4}\mathrm{tr}\,\sigma_3^2 = \pm\frac{i}{2}$.

The basic monopole line bundle is defined as the associated bundle to the bundle $SU(2) \to S^2$, constructed using the natural one dimensional representation of $U(1)$ in $\mathbb{C}$.

Embedding $S^2 \subset \mathbb{R}^3$ and using Cartesian coordinates $\{x_1, x_2, x_3\}$ we can write the curvature form as

$$\Omega = \frac{1}{4r^3}\varepsilon^{ijk}x_i dx_j \wedge dx_k,$$

where $r^2 = x_1^2 + x_2^2 + x_3^2$ is equal to 1 on S^2. However, we can extend Ω to the space $\mathbb{R}^3 \setminus \{0\}$ using the above formula. The coefficients of the linearly independent forms $dx_2 \wedge dx_3$, $dx_3 \wedge dx_1$ and $dx_1 \wedge dx_2$ form a vector $\vec{B} = \frac{1}{2r^3}(x_1, x_2, x_3) = \frac{\vec{x}}{r^3}$. The field $\vec{B}$ satisfies

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{1}$$

$$\vec{\nabla} \times \vec{B} = 0, \tag{2}$$

i.e., it satisfies Maxwell's equations in vacuum. On the other hand,

$$\int_{S^2} \vec{B} \cdot d\vec{S} = 2\pi \tag{3}$$

for *any* sphere containing the origin. Because of these properties, the field $\vec{B}$ can be interpreted as the magnetic field of a magnetic monopole located at the origin. The integral (3) multiplied by the dimensional constant $1/e$ (e is the unit electric charge) is called the *monopole strength.*

The first Chern class

The magnetic field of the monopole is the curvature of a circle bundle over the unit sphere S^2. The circle bundle we have constructed is a "generator" for the set of all circle bundles over S^2. In general, a principal $U(1)$ bundle over S^2 can be constructed from the transition function $\xi : S_- \cap S_+ \to U(1)$ (cf. 3.2.3). The intersection of the coordinate neighborhoods $S_\pm$ is homeomorphic with the product of an open interval with the circle S^1. It follows that the set of maps ξ decomposes to connected components labelled by the winding number of a map $S^1 \to U(1)$. Let ξ_1 be the transition function of the bundle $SU(2) \to S^2$ with respect to some fixed local trivializations on $S_\pm$. The winding number of ξ_1 is equal to one. The winding number of $\xi_n = (\xi_1)^n$ is equal to n. Let $P(n)$ be the bundle constructed from ξ_n. Let $A_\pm$ be the vector potentials on $S_\pm$ corresponding to the chosen local trivializations and the connection in $SU(2)$ described above.

We have $A_+ = A_- + \xi^{-1}d\xi$ on $S_- \cap S_+$ and therefore $nA_+ = nA_- + \xi_n^{-1}d\xi_n$. Thus nA is a connection in the bundle $P(n)$ and the curvature of $P(n)$ is n times the curvature form Ω of the (basic) monopole bundle. The monopole strength of the bundle $P(n)$ is $2\pi n/e$.

The cohomology class $[\Omega] \in H^2(S^2, \mathbb{R})$ is the *first Chern class* of the bundle. It depends only on the equivalence class of the bundle and not on the chosen connection; we shall return to the proof of the topological invariance of the Chern classes in a more general context later, but as an illustration of the general ideas we give a simple proof for the case at hand. Let $B_\pm$ be the vector potentials on $S_\pm$ of some connection in the bundle $P(n)$. We have $B_+ = B_- + n\xi^{-1}d\xi$ and therefore $A_+ - B_+ = A_- - B_-$ on $S_+ \cap S_-$. It follows that $A - B$ is a globally defined one-form on S^2; the difference of the curvatures corresponding to the connections A and B is equal to $d(A-B)$.

The first Chern class of a circle bundle (or an associated complex line bundle) over a manifold M can be evaluated from the knowledge of the $U(1)$ valued transition functions [R. Bott and L.W. Tu: Differential forms in algebraic topology]. In the example above we needed only one transition function ξ. A representative Ω for the Chern class can be constructed from a vector potential (A_+, A_-) such that $A_- = 0$ for $x_3 < \frac{1}{2}$, A_+ is equal to $\xi^{-1} d\xi$ on the strip $-\frac{1}{2} < x_3 < \frac{1}{2}$, and A_+ is contracted smoothly to zero when approaching the north pole $x_3 = 1$. The first Chern class is always quantized in the sense that the integral of the two-form Ω over any two-dimensional compact surface is 2π times an integer.

4.7. Chern classes

We shall consider polynomials $P(A)$ of a complex $N \times N$ matrix variable A which are invariant in the sense that $P(gAg^{-1}) = P(A)$ for all $g \in GL(N, \mathbb{C})$. For example, if we expand

$$\det\left(1 + \frac{\lambda}{2\pi i} A\right) = \sum_{n=0}^{N} \lambda^n P_n(A) \tag{4.7.1}$$

then the coefficients $P_n(A)$ are homogeneous invariant polynomials of degree n in A. These polynomials will play a special role in the following discussion.

To each homogeneous polynomial $P(A)$ one can associate a unique *symmetric* multilinear form $P(A_1, \dots A_n)$ such that $P(A, \dots, A) = P(A)$. The general formula for the n linear form in terms of $P(A)$ is

$$\begin{aligned} P(A_1, \dots, A_n) = & \frac{1}{n!} \{ P(A_1 + \dots + A_n) \\ & - \sum_j P(A_1 + \dots + \hat{A}_j + \dots + A_n) \\ & + \sum_{j<j'} P(A_1 + \dots \hat{A}_j + \dots \hat{A}_{j'} \dots + A_n) - \dots \}, \end{aligned}$$

with $\hat{A}_j$ deleted. When $P(A)$ is invariant we clearly have $P(gA_1g^{-1}, \dots, gA_ng^{-1}) =$

$P(A_1, \dots, A_n)$. Writing $g = g(t) = \exp(tX)$ we get the useful formula

$$\begin{aligned} 0 &= \frac{d}{dt} P(g(t)A_1 g(t)^{-1}, \dots, g(t)A_n g(t)^{-1})|_{t=0} \\ &= \sum_j P(A_1, \dots, [X, A_j], \dots, A_n). \end{aligned} \tag{4.7.2}$$

If F_i is a $N \times N$ matrix valued differential form of degree k_i on a manifold M, $1 \le i \le n$, and P a symmetric n linear form then we can define a complex valued differential form $P(F_1, \dots, F_n)$ of degree $k_1 + \cdots + k_n = p$ by

$$\begin{aligned} &P(F_1, \dots, F_n)(t_1, \dots, t_p) = \\ &\left(\prod \frac{1}{k_i!}\right) \sum_\sigma \epsilon(\sigma) P(F_1(t_{\sigma(1)}, \dots, t_{\sigma(k_1)}), \dots, F_n(t_{\sigma(p-k_n+1)}, \dots, t_{\sigma(p)})) \end{aligned}$$

where the sum is taken over all permutations of the indices $1, 2, \dots, p$.

Let F be the curvature form of a vector bundle E over M with fiber $\mathbb{C}^N$. The curvature transforms in a change of a local trivialization as $F \mapsto gFg^{-1}$ and therefore $P(F, \dots, F)$ is well-defined, independent of the local trivialization, for any invariant symmetric polynomial P.

Proposition 4.7.3. *The symmetric homogeneous polynomial $P(F, \dots, F)$ of degree n in the curvature F is a closed form of degree $2n$.*

Proof. Locally we can write $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$. Using the property $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{deg\alpha} \alpha \wedge d\beta$ of differential forms we have

$$\begin{aligned} dP(F, \dots, F) &= \sum_j P(F, \dots, dF, \dots, F) \\ &= \sum_j \{P(F, \dots, DF, \dots, F) - P(F, \dots, [A, F], \dots, F)\}. \end{aligned} \tag{4.7.4}$$

The covariant derivative $DF = 0$ by the Bianchi identity and the sum of the terms involving $[A, F]$ is zero by (4.7.2).

In particular, the class in $H^{2n}(M, \mathbb{R})$ defined by the closed $2n$ form Re $P_n(F)$ is called the nth *Chern class* of the bundle E and is denoted by $c_n(E)$.

Theorem 4.7.5. *The Chern classes are topological invariants: They do not depend on the choice of connection in the vector bundle E.*

Proof. Let A_0 and A_1 be two connections in E and F_0, F_1 the corresponding curvatures. Define a one-parameter family $A_t = A_0 + t\eta$ of connections with $\eta = A_1 - A_0$;

note that the difference η transforms homogeneously in a change of local trivialization, $\eta \mapsto g\eta g^{-1}$. Let us introduce the notation $Q(A,B) = nP(A,B,\dots,B)$ when B is repeated $n-1$ times. Using

$$F_t = dA_t + \frac{1}{2}[A_t, A_t] = F_0 + tD\eta + \frac{1}{2}t^2[\eta,\eta],$$

where D is the covariant derivative determined by A_0, we get

$$\frac{d}{dt}P(F_t) = Q\left(\frac{d}{dt}F_t, F_t\right) = Q\left(D\eta, F_t\right) + tQ([\eta,\eta], F_t). \tag{4.7.6}$$

On the other hand,

(4.7.7)
$$\begin{aligned} dQ(\eta, F_t) =& Q(d\eta, F_t) - n(n-1)P(\eta, dF_t, F_t, \dots, F_t) \\ =& Q(d\eta, F_t) - n(n-1)P(\eta, dF_t, F_t, \dots, F_t) \\ & + nP([A_0,\eta], F_t, \dots, F_t) - n(n-1)P(\eta, [A_0, F_t], \dots, F_t) \\ =& Q(D\eta, F_t) - n(n-1)P(\eta, DF_t, F_t, \dots, F_t) \\ =& Q(D\eta, F_t) + tn(n-1)P(\eta, [\eta, F_t], F_t, \dots, F_t) \end{aligned}$$

where we have used $DF_t = DF_0 + tD^2\eta + t^2[D\eta,\eta] = t[F_0,\eta] + t^2[D\eta,\eta] = t[F_t,\eta]$, since $[[\eta,\eta],\eta] = 0$ by Jacobi identity. By (4.7.2) we have

$$P([\eta,\eta], F_t, \dots, F_t) - (n-1)P(\eta, [\eta, F_t], F_t, \dots, F_t) = 0$$

or in other words,

$$Q([\eta,\eta], F_t) - n(n-1)P(\eta, [\eta, F_t], F_t, \dots, F_t) = 0.$$

Using (4.7.7) we get

$$dQ(\eta, F_t) = Q(D\eta, F_t) + tQ([\eta,\eta], F_t)$$

and with (4.7.6) we obtain

$$\frac{d}{dt}P(F_t) = dQ(\eta, F_t). \tag{4.7.8}$$

Integrating this result over the interval $0 \le t \le 1$ we get

$$P(F_1) - P(F_0) = d\int_0^1 Q(\eta, F_t)dt$$

which shows explicitly that the difference of the differential forms $P(F_1)$ and $P(F_0)$ is an exact form.

Given a Hermitian inner product in the fibers of the vector bundle E it is always possible to choose a Hermitian connection, that is, a connection such that in an orthonormal basis the vector potential takes values in the Lie algebra of the unitary group $U(N)$. In that case the determinant $\det(1 + \frac{\lambda}{2\pi i}F)$ is real for any real parameter λ and the Chern classes are given by the expansion in powers of λ; the first two positive powers lead to

$$\begin{aligned} c_1(F) &= \frac{1}{2\pi i}\mathrm{tr}\, F \\ c_2(F) &= \frac{1}{2(2\pi i)^2}[-\mathrm{tr}\, F^2 + (\mathrm{tr}\, F)^2]. \end{aligned}$$

The coefficients in the expansion can be best computed by diagonalizing the matrix F. Writing $F = diag(\alpha_1, \ldots, \alpha_N)$ one obtains

$$\det\left(1 + \frac{\lambda}{2\pi i}F\right) = \prod_k \left(1 + \frac{\lambda\alpha_k}{2\pi i}\right) = \sum_n \left(\frac{\lambda}{2\pi i}\right)^n S_n(\alpha)$$

with

$$\begin{aligned} S_0 = 1,\, S_1 = \mathrm{tr}\,\alpha,\, S_2 &= \frac{1}{2}(\mathrm{tr}\,\alpha)^2 - \frac{1}{2}\mathrm{tr}\,\alpha^2 \\ S_3 &= \frac{1}{6}(\mathrm{tr}\,\alpha)^3 - \frac{1}{2}\mathrm{tr}\,\alpha^2\,\mathrm{tr}\,\alpha + \frac{1}{3}\mathrm{tr}\,\alpha^3 \end{aligned}$$

etc. Note that c_n vanishes identically if $n > \frac{1}{2}\dim M$ or $n > N$. If $n = \frac{1}{2}\dim M$ then we can integrate the form $c_n(E)$ over M and the value of the integral is called *the Chern number* associated to the vector bundle E.

Example 4.7.9. Consider a vector bundle E over $M = S^4$ such that the transition functions take values in the group $SU(N)$, $N \geq 2$. Dividing S^4 to the upper and lower hemispheres $S^4_\pm$ the bundle is given by the transition function ϕ along the equator S^3. The vector potentials $A_\pm$ are then related by $A_- = \phi A_+ \phi^{-1} - d\phi\phi^{-1}$ on the equator. Using the formula $\mathrm{tr}\, F^2 = d\mathrm{tr}\,(F\wedge A - \frac{1}{3}A^3)$ we can compute the Chern number corresponding to the second Chern class,

$$\begin{aligned}\frac{1}{8\pi^2}\int_{S^4_+}\mathrm{tr}\, F_+^2 + \frac{1}{8\pi^2}\int_{S^4_-}\mathrm{tr}\, F_-^2 \\ &= \frac{1}{8\pi^2}\int_{S^3}\left[\mathrm{tr}\,(F_+\wedge A_+ - \tfrac{1}{3}A_+^3) - \mathrm{tr}\,(F_-\wedge A_- - \tfrac{1}{3}A_-^3)\right] \\ &= \frac{1}{8\pi^2}\int_{S^3}\left[\mathrm{tr}\,\tfrac{1}{3}(d\phi\phi^{-1})^3 - d\mathrm{tr}\,(A_+\wedge d\phi\phi^{-1})\right] \\ &= \frac{1}{24\pi^2}\int_{S^3}\mathrm{tr}\,(d\phi\phi^{-1})^3.\end{aligned}$$

Remark 4.7.10. The value of the integral above is an integer which depends only on the homotopy class of the map $\phi: S^3 \to SU(N)$. This follows from the fact that the form $\mathrm{tr}\,(dgg^{-1})^3$ on any Lie group is closed (section 4.6) and from Stokes' theorem applied to the integral $\int_0^1 dt\frac{d}{dt}\int_{S^3}\mathrm{tr}\,(d\phi_t{\phi_t}^{-1})^3$ for a 1-parameter family of maps $\phi_t; S^3 \to SU(N)$.

Since the equivalence class of the bundle E depends only on the homotopy class of the transition function ϕ, the Chern number $\int c_2(E)$ gives a complete topological characterization of E.

The *Chern character* $ch(E)$ of a vector bundle is defined as follows. It is a formal sum of differential forms of different degrees,

$$ch(E) = \mathrm{tr}\ \exp\left(\frac{1}{2\pi i}F\right),$$

where again F is the curvature form of E. When the exponential is evaluated as a power series we obtain

$$ch(E) = \sum_{k=0}^{\infty}\frac{1}{(2\pi i)^k k!}\mathrm{tr}\, F^k.$$

Clearly all the terms can be expressed using the Chern classes; the three first terms are

$$ch(E) = N + c_1(E) + \frac{1}{2}c_1(E)\wedge c_1(E) - c_2(E) + \dots.$$

The Chern character is a convenient tool because one has

$$ch(E\oplus E') = ch(E) + ch(E') \quad ch(E\otimes E') = ch(E)\cdot ch(E').$$

This follows immediately from the definition and the elementary properties of the exponential function.

Above we have studied characteristic classes of complex vector bundles. The most important characteristic classes for real vector bundles are the *Pontrjagin classes* and they are constructed as follows.

Let $\pi : E \to M$ be a real vector bundle over the manifold M. We can always think a real vector bundle as an associated vector bundle to a principal bundle P with structure group $GL(n, \mathbb{R})$. We fix a metric in the fibers of E, so it makes sense to speak about orthonormal frames in the fibers. This means that we can consider E as an associated bundle to a principal $O(n)$ bundle; the principal bundle is simply the bundle of orthonormal frames.

Thus we are led to studying connections in principal $O(n)$ bundles. A connection form takes values in the Lie algebra of $O(n)$, that is, in the Lie algebra of real antisymmetric $n \times n$ matrices.

If we choose a local section of the principal $O(n)$ bundle then the curvature form F is a local matrix form on the base space M and in gauge transformations $F' = g^{-1}Fg$.

A real antisymmetric matrix can be brought to the canonical form

$$T^{-1}FT = \begin{pmatrix} 0 & \lambda_1 & 0 & \dots & 0 \\ -\lambda_1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_2 \dots & 0 \\ 0 & 0 & -\lambda_2 & 0 \dots & 0 \\ \dots\dots\dots & & & & \\ \dots\dots\dots & & & & \end{pmatrix}.$$

When $n = 2k$ is even the matrix consist of k antisymmetric 2×2 matrices on the diagonal; when $n = 2k+1$ then the last column and the last row consists of only zeros. The eigenvalues of the matrix are $\pm i\lambda_j$.

We define

$$p(F) = \det\left(1 + \frac{F}{2\pi}\right) = \prod_{i=1}^{k}\left(1 + \frac{\lambda_i^2}{4\pi^2}\right).$$

Clearly $p(F) = p(-F)$ so that p is a polynomial of even degree in the curvature tensor F. We write

$$p(F) = 1 + p_1(F) + p_2(F) + \dots$$

as a sum of homogeneous terms $p_j(F)$ of degree $2j$ in the curvature. Since F is a 2-form, $p_j(F)$ is a differential form on M of degree $4j$.

Note that $p_j(F)$ depends only on the eigenvalues of $F_{\mu\nu}$ and therefore it is invariant with respect to gauge transformations and thus gives a globally well-defined form on M.

Exactly as in the case of Chern classes we expand each p_j in powers of the curvature tensor F. The lowest Pontrjagin classes are

$$\begin{aligned}
p_1(F) &= -\frac{1}{2}(\frac{1}{2\pi})^2 \mathrm{tr}\, F^2 \\
p_2(F) &= \sum_{i<j} (\frac{\lambda_i}{2\pi})^2 (\frac{\lambda_j}{2\pi})^2 = \frac{1}{2}\left[(\sum_i \frac{\lambda_i^2}{(2\pi)^2})^2 - \sum_i (\frac{\lambda_i}{2\pi})^4 \right] \\
&= \frac{1}{8}(\frac{1}{2\pi})^4 \left[(\mathrm{tr}\, F^2)^2 - 2\mathrm{tr}\, F^4 \right] \\
p_3(F) &= \sum_{i<j<k} (\frac{\lambda_i}{2\pi})^2 (\frac{\lambda_j}{2\pi})^2 (\frac{\lambda_k}{2\pi})^2 = \dots \\
&= \frac{1}{48}(\frac{1}{2\pi})^6 \left[-(\mathrm{tr}\, F^2)^3 + 6\mathrm{tr}\, F^2 \cdot \mathrm{tr}\, F^4 - 8\mathrm{tr}\, F^6 \right].
\end{aligned}$$

We shall meet later another set of characteristic classes, called the *A-roof genus,* which are actually formed from the Pontrjagin classes. The definition is best set up in terms of eigenvalues of the matrix form F,

$$\hat{A}(F) = \prod_j \frac{x_j/2}{\sinh(x_j/2)} = \prod_j \left(1 + \sum_\ell (-1)^\ell \frac{2^{2\ell} - 2}{(2\ell)!} B_\ell \, x_j^{2\ell} \right),$$

where B_ℓ are the Bernoulli numbers and $x_j = \lambda_j / 2\pi$. In terms of Ponrjagin classes,

$$\hat{A}(F) = 1 - \frac{1}{24} p_1 + \frac{1}{5760}(7p_1^2 - 4p_2) + \frac{1}{967680}(-31p_1^3 + 44 p_1 p_2 - 16 p_3) + \dots.$$

Further reading: Nakahara, Chapters 9-11. The proof above of the topological invariance of the Chern classes follows S.S. Chern: *Complex Manifolds without Potential Theory.* Princeton University Press (1979). On characteristic classes see also: J.W. Milnor and J.D. Stasheff: *Characteristic Classes.* Princeton University Press (1974).

CHAPTER 5: YANG-MILLS THEORY

5.1 Yang-Mills equations.

Let M be a Riemann manifold with Riemann metric g. In local coordinates the metric is represented as a symmetric nondegenerate tensor field $g_{ij}(x)$ with $i, j = 1, 2, \ldots, n$, where $n = \dim M$. Let $\pi : P \to M$ be a principal G bundle over M. Let $\rho : G \to Aut(V)$ be a unitary finite-dimensional representation of G in V. This defines an associated vector bundle $E = P \times_\rho V$ and the curvature tensor F of a connection in P is represented (locally) by matrix functions $F_{ij}(x) = \partial_i A_j - \partial_j A_i + [A_i, A_j]$ acting on vectors in V.

We shall define raising and lowering of space-time indices (i.e., coordinate indices in M) as usual, $A^i = g^{ij} A_j, B_i = g_{ij} B^j$, where the matrix (g^{ij}) is the inverse of (g_{ij}). Recall also that the metric g defines a volume form on M, $d(vol_M) = \sqrt{\det(g)} dx_1 \wedge dx_2 \cdots \wedge dx_n$. We define the *Yang-Mills functional*

$$Y(A) = \frac{1}{4} \int_M \operatorname{tr} F_{\mu\nu} F^{\mu\nu} d(vol_M).$$

The Yang-Mills action is invariant under gauge transformations $F' = g^{-1} F g$. There is an alternative way to write the YM action as

$$Y(A) = -\frac{1}{2} \int_M \operatorname{tr} F \wedge *F.$$

The action leads to field equations through Euler-Lagrange variational priciple. Let $A + tB$ be a 1-parameter family of vector potentials:

$$\frac{d}{dt} Y(A + tB)|_{t=0} = \frac{1}{2} \int_M \operatorname{tr} F^{\mu\nu} (\partial_\mu B_\nu - \partial_\nu B_\mu + [A_\mu, B_\nu] + [B_\mu, A_\nu]) d(vol_M).$$

When M is a manifold without boundary, we can integrate by parts and we get

$$\delta Y(A) = -\int_M \operatorname{tr} B_\nu (\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}]) d(vol_M).$$

If A is an extremal the YM action then we obtain *the Yang-Mills equations*

$$D_\mu F^{\mu\nu} = \partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0.$$

When G is abelian this gives the Maxwell's equations $\partial_\mu F^{\mu\nu} = 0$ in vacuum. In addition, we have the Bianchi identities

$$D_\mu F_{\nu\lambda} + D_\lambda F_{\mu\nu} + D_\nu F_{\lambda\mu} = 0$$

for all indices λ, μ, ν. If there are external sources we have instead

$$D_\mu F^{\mu\nu} = j^\nu$$

for some Lie algebra valued current j^ν.

The Yang-Mills equations is a complicated nonlinear system of second order partial differential equations. Not much is known about the general solutions. However, there is a class of solutions which is well understood. These so-called (anti) instantons are characterized by the (anti) self-duality property $F = *F$ ($F = -*F$) in the case of a Riemannian 4-manifold M. Recall that

$$* : \Omega^k(M) \to \Omega^{n-k}(M)$$

is a linear map and $** = \pm 1$. When $n = 4$ and $k = 2$ the sign is $+$ (exeercise) For this reason the eigenvalues of $*$ are ± 1, when restricted to 2-forms on a 4-manifold. In the case of Lorentzian metric $** = -1$ on 2-forms and therefore in this case there are no real eigenvalues (and no real (anti) instantons).

In the case of an instanton we have

$$Y(A) = -\frac{1}{2} \int_M \mathrm{tr}\, F \wedge F$$

and so the value of the YM functional is given by the second Chern class. In particular, when $M = S^4$ we get

$$Y(A) \sim \int_{S^3} \mathrm{tr}\, (g^{-1} dg)^3,$$

where $g : S^3 \to G$ is the transition function on the equator. Thus for self-dual solutions the YM functional is quantized in units $(2\pi)^2 k$ with $k \in \mathbb{Z}$.

5.2 Dirac equation For each positive integer n we construct a set of complex $2^{[n/2]} \times 2^{[n/2]}$ matrices γ_i, $i = 1, 2, \dots n$, with the relations

$$\gamma_{\gamma j} + \gamma_j \gamma_i = 2\eta_{ij},$$

where η is either the Minkowski or the Euclidean metric. Here $[x]$ is the integer part of a real number x. The matrices are constructed by induction on n. For $n = 1$ there is only one 1×1 matrix $\gamma_1 = 1$. The induction from odd to even n is as follows. The dimension of the matrices is increased by a factor 2. In the case of Euclidean metric we set

$$\gamma_i \mapsto \begin{pmatrix} 0 & \gamma_i \\ \gamma_i & 0 \end{pmatrix}$$

for $i = 1, 2, \ldots, n$ and we add

$$\gamma_{n+1} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

where the blocks are $2^{(n-1)/2} \times 2^{(n-1)/2}$ matrices. The induction from even n to $n + 1$ is defined by adding the matrix

$$\gamma_{n+1} = (-1)^{n/4} \gamma_1 \gamma_2 \ldots \gamma_n$$

of same dimension.

In the case of the Minkowski metric there will be some sign changes. In the induction from odd to even $n = 2k$ we have

$$\gamma_i \mapsto \begin{pmatrix} 0 & \gamma_i \\ -\gamma_i & 0 \end{pmatrix}$$

where γ_i for $i = 1, 2, \ldots, 2k-1$ are the Euclidean $\gamma-$ matrices in $2k-1$ dimensions, and the 'time like' matrix is defined as

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In the induction from even $n = 2k$ to $2k + 1$ one can take γ_j for $j = 1, 2, \ldots, 2k$ as $i\times$ the Euclidean $\gamma-$ matrices in $2k$ dimensions and then set $\gamma_0 = e^{i\alpha} \gamma_1 \cdots \gamma_{2k}$ for an appropriate phase factor $e^{i\alpha}$.

The Dirac equation in the flat space $\mathbb{R}^n$ is then

$$(i\gamma^\mu \partial_\mu + m)\psi = 0,$$

where $\psi : \mathbb{R}^n \to \mathbb{C}^N$, with $N = 2^{[n/2]}$ and m is a constant which is interpreted as a mass of the field ψ. Because of the anticommutation relations of the $\gamma-$ matrices any solution of the Dirac equation satisfies also the Klein-Gordon equation

$$(\partial^\mu \partial_\mu + m^2)\psi = 0.$$

Using Fourier transform we can look for solutions of the form $u \cdot e^{-ip_\mu x^\mu}$, for some constant vector $u \in \mathbb{C}^N$ and a momentum vector $p \in \mathbb{R}^n$. The Dirac equation becomes now

$$(\gamma^\mu p_\mu + m)u = 0.$$

The Klein-Gordon equation requires that $-p^2 + m^2 = 0$.

Exercise Show that for $n = 4$, Minkowski metric, and for each momentum vector $p^2 = m^2$ there are exactly two linearly independent solutions of the matrix equation $(\gamma^\mu p_\mu + m)u = 0$.

In a curved space the Dirac equation needs a modification. First, we have to assume that the space M has a *spin structure.* This is defined as follows. We shall discuss the case of a Riemann metric. To begin with, the rotation group $SO(n)$, for $n > 2$, has a simply connected double covering $Spin(n)$. That is, there is a onto 2-1 group homomorphism $\phi : Spin(n) \to SO(n)$. We are looking for a principal bundle $P \to M$ with structure group $Spin(n)$ such that there is a 2-1 onto map $\theta : P \to LM$, where LM is the bundle of orthonormal oriented frames in the tangent bundle TM. The structure group of the principal bundle $LM \to M$ is $SO(n)$. The map θ should take the fiber P_x to the fiber $L_x M$ for each $x \in M$ and we require that

$$\theta(pg) = \theta(p)\phi(g)$$

for all $p \in P$ and $g \in Spin(n)$. Such a bundle P is called a spin structure on M. Not every manifold has a spin structure and if there is a spin structure it does not need to be unique.

We shall from now on assume that M is a Riemannian spin manifold with a fixed spin structure.

The group $Spin(n)$ has a (faithful) representation in $\mathbb{C}^N$ for $N = 2^{[n/2]}$. The Lie algebra of $Spin(n)$ is isomorphic with the Lie algebra of $SO(n)$ since the groups are locally isomorphic. The Lie algebra $Lie(SO(n))$ consists of real antisymmetric $n \times n$ matrices and it is spanned by matrices $L_{ij} = -L_{ji}$ with commutation relations

$$[L_{ij}, L_{kl}] = \delta_{jk} L_{il} + \delta_{il} L_{jk} - \delta_{ik} L_{jl} - \delta_{jl} L_{ik}.$$

The representation in $\mathbb{C}^N$ is obtained by the mapping $L_{ij} \mapsto M_{ij} = \frac{1}{4}[\gamma_i, \gamma_j]$. One can then check by direct computation, using the anticommutation relations of

γ-matrices, that the matrices M_{ij} satisfy the same commutation relations as the matrices L_{ij} and thus we indeed have a representation of the Lie algebra $Lie(Spin(n))$ in $\mathbb{C}^N$. The group $Spin(n)$ is simply connected and for this reason the representation of its Lie algebra can be exponentiated to give a representation of the group $Spin(n)$. Denote this representation by ρ.

We can now define the *Dirac spinor bundle* S as the associated bundle $S = Spin(M) \times_\rho \mathbb{C}^N$. Sections of this vector bundle are *Dirac spinor fields.* We define a covariant derivative ∇_μ acting on Dirac spinor field. First, let e_a with $a = 1, 2, \ldots, n$ be a local oriented orthonormal basis in the tangent bundle TM. Then

$$\nabla_\mu e_a = \Gamma^b_{\mu a} e_b$$

defines the Christoffel symbols in the basis e_a. The Christoffel symbols can be computed from the fact that $e_a = e^\mu_a \partial_\mu$, and so

$$\nabla_\mu e_a = (\partial_\mu e^\nu_a)\partial_\nu + e^\nu_a \nabla_\mu \partial_\nu = (\partial_\mu e^\nu_a) e^b_\nu e_b + e^\nu_a \Gamma^\lambda_{\mu\nu} e^b_\lambda e_b,$$

where (e^a_μ) is the inverse to the matrix (e^μ_a). Inserting

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha}(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu})$$

we get an explicit expression for the symbols $\Gamma^b_{\mu a}$. We set

$$\omega_\mu = \frac{1}{2}\Gamma^b_{\mu a} M_{ab}.$$

The Dirac equation on a curved manifold is then

$$i\gamma^\mu(\partial_\mu + \omega_\mu)\psi + m\psi = 0.$$

Here we have defined $\gamma^\mu = e^\mu_a \gamma^a$. They satisfy the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu},$$

whereas

$$\gamma_a\gamma_b + \gamma_b\gamma_a = 2\delta_{ab}.$$

Let E be a complex vector bundle over M with a connection, described locally by a vector potential A_μ. Consider the tensor product bundle $S \otimes E$. The Dirac equation for sections of the extended bundle is

$$i\gamma^\mu(\partial_\mu + \omega_\mu + A_\mu)\psi + m\psi = 0.$$

Here ω_μ acts on the first factor and A_μ on the second factor in the tensor product. To be precise, we should write $A_\mu \mapsto 1 \otimes A_\mu$ and $\omega_\mu \to \omega_\mu \otimes 1$.

We also need to fix a hermitean inner product $< \cdot, \cdot >_x$ in the fibers $(S \otimes E)_x$. We can then define the Hilbert space $H = L_2(S \otimes E)$ of square integrable sections of the bundle $S \otimes E \to M$. The inner product for a pair of sections ψ, ϕ is

$$< \psi, \phi >= \int_M < \psi(x), \phi(x) >_x d(vol_M).$$

Assuming that the gauge group is unitarily represented (the Lie algebra elements $A_\mu(x)$ are antihermitean matrices), in the case of Riemann metric the Dirac operator is self-adjoint (in an appropriate dense domain) in the Hilbert space H.

5.3 The index of the Dirac operator

Let $T : H \to H$ be a linear operator in a Hilbert space H. We set $\ker T = \{x \in H | Tx = 0\}$ and $\operatorname{coker} T = (TH)^\perp$. When $\ker T$ and $\operatorname{coker} T$ are finite dimensional then T is a *Fredholm operator* and its Fredholm index is

$$\operatorname{ind} T = \dim \, \ker T - \dim \, \operatorname{coker} T.$$

Example Let $T : H \to H$ be defined as $Te_n = e_{n+1}$, where $\{e_n\}_{n=1,2,\ldots}$ is an orthonormal basis. Now $\ker T = 0$ and $\operatorname{coker} T$ consists of $\mathbb{C} \cdot e_1$. Thus $\operatorname{ind} T = -1$.

If T is a Fredholm operator then $\operatorname{ind} T = -\operatorname{ind} T^*$. This follows from $< y, T^* x >=< Ty, x >$ and so $T^* x = 0$ if and only if $< Ty, x >= 0$ for all y, which is equivalent to $x \in (TH)^\perp$ and so $\ker T^* = \operatorname{coker} T$. In particular, when T is self-adjoint the Fredholm index is zero.

Let us study the case of the Dirac operator on an even dimensional manifold M. In this case we have an hermitean operator Γ with $\Gamma^2 = 1$ which anticommutes with D. The *chirality operator* Γ is defined as $\gamma_{n+1} = e^{i\alpha} \gamma_1 \ldots \gamma_n$ for an appropriate phase factor $e^{i\alpha}$.

We observe that the nonzero eigenvalues of D come in pairs $\pm\lambda$ because of $D(\Gamma\psi) = -\Gamma D\psi = -\lambda(\Gamma\psi)$ if $D\psi = \lambda\psi$.

The case of the zero eigenvalue is different. We can split the kernel of D to a pair of subspaces $\ker D = V_- \oplus V_+$ where $V_\pm$ are defined by diagonalizing Γ; the eigenvalues of Γ are ± 1. In general, the dimensions of $V_\pm$ are different. However, in the case of a compact manifold M all eigenspaces of D are finite dimensional and we can define the index $\mathrm{dim} V_+ - \mathrm{dim} V_-$. This is in fact the Fredholm index of a certain operator. Diagonalizing Γ we can write

$$D = \begin{pmatrix} 0 & D_+ \\ D_- & 0 \end{pmatrix},$$

where D_+ maps to eigenspace of Γ corresponding to the eigenvalue -1 to the eigenspace $+1$ and vice-versa for D_-.

We have now

$$\mathrm{ind} D_- = \dim V_+ - \dim V_-.$$

We have $D_+ = D_-^*$ and so $\mathrm{ind}\, D_+ = -\mathrm{ind}\, D_-$.

In functional analysis one proves that for bounded operators the index of a Fredholm operator is a continuous function (in operator norm) of the operator. For this reason the index is a *topological invariant.* It remains constant under continuous deformations of the operator. Although a Dirac operator is always unbounded, one can still show that its index is a continuous function of the parameters: vector potentials, choice of Riemann metric etc. For this reason the index $\mathrm{ind}\, D_-$ depends only on the homotopy class of the vector bundles S, E and defines a topological invariant for the bundles under consideration. On the other hand, we have explicitly used a metric and a vector potential in the construction of D.

We already know that there are characteristic classes (Chern classes, Pontrjagin classes) which are topological invariants. For this reason it is not so big surprize that the index can be expressed in terms of these classes.

Theorem. *(Atiyah-Singer)*

$$ind\, D_+ = \int_M \hat{A}(TM) \wedge ch(E).$$

We cannot prove this theorem here but instead we illustrate the philosophy behind the index theorems by an explicit calculation.

Let $H = H_+ \oplus H_-$ be a sum of two infinite-dimensional complex Hilbert spaces and denote by P_+ the projection on H_+. Let $u : H \to H$ be an invertible operator such that $P_+ u P_+$ is a Fredholm operator. We also assume that the trace of $[P_+, u]$ is absolutely converging. We claim that

$$\mathrm{ind}(P_+ u P_+) = \mathrm{tr}\, u^{-1}[u, P_+].$$

To prove this index theorem we need to check that the formula holds for some selection of operators u_n with $\operatorname{ind} P_+ u_n P_+ = n$ for $n \in \mathbb{Z}$. This is sufficient by the continuity of the index! So we may choose $H = L_2(S^1)$ and H_+ is defined by nonnegative Fourier modes and H_- by negative Fourier modes. We select u_n as the multiplication operator by the Fourier mode e^{-inx}. Then it is easy to see that $\operatorname{ind} P_+ u_n P_+ = n$. On the other hand u_n^{-1} is the multiplication operator by the function e^{inx} and by a simple computation $\operatorname{tr} u^{-1}[u, P_+] = n$.