The Geometry of Gauge Theory

Christopher La Fond

December 12, 2023

Acknowledgements

This work began out of a desire to write a paper about physics in the language of geometry. Since I wanted it to be largely self contained, it has turned into a much larger project than I could ever have imagined. That being said, without the help of many, I would never have been able to complete this work.

In particular, I would like to thank my advisor, Matt Szczesny, for his patience, and invaluable help throughout the course of the writing process. Furthermore, I am incredibly thankful for all the help that Mark JD Hamilton has provided me via email. I also thank my friends, Sid, Annie, and Malachi, for our long conversations about the topics in this paper, and for their tremendous help in working out the kinks in certain proofs, and in organizing my thoughts.

Finally, I would be remiss if I did not mention my family, friends, and girlfriend, who provided me with encouragement, and emotional support throughout the writing process.

Contents

Introduction

The core idea behind gauge theory was first stumbled upon in the context of classical electromagnetism. Given an electric potential $V(\mathbf{x}, t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$, and a vector potential $\mathbf{A}(\mathbf{x}, t) : \mathbb{R}^3 \times \mathbb{R} \to T\mathbb{R}^3$, the physical fields $\mathbf{E}$, and $\mathbf{B}$ corresponding to these potentials are given by the following relations:

$$\begin{aligned} \mathbf{E} =& -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} =& \nabla \times \mathbf{A} \end{aligned}$$

Playing around with these relations, one finds that given any smooth function $\lambda : \mathbb{R}^4 \to \mathbb{R}$, the new potentials:

$$V' = V - \frac{\partial \lambda}{\partial t} \qquad \text{and} \qquad \mathbf{A}' = \mathbf{A} + \nabla \lambda$$

correspond to the same physical fields $\mathbf{E}$ and $\mathbf{B}$. In electromagnetism, this invariance under such a transformation $(V, \mathbf{A}) \to (V', \mathbf{A}')$, is called gauge freedom, or gauge invariance. The idea of objects in physics being invariant under a certain group of transformations was further expanded on, and generalized by the physicist's Yang and Mills in their 1954 paper *Conservation of Isotopic Spin and Isotopic Gauge Invariance*. Nowadays, the best model of the universe that physics has to offer is a gauge theory, namely the Standard Model, which has held up remarkably well under experimental verification. All of this then begs the question, what exactly *is* a gauge theory?

Physically, a gauge theory is a field theory with a symmetry that can be attributed to a symmetry group acting on fields, while leaving the Lagrangian invariant; in classical electromagnetism this symmetry group is the Lie group $U(1)$, while in the original Yang-Mills theory, as presented in their 1954 paper, it is the Lie group $SU(2)$, and finally in the Standard Model it is the product Lie Group $U(1) \times SU(2) \times SU(3)$. Mathematically, these types of field theories are intimately related to Cartan's study of principal bundles, and connections on said bundles. Indeed, with a proper background in Cartan's geometry, one could read Yang and Mills paper and come away with the conclusion that they must have been aware of the underlying geometry in their work, though this was certainly not the case. Furthermore, the mathematical study of gauge theory, has proved incredibly fruitful in producing deep geometric and topological results, such as the existence of multiple distinct smooth structures on $\mathbb{R}^4$.

The goal of this paper is to understand classical gauge theories from a geometric perspective, and is split into three main parts, each consisting of two chapters. The first part gives a brief primer on differential topology and Lie theory. Specifically, chapter 1.1 includes a brief introduction to smooth manifolds, and the various objects one encounters on them such as vector fields, differential forms, and (pseudo)-Riemannian metrics. In chapter 1.2, we introduce Lie groups, Lie algebras, group actions on manifolds, and the representation theory of Lie group and Lie algebras.

The second part is focused on mathematical gauge theory and spinors. Chapter 2.1 introduces the main objects of study: the principal bundles, and connections characteristic of Cartan's geometry. We then go on to discuss associated vector bundles, gauge transformations, curvature, and covariant derivatives. In chapter 2.2, we develop Clifford algebras, and the Spin group, with the goal of studying a very special type of associated vector bundle: the spinor bundle.

The final part is dedicated to Yang-Mills theory and applications to physics. In chapter 3.1, we will introduce the necessary geometric constructions to define the Yang-Mill's Lagrangian, and then derive the Yang-Mills equation. The rest of the chapter will be focused on demonstrating how this equation, and the Bianchi identity, relate to classical electromagnetism, and how to then modify the Yang-Mills Lagrangian to incorporate various types of sources into our theory of electromagnetism. In particular, we will the give classical descriptions of two quantum field theories: Scalar Electrodynamics, and Quantum Electrodynamics, the latter of which incorporates fermionic sources.

Background

1.1 Smooth Manifolds

Smooth manifolds are generalizations of smooth curves and surfaces in $\mathbb{R}^3$ to higher dimensions. There are two main ingredients baked into the definition of a smooth manifold, namely a specific topology that allows us to locally identify the manifold with $\mathbb{R}^n$, and a smooth structure, which allows us to perform calculus on the space. Once these two ingredients are well understood, we are in a position two study these objects in depth with a wide variety of tools from analysis and algebra. In this chapter, we give a brief overview of the definitions, objects, and operations necessary for building up gauge theory. We closely follow the treatment found in the text Lee's *Smooth Manifolds*.

1.1.1 Topological Manifolds and Smooth Structures

As mentioned before, a (topological) manifold is a set, endowed with a topology that allows us locally identity it with $\mathbb{R}^n$. In order to make this precise we employ the following definition:

Definition 1.1.1. A topological space M is a **topological manifold** of dimension n if the following conditions hold:

- M is Hausdorff, i.e. for every points $p, q \in M$ there exist disjoint open sets $U_p, U_q \subset M$, such that $p \in U_p$ and $q \in U_q$.
- M is second countable, i.e. M has a countable basis for it's topology.
- M is locally Euclidean of dimension n, i.e. for every point $p \in M$ there exists and open neighborhood U_p of p, such that U_p is homeomorphic to some open subset of $\mathbb{R}^n$.

It is easy to see that in the third condition, we could equivalently require U_p to be homeomorphic to an open ball in $\mathbb{R}^n$, as given a homeomorphism $\phi : U_p \to U \subset \mathbb{R}^n$, we can find an open ball $B_{\phi(p)} \subset U$ centered at $\phi(p)$. The restriction of ϕ to the inverse image $\phi^{-1}(B_{\phi(p)})$ then maps homeomorphicly to an open ball in $\mathbb{R}^n$. Consequently, as open balls in $\mathbb{R}^n$ are homeomorphic in $\mathbb{R}^n$, we could equivalently require U_p to homeomorphic to $\mathbb{R}^n$ itself.

Furthermore, the third condition allows us to obtain the notion of a coordinate chart:

Definition 1.1.2. A **coordinate chart** on a topological manifold M, is a pair (U, ϕ), where $U \subset M$ is open, and $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a homeomorphism.

Clearly, every point $p \in M$ is contained in some coordinate chart, and we say ϕ is centered at p if $\phi(p) = 0$. We call the component functions $\phi(p) = (x^1(p), \dots, x^n(p))$ local coordinates on U. Though cumbersome, at times it is quite useful to define the component of functions of ϕ for some U_p, and 'work in coordinates'. Trivially any space homeomorphic to $\mathbb{R}^n$ is a topological manifold; let us now look at a non trivial example of a topological manifold:

Example 1.1.1. The 2-sphere, denoted $\mathbb{S}^2$ defined as the set:

$$\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$$

where $\|x\|$ denotes the standard Euclidean norm on $\mathbb{R}^3$, is a topological manifold. Endowed with the subspace topology from $\mathbb{R}^3$, $\mathbb{S}^2$ is clearly second countable and Hausdorff. We then must show it is locally Euclidean; consider the following parameterization of $\mathbb{S}^2$:

$$\psi^{-1}(u, v) = \left(\frac{2u}{1 + u^2 + v^2}, \frac{2v}{1 + u^2 + v^2}, \frac{u^2 + v^2 - 1}{1 + u^2 + v^2} \right)$$

It easy to convince oneself that this is a parametrization for $\mathbb{S}^2$ by checking that $\|\psi^{-1}(u, v)\| = 1$ for any u, v. Furthermore have that for a point $p = (x, y, z) \in \mathbb{S}^2$, ψ is given by:

$$\psi(x, y, z) = \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right)$$

ψ is then a homeomorphism from $\mathbb{S}^2 \setminus (0, 0, 1) \to \mathbb{R}^2$, i.e. every point $p \in \mathbb{S}^2$ is contained in the defined by ψ. Consider now the following parameterization instead:

$$\phi^{-1}(u, v) = \left(\frac{2u}{1 + u^2 + v^2}, \frac{2v}{1 + u^2 + v^2}, \frac{1 - u^2 - v^2}{1 + u^2 + v^2} \right)$$

with inverse:

$$\phi(x, y, z) = \left(\frac{x}{1+z}, \frac{y}{1+z}\right)$$

ϕ is then a homeomorphism from $\mathbb{S}^2 \setminus (0,0,-1) \to \mathbb{R}^2$. The coordinate charts $(\mathbb{S}^2 \setminus (0,0,-1), \phi)$ and $(\mathbb{S}^2 \setminus (0,0,1), \psi)$ cover $\mathbb{S}^2$, and thus every point in $\mathbb{S}^2$ is contained in an open set U that is homeomorphic to an open set in $\mathbb{R}^2$, or rather $\mathbb{R}^2$ itself. We have therefore shown that $\mathbb{S}^2$ is locally Euclidean, hence a topological manifold of dimension 2.[1]

We now turn to defining a notion of 'smoothness' on a topological manifold M. Recall that given open sets $U \subset \mathbb{R}^n$, $V \in \mathbb{R}^m$, a function $F : U \to V$ is smooth if each of the component functions are *smooth*, i.e. every component function has continuous partial derivatives of all orders. Further, if $m = n$, and F is a bijection with smooth inverse, we call F a *diffeomorphism* from U to V. Consequently, if F is a diffeomorphism, it is also a homeomorphism. Given two charts (U, ϕ) and (V, ψ) for a topological manifold M, such that $U \cap V \neq \emptyset$, we say they are *smoothly compatible* if the *transition function*:

$$\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$$

is a diffeomorphism. A collection of smoothly compatible charts $\mathcal{A}$ which cover M is called a *smooth atlas* for M, and is said to be maximal if $\mathcal{A}$ is not contained in any other smooth atlas. A *smooth structure* on M is a maximal smooth atlas $\mathcal{A}$.

Definition 1.1.3. A **smooth manifold** M is a pair $(M, \mathcal{A})$, where M is a topological manifold and $\mathcal{A}$ is a smooth structure.

A topological manifold M has many distinct smooth structures, however these distinct smooth structures are usually only unique up to diffeomorphism, though counter examples do exist such as $\mathbb{R}^4$, and $\mathbb{S}^7$. Furthermore, some topological manifolds do not admit any smooth structure. Going forward, we assume all manifolds we work with are smooth. Let us now examine the following examples:

Example 1.1.2. The Euclidean space $\mathbb{R}^n$ for all $n \in \mathbb{N}$ is a smooth manifold. The standard smooth structure on $\mathbb{R}^n$ is determined by the atlas consisting of the chart $(\mathbb{R}^n, \mathrm{Id}_n)$. Including all such charts that are smoothly compatible with Id_n (i.e. smooth local coordinate transformations) we obtain a maximal smooth atlas for $\mathbb{R}^n$. In this paper, when working in $\mathbb{R}^n$ we always assume the standard smooth structure.

Example 1.1.3. Let V be a finite dimensional vector space over $\mathbb{R}$ of dimension n. A norm on V induces a topology such that scalar multiplication $V \times \mathbb{R} \to V$, and vector addition $V \times V \to V$ are both continuous functions, where $V \times V$ and $V \times \mathbb{R}$ have the product topology. Furthermore, this topology is independent of the choice of norm. It then follows that any isomorphism $T : W \to W$ is also a homeomorphism as T is a continuous bijection with continuous inverse. In particular, a choice of basis $\{e_i\}$ for V admits the following vector space isomorphism $T : \mathbb{R}^n \to V$:

$$T(x) = x^i e_i$$

where we have employed the Einstein summation notation. Thus we can view V as a topological manifold since it is homeomorphic to $\mathbb{R}^n$, with chart (V, T^{-1}). Given any other basis $\{f_i\}$ of V, with with chart (V, S^{-1}) there exists an invertible linear transformation A, such that:

$$e_i = A_i^j f_j$$

The transition function for the two charts is then given by:

$$S^{-1} \circ T(x) = S^{-1}(x^i e_i) = S^{-1}(x^i A_i^j f_j) = A_i^j x^i$$

Hence the transition function is an invertible linear map, and thus a diffeomorphism. The collection of all such charts determines a smooth structure, and thus every finite dimensional vector space is a smooth manifold.

[1]This example can be generalized to a sphere of n dimensions, defined as the set $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$

Example 1.1.4. Continuing from **Example 1.1.1**, we calculate the transition function of the charts $(\mathbb{S} \setminus (0,0,-1), \phi)$ and $(\mathbb{S} \setminus (0,0,1), \psi)$:

$$\phi \circ \psi^{-1} = \left(\frac{u}{u^2+v^2}, \frac{v}{u^2+v^2} \right)$$

which is a diffeomorphism $\mathbb{R}^2 \setminus (0,0) \to \mathbb{R}^2 \setminus (0,0)$, hence the aforementioned charts determine a smooth structure, and thus the two sphere is a smooth manifold.

Example 1.1.5. Given a smooth manifold n dimensional manifold M, any open subset $U \subset M$ is also a smooth manifold, namely it is an *open submanifold.* Any open set $U \subset M$ inherits the subspace topology from M, making it into a topological manifold of dimension of n. We can then define a smooth structure on U by removing the charts (V, ψ) from $\mathcal{A}_M$ such that $V \cap U = \emptyset$, and restricting the rest to $V \cap U$. This atlas, $\mathcal{A}_U$, then covers U, and determines a smooth structure on U.

Example 1.1.6. Given n smooth manifolds $M_1, \dots, M_n$, the product space $M_1 \times \cdots M_n$ is a smooth manifold of dimension of $\dim(M_1) + \cdots + \dim(M_n)$. As a topological space, the product is Hausdorff and second countable, so we need only check that the locally Euclidean property, and then determine a smooth structure. Given point $(p_1, \dots, p_n) \in M_1 \times \cdots \times M_n$ we choose a coordinate (U_i, ϕ_i) around each $p_i \in U_i$. The product map:

$$\begin{aligned} \phi_1 \times \cdots \times \phi_n : U_1 \times \dots U_i &\longrightarrow \mathbb{R}^{\dim(M_1)} \times \cdots \mathbb{R}^{\dim(M_n)} \\ (p_1, \dots, p_n) &\longmapsto (\phi_1(p_1), \dots, \phi_n(p_n)) \end{aligned}$$

is then a homeomorphism onto it's image, which is a product open subset of $\mathbb{R}^{\dim(M_1)} \times \cdots \mathbb{R}^{\dim(M_n)}$, hence $M_1 \times \cdots \times M_n$ is a topological manifold. Take each smooth atlas $\mathcal{A}_i$ and select charts which cover each M_i, then construct product charts from these charts as above; each product chart is clearly smoothly compatible with one another, and the set of all such smoothly compatible product charts covers $M_1 \times \cdots \times M_n$, thus these charts determine a smooth structure on the product manifold, therefore $M_1 \times \cdots \times M_n$ is a smooth manifold.

1.1.2 Smooth Maps, Tangent Vectors, and Vector Fields

We now wish to develop of a notion of smoothness for maps between two manifolds M and N. To do so, we employ the following definition:

Definition 1.1.4. Given m and n dimensional manifolds M and N, a map $F : M \to N$ is a **smooth map** if for all $p \in M$ there exist smooth charts (U_p, ϕ) containing p, and smooth charts $(U_{F(p)}, \psi)$ containing $F(p)$ such that the composition $\psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(U_{F(p)})$ is smooth.

Furthermore, we define the following:

Definition 1.1.5. A **diffeomorphism** is a smooth map $F : M \to N$, such that F is bijective, with smooth inverse.

In particular, the coordinate charts (U, ϕ) for an n-manifold, and their inverses, are diffeomorphisms from open sets in M to open sets in $\mathbb{R}^n$, and vice versa. Given this notion of smoothness on a manifold, we now motivate 'derivatives' on a manifold by first considering tangent vectors.

Consider the real vector space $\mathbb{R}^n$; given a point $x \in \mathbb{R}^n$ we denote the *tangent space* to $\mathbb{R}^n$ at the point x as $T_x\mathbb{R}^n = \{x\} \times \mathbb{R}^n$. Clearly, the tangent space is itself a vector space isomorphic to $\mathbb{R}^n$, and can geometrically be pictured as arrows starting at the point x and pointing in the direction of some vector v. This also allows us to write the directional derivative of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ at the point x; let $v_x \in T_x R$, then the directional derivative, $D_x f : C^\infty(\mathbb{R}^n) \to \mathbb{R}$, at the point x in the direction v_x is given by:

$$D_x f(v) = \lim_{t \to 0} \frac{d}{dt} f(x + tv) = v^i \frac{\partial f}{\partial x^i}(x) \tag{1.1.1}$$

where the last equality operates under the assumption that in the standard basis for $\mathbb{R}^n$, $v_x \in T_x\mathbb{R}^n$ is given by $v_x = v^i e_i$. By the properties of the derivative, this map is $\mathbb{R}$-linear and satisfies the Leibniz rule, i.e. for $f, g \in C^\infty(\mathbb{R}^n)$, and $a, b \in \mathbb{R}$, the following hold:

$$\begin{aligned} D_x(af + bg)(v) &= aD_x f(v) + bD_x g(v) \\ D_x(fg)(v) &= fD_x g(v) + gD_x f(v) \end{aligned}$$

With the motivation of (1.1.1), we identify the standard basis vectors $\{e_i\}$ for $\mathbb{R}^n$ with the partial derivative operators:

$$v_x \in T_x\mathbb{R}^n = v^i e_i|_x = v^i \left.\frac{\partial}{\partial x^i}\right|_x$$

where $v^i \in \mathbb{R}$. Thus we obtain the following notational conveniences:

$$v_x(f) = v^i \frac{\partial f}{\partial x^i}(x) = D_x f(v)$$

Therefore, the object $v_x \in T_x\mathbb{R}^n$ is a map $C^\infty(\mathbb{R}^n) \to \mathbb{R}$ that is $\mathbb{R}$-linear, and satisfies the Leibniz rule; we call such and object a *derivation at a point*. Additionally, $T_x\mathbb{R}^n$ with this identification is clearly a vector space under the following operations:

$$\begin{aligned}(v_x + w_x)(f) =& v_x(f) + w_x(f) \in T_x\mathbb{R}^n \\ (cv_x)(f) =& c \cdot v_x(f)\end{aligned}$$

where $c \in \mathbb{R}$.

We can now adequately define tangent vectors on manifolds in the following way:

Definition 1.1.6. Given an n dimensional smooth manifold M, the **tangent space**, T_pM, is the $\mathbb{R}$-linear vector space of all derivations at the point $p \in M$, $v_p : C^\infty(M) \to \mathbb{R}$, i.e. for $f, g \in C^\infty(M)$, $a, b \in \mathbb{R}$, and $v_p \in T_pM$ we have:

$$\begin{aligned}v_p(fg) =& f(p)v_p(g) + g(p)v_p(f) \\ v_p(af + bg) =& av_p(f) + bv_p(g)\end{aligned}$$

and, for $v_p, w_p \in T_pM$, $c \in \mathbb{R}$:

$$\begin{aligned}(v_p + w_x)(f) =& v_x(f) + w_x(f) \in T_pM \\ (cv_x)(f) =& c \cdot v_x(f) \in T_P M\end{aligned}$$

We call $v_p \in T_pM$ a **tangent vector**.

In the same way that we visualized the tangent vectors in $T_x\mathbb{R}^n$ as arrows tangent to $\mathbb{R}^n$, originating at the point x, we visual tangent vectors in T_pM, as arrows tangent to M, starting at the point p. As a consequence of the preceding definition, we obtain the following lemma:

Lemma 1.1.1. *Suppose M is a smooth manifold, $p \in M$, $v_p \in T_pM$, and $f, g \in C^\infty(M)$. The following properties hold:*

- *if f is a constant function, then $v_p(f) = 0$.*
- *if $f(p) = g(p) = 0$, then $v(fg) = 0$*

Proof. For all $p \in M$, let $f(p) = c$ for some $c \in R$. Then we have:

$$v_p(f \cdot f) = 2f(p)v_p(f) \Rightarrow c^2 v_p(1) = 2c^2 v_p(1) \Rightarrow v_p(1) = 0$$

Since $v_p(f) = cv_p(1)$ it follows that if f is constant on M, then $v_p(f) = 0$, thus the first property holds. Furthermore if $f(p), g(p) = 0$ then we have:

$$v_p(f \cdot g) = f(p)v_p(f) + g(p)v_p(f) = 0$$

thus the second property holds. □

Recall that given a smooth map $F : \mathbb{R}^n \to R^m$, the differential of F is given by the Jacobian:

$$(D_x F)^i_j = \frac{\partial F^i}{\partial x^j}(x) = \begin{pmatrix} \frac{\partial F^1}{\partial x^1}(x) & \cdots & \frac{\partial F^1}{\partial x^n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial F^m}{\partial x^1}(x) & \cdots & \frac{\partial F^m}{\partial x^n}(x) \end{pmatrix}$$

which is an $m \times n$ matrix, and hence a linear map $\mathbb{R}^n \to \mathbb{R}^m$. The Jacobian is often interpreted as the best linear approximation to the map F at the point x, hence, we can view the Jacobian as a linear map over tangent spaces, i.e. $D_xF : T_x\mathbb{R}^n \to T_{F(x)}\mathbb{R}^m$. Since there is no reason for two manifolds M and N to be vector spaces, we opt for the latter view , that is, given a smooth map $F : M \to N$, we view the differential of F at the point p to be the linear map:

$$D_pF : T_pM \to T_{F(p)}N$$

Given a smooth map $F : M \to N$, some $v_p \in T_pM$, and some $f \in C^\infty(N)$, we define:

$$D_pF(v_p)(f) = v_p(f \circ F) \tag{1.1.2}$$

where $f \circ F$ belongs to $C^\infty(M)$. It is not difficult to see that $D_pF(v_p)$ is a derivation belonging to $T_{F(p)}$; $D_pF(v_p)$ is linear by (1.1.2), and given $f, g \in C^\infty(N)$ we have:

$$\begin{aligned} D_pf(v_p)(fg) =& v_p((f \circ F) \cdot (g \circ F)) \\ =& f \circ F(p)v_p(f \circ F) + g \circ F(p)v_p(g \circ F) \\ =& f(F(p))D_pF(v)(g) + g(F(p))D_pF(v_p)(g) \end{aligned}$$

hence $D_pF(v_p) \in T_pN$. Furthermore the map D_pF is a linear map from T_pM to $T_{F(p)}N$, as given $v_p, w_p \in T_pM$, $a, b \in R$, and $f \in C^\infty(N)$ we have that:

$$\begin{aligned} D_pF(av_p + bw_p)f =& (av_p + bw_p)(f \circ F) \\ =& (av_p)(f \circ F) + (bw_p)(f \circ F) \\ =& a \cdot v_p(f \circ F) + b \cdot w_p(f \circ F) \end{aligned}$$

We now prove three important propositions:

Proposition 1.1.1. *Given manifolds M, N, and P, and smooth maps $F : M \to N$, $G : N \to P$, the differential $D_p(G \circ F) : T_pM \to T_{F\circ G(p)}M$ at a point $p \in M$ is given by:*

$$D_{F(p)}G \circ D_pF$$

This is the chain rule for smooth maps between manifolds.

Proof. Let $f \in C^\infty(P)$, and $v_p \in T_pM$, we then have that:

$$\begin{aligned} D_p(G \circ F)(v_p)(f) =& v_p(f \circ G \circ F) \\ =& v_p((f \circ G) \circ F) \\ =& D_pF(v_p)(f \circ G) \end{aligned}$$

Let $D_pF(v_p) = w_{F(p)}$ for some $w_{F(p)} \in T_{F(p)}N$, then:

$$\begin{aligned} D_p(G \circ F)(v_p)(f) =& w_{F(p)}(f \circ G) \\ =& D_{F(P)}G(w_{F(p)})(f) \\ =& D_{F(P)}G(D_pF(v_p))(f) \\ =& D_{F(p)}G \circ D_pF(v_p)(f) \end{aligned}$$

as desired. □

For the next proposition we require this lemma:

Lemma 1.1.2. *Let M be a smooth manifold, and Id be the identity map $M \to M$, then D_pId is the identity map $T_pM \to T_pM$.*

Proof. Let $v_p \in T_pM$, and $f \in C^\infty(M)$, then:

$$\begin{aligned} D_p\mathrm{Id}(v_p)(f) =& v_p(f \circ \mathrm{Id}) \\ =& v_p(f) \end{aligned}$$

thus $D_p\mathrm{Id}$ takes any element v_p to itself, and is thus the identity map. □

Proposition 1.1.2. *Given smooth manifolds M and N, and a diffeomorphism $F : M \to N$, the differential of F at a point $p \in M$, is an isomorphism of vector spaces $D_pF : T_pM \to T_{F(p)}N$. Conversely, if F is a smooth bijection, such that D_pF is an isomorphism for all $p \in M$, then F is a diffeomorphism.*

Proof. First let $v_p \in T_pM$ and $f \in C^\infty(N)$. As F is a diffeomorphism, it has a smooth inverse, namely F^{-1}, such that $F^{-1} \circ F = \mathrm{Id}$. By the chain rule we then see that:

$$\begin{aligned} D_p\mathrm{Id}(v_p)(f) =& D_pF^{-1} \circ F(v_p)(f) \\ =& D_{F(p)}F^{-1} \circ D_pF(v_p)(f) \end{aligned} \tag{1.1.3}$$

By **Lemma (1.1.1)**, we know that (1.1.3) must be the identity, it follows that D_pF is an invertible linear map with inverse $D_{F(p)}F^{-1}$ and thus an isomorphism of vector spaces $T_pM \to T_{F(p)}N$.

Now suppose that F is a smooth bijection, and that D_pF is an isomorphism for all $p \in M$. Since F is a bijection, it follows that there exists a unique inverse map F^{-1}, we want to show that this map is smooth. Let $p \in M$ and (U, ϕ), (V, ψ) be charts around p and $F(p)$ respectively. Since D_pF is an isomorphism it follows that $\dim T_pM = \dim T_{F(p)}N$, and further that $\dim T_{\phi(p)}M = \dim T_{\psi(F(p))}$ as ϕ and ψ are local diffeomorphisms. Suppose that $\dim T_{\phi(p)}M = n$, then $\phi(U)$ and $\psi(V)$ are subsets of $\mathbb{R}^n$. It follows that:

$$\psi \circ F \circ \phi^{-1} : \phi(U) \subset \mathbb{R}^n \longrightarrow \psi(V) \subset \mathbb{R}^n$$

Since D_pF is an isomorphism, we have that $D_{\phi(p)}(\psi \circ F \circ \phi^{-1})$ is an isomorphism by the chain rule. By the inverse function theorem, we then have there exists an inverse function theorem, there exists open neighborhoods U' and V' of $\phi(p)$ and $\psi(F(p))$ respectively, such that $\psi \circ F \circ \phi^{-1}$ has a smooth inverse. By the uniqueness of inverse maps, it follows that this inverse must be given on V' by $\phi \circ F^{-1} \circ \psi^{-1}$, where the coordinate charts ϕ, and ψ are now restricted to $\phi^{-1}(U')$ and $\psi^{-1}(V')$. This shows that in an open neighborhood of $F(p)$, F^{-1} is smooth. Since for all $p \in M$ D_pF is an isomorphism, and since F is a bijection, we have that for all $F(p) \in N$, there exist smooth charts $(\psi^{-1}(V'), \psi)$ and $(\phi^{-1}(U'), \phi)$ around $F(p)$ and p such that $\phi \circ F^{-1} \circ \psi^{-1}$ is smooth, so F^{-1} is a smooth map, implying the claim. □

Corollary 1.1.1. *Let M be a smooth manifold of dimension n, then for all $p \in M$, the tangent space T_pM is isomorphic to $\mathbb{R}^n$.*

Clearly, from **Corollary 1.1.1** , we see that $\dim M = \dim T_pM$ for all $p \in M$, implying that diffeomorphisms preserve the dimension of smooth manifolds. Furthermore, a basis for T_pM can be constructed in the following way; let (U, ϕ) be a chart for M, and let:

$$\phi = (x^1, \dots, x^n)$$

be the coordinate functions on M. In these coordinates we have that a basis for $T_{\phi(p)}\phi(U)$ is given by the set:

$$\left\{ \frac{\partial}{\partial x^1}\Big|_{\phi(p)}, \dots, \frac{\partial}{\partial x^n}\Big|_{\phi(p)} \right\}$$

As $D_p\phi$ is an isomorphism, with inverse given by $D_{\phi(p)}\phi^{-1}$, we see that a basis for T_pM, is given by the set:

$$\left\{ D_{\phi(p)}\phi^{-1}\left(\frac{\partial}{\partial x^1}\Big|_{\phi(p)} \right), \dots, D_{\phi(p)}\phi^{-1}\left(\frac{\partial}{\partial x^n}\Big|_{\phi(p)} \right) \right\}$$

For brevity we write:

$$\frac{\partial}{\partial x^i}\Big|_p = D_{\phi(p)}\phi^{-1}\left(\frac{\partial}{\partial x^i}\Big|_{\phi(p)} \right)$$

We turn to an example:

Example 1.1.7. Let $M = \mathbb{S}^2$, it stands to reason that from the definition of $\mathbb{S}^2$ that the tangent space at any point p would be the set of vectors orthogonal to $p \in \mathbb{S}^2$, let us see this in a chart. We use the following parameterization of the sphere:

$$\psi^{-1}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

where $\psi(U) = (\theta, \phi) \in (0, \pi) \times (0, 2\pi) \subset \mathbb{R}^2$. Here our coordinates for $\mathbb{S}^2$ are θ, ϕ, hence our basis vectors in $T_{(\theta,\phi)}\psi(U)$ are represented by:

$$\left\{ \frac{\partial}{\partial\theta}\Big|_{(\theta,\phi)}, \frac{\partial}{\partial\phi}\Big|_{(\theta,\phi)} \right\}$$

We now calculate $D_{(\theta,\phi)}\psi^{-1}$ as a matrix of partial derivatives:

$$D_{(\theta,\phi)}\psi^{-1} = \begin{pmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{pmatrix}$$

thus we obtain:

$$\begin{aligned} D_{(\theta,\phi)}\psi^{-1}(\partial_\theta) &= \begin{pmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix} \end{aligned}$$

and similarly that:

$$D_{(\theta,\phi)}\psi^{-1}(\partial_\phi) = \begin{pmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{pmatrix}$$

Via a brief computation one sees that:

$$\langle \psi^{-1}(\theta,\phi), D_{\theta,\phi}\psi^{-1}(\partial_\theta) \rangle = \langle \psi^{-1}(\theta,\phi), D_{\theta,\phi}\psi^{-1}(\partial_\phi) \rangle = 0$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product on $\mathbb{R}^3$. Since $D_{(\theta,\phi)}\psi^{-1}(\partial_\theta)$ and $D_{(\theta,\phi)}\psi^{-1}(\partial_\phi)$ form a basis for $T_{(\theta,\phi)}\mathbb{S}^2$, we see that $T_{(\theta,\phi)}\mathbb{S}^2$ is the set of vectors in $\mathbb{R}^3$ orthogonal to $\psi^{-1}(\theta, \phi)$.

Now let us work more generally. Let F be a smooth map from $M \to N$, and (U, ϕ), (V, ψ) be coordinate charts for M and N, where $\phi = (x^1, \ldots, x^n)$ and $\psi = (y^1, \ldots, y^m)$. The coordinate representation for F, denoted F^c, is given by:

$$F^c = \psi \circ F \circ \phi^{-1}$$

We now see that:

$$\begin{aligned} D_pF\left(\frac{\partial}{\partial x^i}\Big|_p\right) =& D_pF\left(D_\phi(p)\phi^{-1}\frac{\partial}{\partial x^i}\Big|_{\phi(p)}\right) \\ =& D_{\psi(F(p))}\psi^{-1} \circ D_{F(p)}\psi \circ D_pF \circ D_{\phi(p)}\phi^{-1}\left(\frac{\partial}{\partial x^i}\Big|_{\phi(p)}\right) \\ =& D_{F^c(\phi(p))}\psi^{-1}(p) \circ D_{\phi(p)}F^c\left(\frac{\partial}{\partial x^i}\Big|_{\phi(p)}\right) \end{aligned} \tag{1.1.4}$$

Since F^c is a map from $\mathbb{R}^n \to \mathbb{R}^m$, we can compute its differential as a Jacobian matrix with entries:

$$\frac{\partial (F^c)^j}{\partial x^i}(\phi(p))$$

thus via matrix multiplication we see that:

$$D_{\phi(p)}F^c\left(\frac{\partial}{\partial x^i}\Big|_{\phi(p)}\right) = \frac{\partial (F^c)^j}{\partial x^i}(\phi(p))\frac{\partial}{\partial y^j}\Big|_{\psi(p)}$$

hence (1.1.4) becomes:

$$\begin{aligned} D_pF\left(\frac{\partial}{\partial x^i}\Big|_p\right) =& \frac{\partial (F^c)^j}{\partial x^i}(\phi(p)) \cdot D_{\psi(F(p))}\psi^{-1}\left(\frac{\partial}{\partial y^j}\Big|_{\psi(F(p))}\right) \\ =& \frac{\partial (F^c)^j}{\partial x^i}(\phi(p))\frac{\partial}{\partial y^j}\Big|_{F(p)} \end{aligned} \tag{1.1.5}$$

At times we refer to D_pF as the pushforward of F, and denote it by F_* to avoid notational clutter.

Let us look at how the basis vectors transform under a change coordinates. Let M be a manifold, and let $\phi = (x^1, \ldots, x^n)$, $\psi = (y^1, \ldots, y^n)$, be the coordinate functions for two charts (U, ϕ), (V, ψ) which overlap. We see that for $y \in \psi(U \cap V)$:

$$\phi \circ \psi^{-1}(y) = (x^1(y), \ldots, x^n(y))$$

is a map from $\mathbb{R}^n \to \mathbb{R}^n$, hence has Jacobian:

$$D_{\psi(p)}(\phi \circ \psi^{-1}) = \frac{\partial x^j}{\partial y^i}(\psi(p)) \tag{1.1.6}$$

Let $v \in T_pV$, then:

$$v = v^i \frac{\partial}{\partial y^i}\Big|_p = v^i D_{\psi(p)}\psi^{-1}\left(\frac{\partial}{\partial y^i}\Big|_{\psi(p)}\right)$$

introducing the identity map $D_{\psi(p)}(\phi^{-1} \circ \phi)$, we by chain rule that:

$$\begin{aligned} v =& v^i D_{\phi(p)}\phi^{-1} \circ D_{\psi(p)}(\phi \circ \psi^{-1})\left(\frac{\partial}{\partial y^i}\Big|_{\psi(p)}\right) \\ =& v^i D_{\phi(p)}\phi^{-1}\left(\frac{\partial x^j}{\partial y^i}(\psi(p))\frac{\partial}{\partial x^j}\Big|_{\phi(p)}\right) \\ =& v^i \frac{\partial x^j}{\partial y^i}(\psi(p))\frac{\partial}{\partial x^j}\Big|_p \end{aligned}$$

Thus if v is a vector in T_pV, and w is the corresponding vector in T_pU, we see that their components are related by:

$$w^j = v^i \frac{\partial x^j}{\partial y^i}(\psi(p)) \tag{1.1.7}$$

We now define three classes of smooth maps based on properties of their differentials, and then state two theorems without proof.

Definition 1.1.7. Let M and N be smooth manifolds, and $F : M \to N$ a smooth map, then:

- If D_pF is injective at all points $p \in M$ we call F a smooth **immersion**.
- If D_pF is surjective at all points $p \in M$ we call F a smooth **submersion**.
- If D_pF is injective at all points $p \in M$, and F is a homeomorphism on to its image in the subspace topology, we call F an **embedding**. In particular we call the image of F an **embedded submanifold** of N.
- A point $p \in M$ is a **regular point** of F if D_pF is a surjection onto $T_{F(p)}N$.
- A point $q \in N$ is a **regular value** of F if each point p in the preimage of $F^{-1}(q) \subset M$ is a regular point.

Theorem 1.1.1 (Regular Value Theorem)**.** *Let M and N be smooth manifolds, and let $q \in N$ be a regular value of the smooth map $F : M \to N$. Then, $L = F^{-1}(q) \subset M$ is an embedded submanifold of M of dimension:*

$$\dim L = \dim M - \dim N$$

Theorem 1.1.2 (Regular Point Theorem)**.** *Let p be regular point of the map F. Then there exist smooth charts (U, ϕ) of M around p, and (V, ψ) around $f(p)$ satisfying:*

$$\phi(p) = 0 \qquad \text{and} \qquad \psi(f(p)) = 0 \qquad \text{and} \qquad f(U) \subset V$$

Furthermore, the map $\psi \circ f \circ \phi^{-1}$ satisfies:

$$\psi \circ f \circ \phi^{-1}(x_1, \dots, x_{n+k}) = (x_1, \dots x_n)$$

where $\dim M = n + k$, *and* $\dim N = n$. *In other words, f 'locally looks like' a projection.*

Example 1.1.8. The n-sphere is an embedded submanifold of $\mathbb{R}^{n+1}$ by the Regular Value Theorem. Indeed, let $F : \mathbb{R}^{n+1} \to \mathbb{R}$ be the map:

$$F(x) = \|x\|^2 = (x^1)^2 + \cdots + (x^n)^2$$

Letting $p = (x^1, \dots, x^n)$, we see that:

$$D_pF = \begin{pmatrix} 2x^1 \\ \vdots \\ 2x^n \end{pmatrix}$$

which is surjective everywhere but $p = (0, \dots, 0)$, hence we have that 1 is a regular value of F, therefore $F^{-1}(1)$ is an embedded submanifold of $\mathbb{R}^{n+1}$ of dimension $n + 1 - 1 = n$. Note that:

$$F^{-1}(1) = \{x \in \mathbb{R}^{n+1} : |x| = 1\} = \mathbb{S}^n$$

thus, $\mathbb{S}^n$ is an embedded submanifold of dimension $\mathbb{R}^n$.

Now that we have sufficiently built up tangent vectors and tangent spaces, we wish to turn our attention to smoothly assigning a $v \in T_pM$ for each $p \in M$; i.e. smooth vector fields. To do this though, we must first examine the space of all tangent vectors to M.

Definition 1.1.8. The **tangent bundle**, denoted TM is the set constructed by the disjoint union of T_pM for each $p \in M$, that is:

$$TM = \coprod_{p \in M} T_pM$$

This set comes equipped with a natural projection map $\pi : TM \to M$. For some $p \in M$ and $v \in T_PM$, we refer to elements of TM as an ordered pair (p, v), where $\pi(p, v) = p$

For the special case $M = \mathbb{R}^n$ we see that via our early identification of $T_x\mathbb{R}^n$ with $\{x\} \times \mathbb{R}^n$:

$$T\mathbb{R}^n = \coprod_{x \in \mathbb{R}^n} T_x\mathbb{R}^n = \coprod_{x \in \mathbb{R}^n} \{x\} \times \mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$$

however this is, in general, not true for any M. For a smooth manifold M, if it has tangent bundle $TM \cong M \times \mathbb{R}^n$, we call TM *trivial.* In the following proposition, we see that the tangent bundle can be thought of as a smooth manifold in a natural way:

Proposition 1.1.3. *For any n-manifold M, the tangent bundle TM has a natural topology and smooth structure that make it into a smooth manifold of dimension $2n$. The projection map π with respect to this smooth structure is smooth.*

Proof. First note that for any smooth chart (U, ϕ), we have that $\pi^{-1}(U) \subset TM$ is the set of all tangent vectors to M at all points $p \in U$, i.e.,

$$\pi^{-1}(U) = \coprod_{p \in U} T_pM$$

As M is second countable, there exists a countable basis for it's topology $\{U_i\}_{i \in \mathbb{N}}$. Likewise we define a basis for TM by noting that:

$$\pi^{-1}(U_i) = U_i \times \mathbb{R}^n$$

since we can choose a basis for M such that every U_i is diffeomorphic to $\mathbb{R}^n$, hence the tangent bundle restricted to U_i is trivial. We then define a basis for TM by taking the countable basis of open balls for $\mathbb{R}^n$ and crossing it with each U_i. This is then a countable basis for TM, hence TM is second countable.

Furthermore, since T_pM is a real vector space isomorphic to $\mathbb{R}^n$, it follows that for points in the same fibre of TM, (p,v) and (p,w), there exist disjoint open sets $V_v, V_w \subset T_pM$, such that $(p,v) \in V_v$ and $(p,w) \in V_w$. Finally, note that for two points $p,q \in M$, there exist disjoint open sets U_p, U_q such that $p \in U_p$, and $q \in U_q$, thus for points which lie in different fibres TM, (p,v) and (q,w), the disjoint open sets $\pi^{-1}(U_p)$ and $\pi^{-1}(U_q)$ contain (p,v) and (q,w) respectively, hence TM is Hausdorff.

Finally, we define coordinates charts on TM in order to a) show that TM is locally Euclidean, and b) determine a smooth structure on TM. Let (U,ϕ) be a chart for M, with coordinate functions $\phi = (x^1, \dots, x^n)$, we define a chart $\phi_{TM} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ by:

$$\phi_{TM}\left(v^i \frac{\partial}{\partial x^i}\bigg|_p \right) = (x^1(p), \dots, x^n(p), v^1, \dots, v^n)$$

with inverse given by:

$$\phi_{TM}^{-1}(x^1, \dots, x^n, v^1, \dots, v^n) = v^i D_x\phi^{-1}\left(\frac{\partial}{\partial x^i}\bigg|_x \right) = v^i \frac{\partial}{\partial x^i}\bigg|_{\phi^{-1}(x)}$$

thus ϕ_{TM} is a smooth bijection onto its image, $\phi(U) \times \mathbb{R}^n$. Given two charts (U,ϕ) and (V,ψ), we have corresponding charts $(\pi^{-1}(U), \phi_{TM})$ and $(\pi^{-1}(V), \psi_{TM})$, which map open set $\pi^{-1}(U) \cap \pi^{-1}(V)$ to $\mathbb{R}^{2n}$ in the following way:

$$\phi_{TM}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \phi(U \cap V) \times \mathbb{R}^n$$
$$\psi_{TM}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \phi(U \cap V) \times \mathbb{R}^n$$

which are both open in $\mathbb{R}^{2n}$. Let $\phi = (x^1, \dots, x^n)$ and $\psi = (y^1, \dots, y^n)$, then the transition map $\phi_{TM} \circ \psi_{TM}^{-1}$ can be explicitly computed by making use of (1.1.4):

$$\phi_{TM} \circ \psi_{TM}^{-1}(y^1, \dots, y^n, v^1, \dots, v^n) = (x^1, \dots, x^n, \frac{\partial x^1}{\partial y^j} v^j, \dots, \frac{\partial x^n}{\partial y^j} v^j)$$

which is smooth. Therefore, TM is locally Euclidean, and thus a topological manifold of dimension $2n$; it has a smooth structure determined by the smooth atlas $\{\pi^{-1}(U_i), \phi_{TM_i}\}_{i \in \mathbb{N}}$, where $\{(U_i, \phi_i)\}_{i \in \mathbb{N}}$ covers M, and is thus a smooth manifold of dimension $2n$, as desired. Furthermore, note that with the charts (U,ϕ) for M, and $(\pi^{-1}(U), \phi_{TM})$ for TM, the coordinate representation of π is given by:

$$\pi(x^1, \dots, x^n, v^1, \dots, v^n) = (x^1, \dots, x^n)$$

hence π is smooth. □

Now that we have shown that the tangent bundle is a smooth manifold, we introduce the following definition:

Definition 1.1.9. Let $F : M \to N$ be a smooth map, the **global differential** of F is then the map $DF : TM \to TN$ who's restriction to each tangent space is $D_pF : T_PM \to T_pN$.

As the next proposition shows this map is smooth.

Proposition 1.1.4. *Let $F : M \to N$ be smooth, then the global differential of F is smooth.*

Proof. In some local chart, let M have coordinates $(x^1, \dots, x^n)$, and let $\dim(N) = m$. Then, by (1.1.5) we see that:

$$DF^c(x^1, \dots, x^n, v^1, \dots, v^n) = \left((F^c)^1(x), \dots, (F^c)^m(x), \frac{\partial (F^c)^1}{\partial x^i} v^i, \dots, \frac{\partial (F^c)^m}{\partial x^i} v^i \right)$$

which is smooth since F is. □

Furthermore, from **Lemma 1.1.2**, **Proposition 1.1.1** and **Proposition 1.1.2** we obtain the following corollary:

Corollary 1.1.2. *Let $F : M \to N$, and $G : N \to P$ be smooth maps. Then the following statements hold:*

a) $D(G \circ F) = DG \circ DF$

b) $DId_M = Id_{TM}$

c) If F is a diffeomorphism then $DF : TM \to TN$ is also a diffeomorphism with inverse $(DF)^{-1} = D(F^{-1})$.

We can now properly define vector fields:

Definition 1.1.10. A **vector field** on a smooth manifold M is a **smooth section** of the map $\pi : TM \to M$, or rather a smooth map:

$$\begin{aligned} X :& M \to TM \\ & p \longmapsto X_p \end{aligned}$$

such that:

$$\pi \circ X = \mathrm{Id}_M$$

We denote the set of all vector fields over M by $\mathfrak{X}(M)$

Choosing coordinates, we can locally write that:

$$X_p = X^i(p) \frac{\partial}{\partial x^i}\bigg|_p$$

where each $X^i \in C^\infty(M)$. At times we don't specify the point p and simply write:

$$X = X^i \frac{\partial}{\partial x^i}$$

where it is understood that the X^i are still smooth functions on M, and that the $\partial/\partial x^i$'s are *coordinate vector fields* on M. Furthermore, we see that $\mathfrak{X}(M)$ is a module over the ring $C^\infty(M)$, as given two vector fields $X, Y \in \mathfrak{X}(M)$, and two smooth functions $f, g \in C^\infty(M)$ we have:

$$(fX + gY)_p = f(p)X_p + g(p)Y_p \in \mathfrak{X}(M)$$

Since a zero section of TM always exists as there is as distinguished 0 element in each fibre[2] we also have that $0 \in \mathfrak{X}(M)$. Allowing f, g to be constant functions shows that $\mathfrak{X}(M)$ is also a $\mathbb{R}$-linear vector space under pointwise addition and scalar multiplication.

Example 1.1.9. In the vector calculus formulation of electromagnetism, it is often common to employ spherical coordinates (r, θ, ϕ) on $\mathbb{R}^3 \setminus \{0\}$ given by:

$$x = r\sin\theta\cos\phi, y = r\sin\theta\sin\phi, z = r\cos\theta$$

The coordinate vector fields are then the set:

$$\left\{\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \phi}\right\}$$

The electric field due to a static point charge q is then given by:

$$\mathbf{E}(r, \theta, \phi) = \frac{q}{r^2}\frac{\partial}{\partial r}\bigg|_{(r,\theta,\phi)}$$

Definition 1.1.11. Let M be a smooth manifold of dimension n, a **global frame** is a set of vector fields $\{X_p^1, \ldots, X_p^n\}$ which span the tangent space T_pM at each point $p \in M$.

[2] One can easily glue together local zero sections with a partition of unity and the transition functions defined in **Proposition 1.1.3**, to obtain a global zero section which is well defined and smooth.

Definition 1.1.12. Let M be a smooth manifold of dimension n, and $U \subset M$, a **local frame** is a set of vector fields $\{X_p^1, \ldots, X_p^n\}$ which span the tangent space T_pM at each point $p \in U$.

It is in general easy to find local frames, indeed take any coordinate chart (U, ϕ) for M, as $D_p\phi^{-1}$ is an isomorphism at each point $p \in M$, the coordinate vector fields form a local frame on U. However, the existence of a global frame is not guaranteed, as given a global frame $\{X_p^1, \ldots, X_p^n\}$, this would allow us to construct a diffeomorphism $M \times \mathbb{R}^n \to TM$, by:

$$(p, x^1, \ldots, x^n) \longmapsto (p, x^1X_p^1, \ldots, x^nX_p^n)$$

which, as we mentioned earlier, is not guaranteed.

Vector fields are also derivations at every point $p \in M$ in the following way:

$$(Xf)(p) = X_p(f) \tag{1.1.8}$$

Since for all $p \in M$, X_p is a derivation at the point p, we see that a vector field X can be viewed as a map $C^\infty(M) \to C^\infty(M)$, that is $\mathbb{R}$-linear, and satisfies the Leibniz rule, we call such a map a *derivation*, which brings us to the following proposition:

Proposition 1.1.5. *Let M be a smooth manifold, a map $D : C^\infty(M) \to C^\infty(M)$ is a derivation if and only if it is of the form $Df = Xf$ for some smooth vector field $X \in \mathfrak{X}(M)$.*

Proof. From (1.1.7) it is clear that any vector field X can be thought of as a derivation, so all that is left is to show the converse. Suppose $D : C^\infty(M) \to C^\infty(M)$ is a derivation, then at a point $p \in M$ and some $f \in C^\infty(M)$ we have that:

$$(Df)(p) = v_p(f) \tag{1.1.9}$$

for some $v_p \in T_pM$. Since $(Df) \in C^\infty(M)$ we see that there exists a v_p in every T_pM such that the (1.1.8) is true. Now let X be a map from $M \to TM$ such that at each point p, X_p is the $v_p \in T_pM$ such that (1.1.8) holds. It then follows that:

$$Xf = Df$$

Finally, since Df is smooth, it follows that v_p must vary smoothly, and hence $X \in \mathfrak{X}(M)$ as desired. □

Definition 1.1.13. Suppose $F : M \to N$ is a smooth map, and suppose there exists a vector field $Y \in \mathfrak{X}(N)$ such that $\forall p \in M$, $D_pF(X_p) = Y_{F(p)}$, then we say X and Y are **F related**.

Proposition 1.1.6. *Suppose M and N is a smooth manifold, and F a diffeomorphism between them. Then for every $X \in \mathfrak{X}(M)$ there is a unique $Y \in \mathfrak{X}(N)$ such that X and Y are F related..*

Proof. Note that if X and Y are F related we necessitate for all $p \in M$:

$$D_pF(X_p) = Y_{F(p)}$$

Since D_pF is an isomorphism of vector spaces at each point p, we simply define the vector field Y at each point $q \in N$ as:

$$Y_q = D_{F^{-1}(q)}F(X_{F^{-1}(q)})$$

Which is indeed smooth as Y is the composition of the smooth maps:

$$Y : N \xrightarrow{F^{-1}} M \xrightarrow{X} TM \xrightarrow{DF} TN$$

□

We often denote Y by F_*X, which can be explicitly computed point-wise via:

$$(F_*X)_q = D_{F^{-1}(q)}F(X_{F^{-1}(q)})$$

It is important to note that F_*X is only well defined when F is a diffeomorphism, as **Proposition 1.1.6** explicitly depended on the existence of F^{-1}. We would also like to see how vector fields which are F related act on smooth functions, motivating the next proposition.

Proposition 1.1.7. *Suppose $F : M \to N$ is a smooth map between manifolds, and $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$. Then X and Y are F related if and only if for every $f \in C^\infty(N)$:*

$$X(f \circ F) = Yf \circ F$$

Proof. For any $p \in M$, and any $f \in C^\infty(N)$:

$$X_p(f \circ F) = D_pF(X_p)(f)$$

while:

$$(Yf) \circ F(p) = Y_{F(p)}(f)$$

Since both hold for all $p \in M$, we have that:

$$X(f \circ F) = (Yf) \circ F$$

holds for all $f \in C^\infty(N)$ if and only if X and Y are F related. □

Combining the two proceeding propositions gives us the following the statement:

$$((F_*X)f) \circ F = X(f \circ F)$$

where $X \in \mathfrak{X}(M)$, $f \in C^\infty(N)$, and $F : M \to N$ is a diffeomorphism. Furthermore, **Proposition 1.1.5** allows us to define the following bracket operation on $\mathfrak{X}(M)$:

$$[X, Y](f) = X \circ Y(f) - Y \circ X(f) \tag{1.1.10}$$

As $Y(f)$ and $X(f)$ are both in $\mathbb{C}^\infty(M)$, $[X, Y]$ is a map $C^\infty(M) \to C^\infty(M)$. We would like this map to also be a derivation, and hence a vector field, motivating our following definition:

Definition 1.1.14. A **Lie algebra** is any vector space V with a bracket operation $[\cdot, \cdot] : V \times V \to V$, called the **Lie bracket**, that satisfies the following properties:

a) $[\cdot, \cdot]$ is bilinear:

$$\begin{aligned}[aX + bY, Z] =& a[X, Z] + b[Y, Z]\\ [X, aY + bZ] =& a[X, Y] + b[X, Z]\end{aligned}$$

b) $[\cdot, \cdot]$ is anticommutative:

$$[X, Y] = -[Y, X]$$

c) $[\cdot, \cdot]$ satisfies the Jacobi Identity:

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

Proposition 1.1.8. *Equipped with the bracket operation defined in* (1.1.10)*, $\mathfrak{X}(M)$ is a Lie algebra over $\mathbb{R}$.*

Proof. We first show that for $X, Y \in \mathfrak{X}(M)$, $[X, Y]$ is also in $\mathfrak{X}(M)$, by showing that is also a derivation. Linearity comes from the fact that for $a, b \in \mathbb{R}$, and $f, g \in C^\infty(M)$:

$$\begin{aligned}[X, Y](af + bg) =& X \circ Y(af + bg) - Y \circ X(af + bg)\\ =& X(aY_p(f) + bY(g)) - Y(X(f) + bX(g))\\ =& aX \circ Y(f) + bX \circ Y(g) - aY \circ X(f) - bY \circ X(g)\\ =& a[X, Y](f) + b[X, Y](g)\end{aligned}$$

Furthermore, $[X, Y]$ satisfies the Leibniz rule by:

$$\begin{aligned}[X, Y](fg) =& X \circ Y(fg) - Y \circ X(fg)\\ =& X_p(fY(g) + gY(f)) - Y(fX(g) + gX(f))\\ =& f(X \circ Y(g)) + g(X \circ Y(f)) + X(f) \cdot Y(g) + X(g) \cdot Y(g)\\ & - f(X \circ Y(g)) - g(Y \circ X(f)) - Y(f) \cdot X(g) - Y(g) \cdot X(f)\\ =& f[X, Y](g) + g[X, Y](f)\end{aligned}$$

thus $[X,Y] \in \mathfrak{X}(M)$. Condition a) follows from the fact that for $f \in C^\infty(M)$, $a, b \in \mathbb{R}$, and $X, Y, Z \in \mathfrak{X}(M)$:

$$\begin{aligned}
[aX + bY, Z](f) =&(aX + bY) \circ Z(f) - Z \circ (aX + bY)(f)\\
=&aX \circ Z(f) + bY \circ Z(f) - Z \circ (aX(f) + bY(f))\\
=&aX \circ Z(f) + bY \circ Z(f) - aZ \circ X(f) - bZ \circ Y(f)\\
=&a[X, Z] + b[Y, Z]\\
[X, aY + bZ](f) =&X \circ (aY + bZ)(f) - (aY + bZ) \circ X(f)\\
=&X \circ (aY(f) + bZ(f)) - aY \circ X(f) - bZ \circ X(f)\\
=&aX \circ Y(f) + bX \circ Z(f) - aY \circ X(f) - bZ \circ X(f)\\
=&a[X, Y] + b[X, Z]
\end{aligned}$$

Condition b) follows from:

$$\begin{aligned}
[X, Y](f) =&X \circ Y(f) - Y \circ X(f)\\
=& - (Y \circ X(f) - X \circ Y(f))\\
=& - [Y, X](f)
\end{aligned}$$

Finally condition c) follows from an involved calculation. We begin by taking:

$$\begin{aligned}
[Y, [Z, X]](f) =&[Y, Z \circ X - X \circ Z]\\
=&[Y, Z \circ X](f) - [Y, X \circ Z](f) \qquad (1.1.11)\\
[Z, [X, Y]](f) =&[Z, X \circ Y](f) - [Z, Y \circ X](f) \qquad (1.1.12)
\end{aligned}$$

where we are employing a mild abuse of notation, as the composition $X \circ Y$ is not a vector field. Adding (1.1.11) and (1.1.11) gives:

$$\begin{aligned}
[Y, [Z, X]](f) + [Z, [X, Y]](f) =&Y \circ Z \circ X(f) - Z \circ X \circ Y(f) - Y \circ X \circ Z(f) + X \circ Z \circ Y(f)\\
&+Z \circ X \circ Y(f) - X \circ Y \circ Z(f) - Z \circ Y \circ X(f) + Y \circ X \circ Z(f)\\
=&Y \circ Z \circ X(f) + X \circ Z \circ Y(f) - X \circ Y \circ Z(f) - Z \circ Y \circ X(f)\\
=&[Y \circ Z, X](f) - [Z \circ Y, X](f)\\
=&[[Y, Z], X]\\
=& - [X, [Y, Z]]
\end{aligned}$$

hence condition (c) holds. □

The following example details how the Lie bracket works in local coordinates:

Example 1.1.10. Let M be a smooth manifold of dimension n, and (U, ϕ) be a chart for M where $\phi = (x^1, \dots, x^n)$, the coordinate vector fields on U then form a local frame for U. Let X and Y be smooth vector fields on U such that:

$$\begin{aligned}
X =&X^i \frac{\partial}{\partial x^i} = X^i \partial_i\\
Y =&Y^j \frac{\partial}{\partial x^j} = Y^j \partial_j
\end{aligned}$$

where $X^i, Y^i \in C^\infty(M)$. We calculate the Lie bracket as follows, let $f \in C^\infty(M)$:

$$\begin{aligned}
[X, Y](f) =&X \circ Y(f) - Y \circ X(f)\\
=&X^i \partial_i (Y^j \partial_j(f)) - Y^j \partial_j (X^i \partial_i f)\\
=&X^i \partial_i (Y^j) \partial_j(f) + X^i Y^j \partial_i \partial_j(f) - Y^j \partial_j X^i \partial_i(f) - Y^j X^i \partial_j \partial_i(f)\\
=&X^i \partial_i (Y^j) \partial_j(f) - Y^j \partial_j (X^i) \partial_i(f) \qquad (1.1.13)
\end{aligned}$$

Equation (1.1.12) is then the formula for the Lie bracket of two vector fields in coordinates.

Before we move on, we prove the following proposition, which will be vital when we discuss the Lie algebra of a Lie group.

Proposition 1.1.9. *Suppose $F : M \to N$ is a smooth map between manifolds, and let $X_1, X_2 \in \mathfrak{X}(M)$ and $Y_1, Y_2 \in \mathfrak{X}(N)$ be vector fields such that X_i is F related to Y_i. Then $[X_1, X_2]$ is F related to $[Y_1, Y_2]$. In particular, if F is a diffeomorphism we have that:*

$$F_*[X_1, X_2] = [F_* X_1, F_* X_2]$$

Proof. Since X_i is F related to Y_i, by **Proposition 1.1.7** we have that for $f \in C^\infty(N)$:

$$X_i X_j (f \circ F) = X_i((Y_j f) \circ F) = (Y_i Y_j f) \circ F$$

Therefore:

$$\begin{aligned}[X_1, X_2](f \circ F) =& X_1 X_2 (f \circ F) - X_2 X_1 (f \circ F) \\ =& (Y_1 Y_2 f) \circ F - (Y_2 Y_1 f) \circ F \\ =& ([Y_1, Y_2] f) \circ F\end{aligned}$$

So $[X_1, X_2]$ is F related to $[Y_1, Y_2]$. Suppose now that F is a diffeomorphism, then, by **Proposition 1.1.6**, $Y_1 = F_* X_1$, $Y_2 = F_* X_2$, and:

$$F_*[X_1, X_2] = [Y_1, Y_2] = [F_* X_1, F_* X_2]$$

□

1.1.3 Differential Forms and Integration

Before we discuss differential forms, we must first define covectors and the exterior algebra of a finite dimensional vector space V.

Definition 1.1.15. Given a finite dimensional real vector space V, a **covector** on V is a real-valued linear function on V, i.e., a linear map $\lambda : V \to \mathbb{R}$. The space of all covectors on V is itself a real vector space under pointwise addition and scalar multiplication. We denote V^*, and and call it the dual space of V.

Proposition 1.1.10. *For any basis for V, $\{e_i, \dots, e_n\}$, we can a obtain dual basis for V^* $\{\lambda^1, \dots, \lambda^n\}$ defined via the relation:*

$$\lambda^i(e_j) = \delta^i_j$$

Proof. Let $\omega \in V^*$, and $v \in V$ then as ω is linear we have:

$$\omega(v) = v^i \omega(e_i)$$

Let $\omega(e_i) = b_i$ for some $b_i \in \mathbb{R}$, then we have that:

$$\omega(v) = v^i b_i$$

Which clearly implies that ω can be written as the sum $b_j \lambda^j$ as:

$$\omega(v) = v^i b_j \lambda^j(e_i) = v^i b_j \delta^j_i = v^i b_i$$

thus the set $\{\lambda^1, \dots, \lambda^n\}$ spans V^*. To see that these are linearly independent, let:

$$\omega = b_i \lambda^i = 0 \in V^*$$

then we have that for all $j \in \{1, \dots, n\}$:

$$\omega(e_j) = b_j = 0$$

thus each b_j must be zero, therefore $\{\lambda^1, \dots, \lambda^n\}$ form a basis for V^*. □

From the preceding proposition we obtain the following corollary:

Corollary 1.1.3. *Let V be a finite dimensional real vector space, and V^* it's dual space. Then, the double dual space $(V^*)^* \cong V$.*

Proof. For all $v \in V$ we define the linear map $\tau_v : V^* \to \mathbb{R}$ by:

$$\tau_v(\omega) = \omega(v)$$

Now consider the map:

$$F : V \longrightarrow (V^*)^*$$
$$v \longmapsto \tau_v$$

We see that this map is linear as for all $a, b \in \mathbb{R}$, $v, w \in V$, and $\omega \in V^*$:

$$\tau_{av+bw}(\omega) = \omega(av + bw) = a\omega(v) + b\omega(w) = a\tau_v(\omega) + b\tau_w(\omega)$$

Furthermore, if for some $v \in V$:

$$\tau_v(\omega) = \omega(v) = 0$$

for all $\omega \in V^*$, then $v = 0$ as the only element $v \in V$ which maps to zero under every linear function on V is 0, implying that F is injective. Since **Proposition 1.1.10** implies that the dual of any finite dimensional vector space V has the same dimension as V, we see that by rank nullity F is an isomorphism, as desired. $\square$

We refer to covectors as *one forms* on V. Using one forms, and vectors we can build a wide variety of multilinear maps called tensors.

Definition 1.1.16. A **tensor** of type (l, k) is a multilinear map:

$$V_1^* \times \ldots \times V_l^* \times V_1 \times \ldots \times V_k \to \mathbb{R}$$

We can construct tensors using the **tensor product**, denoted $\otimes$. For $v_1, \ldots, v_l, w_1, \ldots, w_k \in V$, and $\omega_1, \ldots, \omega_k, \lambda_1, \ldots, \lambda_l \in V^*$, we construct a tensor of type (l, k) via the defining relation:

$$v_1 \otimes \cdots \otimes v_l \otimes \omega_1 \otimes \cdots \otimes \omega_k(\lambda_1, \ldots, \lambda_l, w_1, \ldots, w_k) = \lambda_1(v_1) \cdots \lambda_l(v_l) \cdot \omega_1(w_1) \cdots \omega_k(w_k)$$

The space of all tensors of the type (l, k), for a vector space V, which we denote by $T^{l,k}(V)$, is itself a vector space of dimension n^{l+k}. For reasons that will become apparent later, the types of multilinear maps we are interested in are ones that are antisymmetric on V, bringing us to our next definition.

Definition 1.1.17. A $(0, k)$ tensor, i.e. a linear map:

$$V_1 \times \ldots \times V_k \to \mathbb{R}$$

is called a **k-form** on V, if it is also totally antisymmetric. That is for $\lambda \in T^{0,k}(V)$, we have:

$$\lambda(\ldots, v_i, \ldots, v_j, \ldots) = -\lambda(\ldots, v_j, \ldots, v_i, \ldots)$$

We denote the space of all such tensors as $\Lambda^k(V^*)$.

It is clear from the preceding definition that for any vector space V of dimension n, the highest order k form that can exist on V is one with $k = n$, as if $k = n + 1$ then at least two of the vectors must linearly depend on one another; we call these types of forms *top forms*, and the set of top forms on V is a one dimensional vector space. The following definition defines an associative multiplicative structure on the set of all forms:

Definition 1.1.18. Let λ be a k-form and ω be an l-form, then the **wedge product** of the two, denoted $\lambda \wedge \omega$, is a $k + l$ form defined by:

$$(\lambda \wedge \omega)(v_1, \ldots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \mathrm{sgn}(\sigma)\lambda(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot \omega(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)})$$

where S_{k+l} denotes the set of permutations of $\{1, 2, \ldots, k + l\}$.

In particular, one can use the definition above to find that:

$$\lambda \wedge \omega = (-1)^{kl} \omega \wedge \lambda$$

and that if l is odd:

$$\omega \wedge \omega = 0$$

Proposition 1.1.11. *Let $\{\omega^i\}$ be a set of k covector's on V. Then:*

$$\omega^1 \wedge \cdots \wedge \omega^k = \sum_{\sigma \in S_k} sgn(\sigma) \omega^{\sigma(1)} \otimes \cdots \otimes \omega^{\sigma(k)}$$

Proof. We proceed by induction, the $k = 1$ case is trivial, so assuming the $k-1$th case, we wish to prove the kth case. Let $v_1, \cdots, v_k \in V$, then, by **Definition 1.1.18** we have that:

$$\omega^1 \wedge \cdots \wedge \omega^{k-1} \wedge \omega^k(v_1, \cdots, v_k) = \frac{1}{(k-1)!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \omega^1 \wedge \cdots \wedge \omega^{k-1}(v_{\sigma(1)}, \ldots, v_{\sigma(k-1)}) \cdot \omega^k(v_{\sigma(k)})$$

Denote the left hand side of the above equation by Ω for brevity, then applying the inductive step we see that:

$$\Omega = \frac{1}{(k-1)!} \sum_{\sigma \in S_k} \sum_{\tau \in S_{k-1}} \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_{\sigma(k-1)}) \cdot \omega^k(v_{\sigma(k)})$$

For each $\sigma \in S_k$ there are $(k-1)!$ factorial $\tilde{\sigma}$'s, which satisfy $\sigma(k) = \tilde{\sigma}(k)$, including σ. We can then split S_k into k subsets of S_k, each consisting of the permutations which satisfy the aforementioned property. Denote each set by A_i, where each $\sigma \in A_i$ satisfies $\sigma(k) = i$, then we can rewrite our sum as:

$$\Omega = \frac{1}{(k-1)!} \sum_{i=1}^{k-1} \sum_{\sigma \in A_i} \sum_{\tau \in S_{k-1}} \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_{\sigma(k-1)}) \cdot \omega^k(v_i)$$

Fix an i, then:

$$\begin{aligned}\sum_{\sigma \in A_i} \sum_{\tau \in S_{k-1}} & \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_{\sigma(k-1)}) \cdot \omega^k(v_{\sigma(k)}) \\ &= \omega^k(v_i) \sum_{\sigma \in A_i} \sum_{\tau \in S_{k-1}} \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_i)\end{aligned}$$

Fix $\tau, \tilde{\tau} \in S_{k-1}$, we claim that:

$$\begin{aligned}\sum_{\sigma \in A_i} & \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) \omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_{\sigma(k-1)}) \\ &= \sum_{\sigma \in A_i} \operatorname{sgn}(\tilde{\tau})\operatorname{sgn}(\sigma) \omega^{\tilde{\tau}(1)}(v_{\sigma(1)}) \cdots \omega^{\tilde{\tau}(k-1)}(v_{\sigma(k-1)})\end{aligned}$$

We proceed by cases, let $\operatorname{sgn}(\tau) = \operatorname{sgn}(\tilde{\tau})$, then τ and $\tilde{\tau}$ differ by an even about of swaps. For any $\sigma \in A_i$, there then exists a unique $\tilde{\sigma} \in A_i$ satisfying:

$$\omega^{\tau(1)}(v_{\sigma(1)}) \cdots \omega^{\tau(k-1)}(v_{\sigma(k-1)}) = \omega^{\tilde{\tau}(1)}(v_{\tilde{\sigma}(1)}) \cdots \omega^{\tilde{\tau}(k-1)}(v_{\tilde{\sigma}(k-1)})$$

which must also differ by an even number swaps, thus $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tilde{\sigma})$, and, consequently $\operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tilde{\tau})\operatorname{sgn}(\tilde{\sigma})$. Therefore, we see that for any term in the left sum, the same term appears in the right sum, and since the order of summation doesn't matter the two sums are equal. If instead $\operatorname{sgn}(\tau) = -\operatorname{sgn}(\tilde{\tau})$, then τ and $\tilde{\tau}$ differ by an odd amount of swaps. The same argument then shows that for every $\sigma \in A_i$, there exists a unique $\tilde{\sigma} \in A_i$ satisfying the same relation, such that $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\tilde{\sigma})$. Thus $\operatorname{sgn}(\tau)\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tilde{\tau})\operatorname{sgn}(\tilde{\sigma})$, and for any term in the left sum, the same term appears in the right sum, so the sums are the same. Since for every τ, the sum over A_i is equal, we obtain:

$$\begin{aligned}\sum_{\sigma\in A_i}\sum_{\tau\in S_{k-1}} \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)\omega^{\tau(1)}(v_{\sigma(1)})\cdots\omega^{\tau(k-1)}(v_{\sigma(k-1)})\\ =&(k-1)!\sum_{\sigma\in A_i}\operatorname{sgn}(\sigma)\omega^1(v_{\sigma(1)})\cdots\omega^{k-1}(v_{\sigma(k-1)})\end{aligned}$$

Therefore:

$$\begin{aligned}\Omega =&\sum_{i=1}^{k-1}\sum_{\sigma\in A_i}\operatorname{sgn}(\sigma)\omega^1(v_{\sigma(1)})\cdots\omega^{k-1}(v_{\sigma(k-1)})\cdot\omega^k(v_i)\\ =&\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\omega^1\otimes\cdots\omega^{k-1}\otimes\omega^k(v_{\sigma(1)},\cdots,v_{\sigma(k)})\end{aligned}$$

Since computationally it doesn't matter whether we permute the vectors, or the covectors, we can write:

$$\Omega = \sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\omega^{\sigma(1)}\otimes\cdots\otimes\omega^{\sigma(k-1)}\otimes\omega^{\sigma(k)}(v_1,\cdots,v_k)$$

Finally, since the set of vectors was chosen arbitrarily , we see that:

$$\omega^1\wedge\cdots\wedge\omega^k = \sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\omega^{\sigma(1)}\otimes\cdots\otimes\omega^{\sigma(k)}$$

as desired. □

The set:

$$\{e^{i_1}\wedge\cdots\wedge e^{i_k} : 1\le i_1<\cdots<i_k\le k\}$$

forms a basis for $\Lambda^k(V^*)$. It is easy to show that this set spans $\Lambda^K(V^*)$, and one shows that is linearly dependent by supposing that:

$$\omega = \sum_{i_1<\cdots<i_k} a_{i_1\cdots i_k}e^{i_k}\wedge\cdots\wedge e^{i_k} = 0$$

and then noting that by **Proposition 1.1.11**:

$$e^{i_k}\wedge\cdots\wedge e^{i_k}(e_{i_1},\ldots,e_{i_k}) = 1$$

so since ω must be the zero alternating multilinear map $V^k\to\mathbb{R}$, we obtain that each $a_{i_1\cdots i_k}=0$. Note that this implies that $\dim\Lambda^k(V^*) = n$ choose k. Furthermore, an element of $\Lambda^k(V)$ is said to *decomposable* if it can be written as in **Proposition 1.1.11**; clearly if $\omega\in\Lambda^k$ is decomposable then:

$$\omega\wedge\omega = 0$$

Definition 1.1.19. The vector space of all forms on V is given by the direct sum:

$$\Lambda(V^*) = \bigoplus_{i=0}^{n}\Lambda^i(V^*)$$

where $\Lambda^0(V^*)$ is the field of scalars over V. Equipped with the wedge product, it is an associative, graded, algebra with unit element $1\in\mathbb{R}$. We call this space the **exterior algebra**, and it has dimension:

$$\dim(\Lambda(V^*)) = \sum_{k=0}^{n}\binom{n}{k} = 2^n$$

Definition 1.1.20. Let V be a finite dimensional vector space, and ω a k-form on V. Given a linear map $A : W \to V$, we can obtain a k-form on W via the **pullback** of ω by A, denoted $A^*\omega$, in the following way:

$$A^*\omega(v_1, \dots, v_k) = \omega(A(v_1), \dots, A(v_k))$$

This definition demonstrates a fundamental difference between covectors (one forms) and vectors; that is vectors can be pushed forward by linear maps, and covectors, as well as higher order forms, are pulled back by linear maps.

Definition 1.1.21. Let V be a finite dimensional vector space, $v \in V$, and ω a k-form on V. We define the **contraction**, or **interior product** of ω, denoted $\iota_v\omega$, or $v \lrcorner \omega$, as a map from $\Lambda^k(V^*) \to \Lambda^{k-1}(v)$ via:

$$\iota_v\omega(v_1, \dots, v_{k-1}) = \omega(v, v_1, \dots, v_{k-1})$$

For simple k-form's contraction yields the following $(k-1)$ form:

$$\iota_v(e^{i_1} \wedge \dots \wedge e^{i_k}) = \sum_{j=1}^{k} (-1)^{j+1} e^{i_j}(v) e^{i_1} \wedge \dots \wedge \hat{e}^{i_j} \wedge \dots \wedge e^{i_k}$$

where $\hat{e}^i$ denotes deletion of the ith component. We turn to an example:

Example 1.1.11. Let $V = \mathbb{R}^3$, the standard basis vectors e_1, e_2 and e_3 admit a dual basis e^1, e^2, and e^3. Define a two form ω on $\mathbb{R}^3$ by:

$$\omega = e^1 \wedge e^2 + e^2 \wedge e^3 + e^1 \wedge e^3$$

For vectors $v = a^i e_i$, and $w = b^i e_i$ we have the following:

$$\begin{aligned}
\iota_v\omega =& a^1 e^2 - a^2 e_1 + a^2 e^3 - a^3 e^2 + a^1 e^3 - a^3 e^1 \\
\iota_w(\iota_v\omega) =& a^1 b^2 - a^2 b^1 + a^2 b^3 - a^3 b^2 + a^1 b^3 - a^3 b^1 = \omega(v, w)
\end{aligned}$$

We can also construct a three form on $\mathbb{R}^3$ given by wedging ω with e^3:

$$\begin{aligned}
\omega \wedge e^3 =& e^1 \wedge e^2 \wedge e^3 + e^2 \wedge e^3 \wedge e^3 + e^1 \wedge e^3 \wedge e^3 \\
=& e^1 \wedge e^2 \wedge e^3
\end{aligned}$$

For another vector $u = c^i e_i$ we have that:

$$\begin{aligned}
u \lrcorner (w \lrcorner (v \lrcorner \omega \wedge e^3)) =& u \lrcorner (w \lrcorner (a^1 e^2 \wedge e^3 - a^2 e^1 \wedge e^3 + a^3 e^1 \wedge e^2)) \\
=& u \lrcorner (a^1(b^2 e^3 - b^3 e^2) - a^2(b^1 e^3 - b^3 e^1) + a^3(b^1 e^2 - b^2 e^1)) \\
=& a^1(b^2 c^3 - b^3 c^2) - a^2(b^1 c^3 - b^3 c^1) + a^3(b^1 c^2 - b^2 c^1) \\
=& \det(v, w, u)
\end{aligned}$$

We see that this aligns with **Definition 1.1.18**, as with $v_1 = v$, $v_2 = w$ and $v_3 = u$:

$$\begin{aligned}
\omega \wedge e^3(v_1, v_2, v_3) =& \frac{1}{2} \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) \omega(v_{\sigma(1)}, v_{\sigma(2)}) e^3(v_{\sigma(3)}) \\
=& \omega(v_1, v_2) e^3(v_3) + \omega(v_2, v_3) e^3(v_1) + \omega(v_3, v_1) e^3(v_2) \\
=& (a^1 b^2 - a^2 b^1) c^3 + (b^1 c^2 - b^2 c^1) a^3 + (c^1 a^2 - c^2 a^1) b^3 \\
=& \det(v_1, v_2, v_3)
\end{aligned}$$

Finally, this aligns with our form of the wedge product for simple k forms as:

$$\begin{aligned}
e^1 \wedge e^2 \wedge e^3(v_1, v_2, v_3) =& \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) e^{\sigma(1)}(v_1) e^{\sigma(2)}(v_2) e^{\sigma(3)}(v_3) \\
=& a^1 b^2 c^3 + b^1 c^2 a^3 + c^1 a^2 b^3 - b^1 a^2 c^3 - a^1 c^2 b^3 - c^1 b^2 a^3 \\
=& \det(v_1, v_2, v_3)
\end{aligned}$$

We will further discuss the exterior algebra and the wedge product when we discuss Clifford algebras later in the paper, but for now we move onwards. Recall from the previous section that for a smooth n manifold M, T_pM is a real vector space of dimension n. We denote the dual space by T_p^*M, and after choosing a coordinate basis $(\partial/\partial x^i|_p)$, denote it's dual basis by dx^i. Any covector $\omega \in T_p^*M$ can thus be written as $\omega_i dx^i$, where:

$$\omega_i = \omega\left(\frac{\partial}{\partial x^i}\Big|_p\right)$$

Let $(\partial/\partial y^i)$ be another coordinate basis for T_pM, recall that for:

$$v = \frac{\partial}{\partial y^i}$$

we have:

$$v = \frac{\partial x^j}{\partial y^i}\frac{\partial}{\partial x^j}$$

in our former coordinate basis. Let $\omega = \omega_l dx^l$, then:

$$\begin{aligned}\omega(v) =&\omega\left(\frac{\partial x^j}{\partial y^i}\frac{\partial}{\partial x^j}\right)\\ =&\omega_l\frac{\partial x^j}{\partial y^i}dx^l\left(\frac{\partial}{\partial x^j}\right)\\ =&\omega_j\frac{\partial x^j}{\partial y^i}\end{aligned}$$

Thus in our y^i coordinates we have:

$$\omega = \omega_j\frac{\partial x^j}{\partial y^i}dy^i \tag{1.1.14}$$

hence the components of covectors transform in an 'opposite' way from vectors. In physics, tensors which transform like vectors are called *contravariant*, and tensors which transform like covectors are called *covariant*.

Definition 1.1.22. The **Cotangent Bundle** is the disjoint union:

$$T^*M = \coprod_{p\in M} T_p^*M$$

It has a natural projection map $\pi : T^*M \to M$, which sends $\omega \in T_p^*M$ to $p \in M$.

Though we omit the proof, T^*M has a natural topology and smooth structure that make it into a smooth manifold of dimension $2n$. When we discuss fibre bundles, and vector bundles, we shall see that both TM, and T^*M are special cases of vector bundles over M. Further, we define covector fields, or differential one forms on M, in a similar way to vector fields:

Definition 1.1.23. Given a smooth manifold M, a **covector field**, often called a **differential one form**, is a smooth section of T^*M, i.e. a map:

$$s : M \to T^*M$$

such that:

$$\pi \circ s = \mathrm{Id}$$

We denote the space of all one forms on M by $\Omega^1(M)$

Much like vector fields, we define local and global frames for T^*M in the same way but instead call them *coframes*; we denote coordinate coframes by $\{dx^i\}$. Furthermore, given a differential one form, $\omega \in T^*M$, and a vector field $X \in \mathfrak{X}(M)$, the contraction of ω with X, is a smooth function on M. Indeed, let (U, ϕ), and (V, ψ) be charts on M, such that $U \cap V \neq \emptyset$, and have coordinates

x^i, and y^j respectively. In $U \cap V$, let $\omega = \omega_i dx^i$, and $X = X^j \partial/\partial x^j$, for $X^j, \omega_i \in C^\infty(U \cap V)$, then:

$$\omega(X) = \omega_i X^i$$

By (1.1.13) and (1.1.6), under the transition map, this becomes:

$$\omega(X) = X^k \frac{\partial y^j}{\partial x^k} \omega_i \frac{\partial x^i}{\partial y^j}$$

Which, by the multivariate chain rule, reduces to:

$$\begin{aligned}\omega(X) =& X^k \omega_i \frac{\partial x^i}{\partial x^k} \\ =& X^k \omega_i \delta^i_k \\ =& X^i \omega_i\end{aligned}$$

hence, $\omega(X)$ is smoothly compatible, and independent of coordinates, and thus a smooth function on M. Given a smooth function $f \in C^\infty$, we can also construct differential one forms, with the differential of f, defined via:

$$df(X) = X(f)$$

In coordinates, with $X = X^i \partial/\partial x^i$, we see that:

$$df(X) = X^i \frac{\partial f}{\partial x^i}$$

implying that:

$$df = \frac{\partial f}{\partial x^i} dx^i$$

Now examine the following disjoint union:

$$\Lambda^k(T^*M) = \coprod \Lambda^k(T^*_p M)$$

Just as TM, and T^*M, $\Lambda^k(T^*M)$ is a smooth manifold, equipped with a projection map, such that $\Lambda^k(T^*M)$ is a vector bundle over M. A differential k-form on M, is then a smooth section of $\Lambda^k(T^*M)$, and we denote the set of all of differential k-forms on M by $\Omega^k(M)$. The wedge product of two differential forms is defined pointwise:

$$(\omega \wedge \eta)_p = \omega_p \wedge \eta_p$$

and just as we saw earlier in exterior algebra, the wedge product of a differential k form and differential l form is a differential $(k+l)$ form. Taking smooth functions on M to be zero forms, such that for $f \in C^\infty(M)$:

$$f \wedge \omega = f\omega$$

we define:

$$\Omega(M) = \bigoplus_{i=0}^{n} \Omega^i(M)$$

which, when equipped with the wedge product, is the associative, anticommutative, graded algebra of differential forms of all order. Let ω be a differential k-form, and I be an ordered multi index $(i_1, \dots, i_k)$[3] then we employ the Einstein summation convention in the following way:

$$\omega = \omega_I dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \omega_I dx^I$$

[3]i.e. $i_1 < \cdots < i_k$

In general, unless stated otherwise, we will understand this summation to be over all possible ordered multi indices, so that we do not over count terms in the wedge product. Let I, and J be two ordered multi indices, then:

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} \left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_k}} \right) = \delta^I_J$$

Therefore the component functions ω_I of ω are found via:

$$\omega_{i_1,\ldots,i_k} = \omega \left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_k}} \right)$$

Example 1.1.12. Let $M = \mathbb{R}^3$, a zero form on M is any smooth function $f : \mathbb{R}^3 \to \mathbb{R}$. We can define a one form via:

$$\omega = \sin(x)dx + \sin(y)dy + \sin(z)dz$$

We have that:

$$\begin{aligned}
\omega \wedge \omega =& (\sin(x)dx + \sin(y)dy + \sin(z)dz) \wedge (\sin(x)dx + \sin(y)dy + \sin(z)dz) \\
=& (\sin(x)\sin(y) - \sin(y)\sin(x))dx \wedge dy \\
&+ (\sin(x)\sin(z) - \sin(z)\sin(x))dx \wedge dz \\
&+ (\sin(y)\sin(z) - \sin(z)\sin(y))dy \wedge dz \\
=& 0
\end{aligned}$$

as expected. Let $\eta = dx \wedge dz$, then:

$$\begin{aligned}
\omega \wedge \eta &= (\sin(y)dy \wedge dx \wedge dz) \\
&= -\sin(y)dx \wedge dy \wedge dz
\end{aligned}$$

which is a three form on M. In particular, any n form ω on an n-dimensional smooth manifold can be written in coordinates as:

$$\omega = f dx^1 \wedge \cdots \wedge dx^n$$

for some $f \in C^\infty(M)$

Let M, and N be smooth manifolds, F a smooth map between them, and η a differential k form on N. The pullback of η by F is then a differential form on M given by:

$$F^*\omega_p(v_1, \ldots, v_k) = \omega_{F(p)}(D_pF(v_1), \ldots, D_pF(v_k))$$

Lemma 1.1.3. *Suppose $F : M \to N$ is smooth. Then in any smooth chart:*

$$F^*(\omega_I dy^{i_1} \wedge \cdots \wedge dy^{i_k}) = (\omega_I \circ F)d(y^{i_1} \circ F) \wedge \cdots \wedge d(y^{i_k} \circ F) \tag{1.1.15}$$

Proof. We first prove the statement for an arbitrary one form on N, $\eta = f_i dy^i$. Let $v \in T_pM$ be arbitrary then:

$$\begin{aligned}
(F^*\eta)_p(v) =& \eta_{F(p)}(D_pF(v)) \\
=& (f_i \circ F(p))(dy^i(D_pF(v)))_{F(p)} \\
=& (f_i \circ F(p))(D_pF(v))(y^i) \\
=& (f_i \circ F(p))v(y^i \circ F) \\
=& (f_i \circ F(p))d(y^i \circ F)(v)
\end{aligned}$$

Now let $\omega = \omega_I dy^I$ be a differential k form on N, by **Definition 1.1.18** we have that for $w_1, \cdots, w_k \in T_qN$:

$$\omega_I dy^{i_1} \wedge \cdots \wedge dy^{i_k}(w_1, \ldots, w_k) = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\omega_I dy^{1_k}(w_{\sigma(1)}) \cdots dy^{i_k}(w_{\sigma(k)})$$

where S_k denotes the set of permutations of $\{1,\dots,k\}$. Let $v_1,\dots,v_k\in T_pM$, then:

$$\begin{aligned}F^*\omega(v_1,\dots,v_k) &= \sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)(\omega_I\circ F)dy^{i_1}(D_pF(v_{\sigma(1)}))\cdots dy^{i_k}(D_pF(v_{\sigma(k)}))\\ &= \sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)(\omega_I\circ F)d(y^{i_1}\circ F)(v_{\sigma(1)})\cdots d(y^{i_k}\circ F)(v_{\sigma(k)})\\ &=(\omega_I\circ F)d(y^{i_1}\circ F)\wedge\cdots\wedge d(y^{i_k}\circ F)(v_1,\dots,v_k)\end{aligned}$$

hence we have that:

$$F^*\omega = (\omega_I\circ F)d(y^{i_1}\circ F)\wedge\cdots\wedge d(y^{i_k}\circ F)$$

as desired. □

Proposition 1.1.12. *If* $\dim M=\dim N=n$, *and* $F: M\to N$ *is a smooth map, then for a top form* $\omega = f dy^1\wedge\cdots\wedge y^n$ *on* N*, the coordinate expression for the pullback of* ω *by* F *is given by:*

$$F^*(f dy^1\wedge\cdots\wedge y^n) = (f\circ F)\det(DF)dx^1\wedge\cdots\wedge dx^n \tag{1.1.16}$$

where $x^1,\dots,x^n$ *are local coordinates on* M*.*

Proof. Let $v_1,\dots v_n\in T_pM$ have basis expansions given by:

$$v_l = v_l^i\frac{\partial}{\partial x^i}$$

then the LHS of (1.1.15) gives:

$$\begin{aligned}F^*(f dy^1\wedge\cdots\wedge dy^n)(v_1,\dots,v_n) &=(f\circ F)\det(D_pF(v_1),\dots D_pF(v_n))\\ &=(f\circ F)\det\left(\frac{\partial F^j}{\partial x^i}v_1^i\frac{\partial}{\partial y^j},\dots,\frac{\partial F^j}{\partial x^i}v_n^i\frac{\partial}{\partial y^j}\right)\\ &=(f\circ F)\det\left(\frac{\partial F^j}{\partial x^i}\right)\det(v_1,\dots,v_n)\\ &=(f\circ F)\det\left(\frac{\partial F^j}{\partial x^i}\right)dx^1\wedge\cdots\wedge dx^n(v_1,\dots,v_n)\end{aligned}$$

The matrix with coefficients $\partial F^j/\partial x^i$ is precisely the Jacobian DF, hence (1.1.15) holds. □

Example 1.1.13. Let $M=\mathbb{S}^2$, then $\psi^{-1}(\theta,\phi)=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ is an embedding of 'most' of the two sphere in $\mathbb{R}^3$, where $\theta\in(0,\pi)$ and $\phi\in(0,2\pi)$. Let $\omega = dx\wedge dy + dy\wedge dz$ be a two form on $\mathbb{R}^3$, if we want to restrict this two form to $\mathbb{S}^2$ we can pull ω back by ψ^{-1}, via (1.1.14):

$$\psi^{-1*}\omega = d(\sin\theta\cos\phi)\wedge d(\sin\theta\sin\phi) + d(\sin\theta\sin\phi)\wedge d(\cos\theta)$$

Examine the first term in the sum:

$$\begin{aligned}d(\sin\theta\cos\phi)\wedge d(\sin\theta\sin\phi) &=(\cos\theta\cos\phi d\theta - \sin\theta\sin\phi d\phi)\wedge(\cos\theta\sin\phi d\theta+\sin\theta\cos\phi d\phi)\\ &=\cos\theta\sin\theta\cos^2\phi d\theta\wedge d\phi - \sin\theta\cos\theta\sin^2\phi d\phi\wedge d\theta\\ &=\cos\theta\sin\theta d\theta\wedge d\phi\end{aligned}$$

Now the second term:

$$\begin{aligned}d(\sin\theta\sin\phi)\wedge d(\cos\theta) &=(\cos\theta\sin\phi d\theta+\sin\theta\cos\phi d\phi)\wedge(-\sin\theta d\theta)\\ &=-\sin^2\theta\cos\phi d\phi\wedge d\theta\end{aligned}$$

thus we have that:

$$\psi^{-1*}\omega = (\cos\theta\sin\theta+\sin^2\theta\cos\phi)d\theta\wedge d\phi$$

Note that the differential of a function takes differential 0-form, f, to a 1-form, df. We now wish to generalize this operation with the following theorem:

Theorem 1.1.3. *Suppose M is a smooth manifold, there exists a unique operator $d : \Omega^k(M) \to \Omega^{k+1}(M)$ for all k, called the **exterior derivative**, satisfying the following five conditions:*

a) d is linear over $\mathbb{R}$

b) if $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M)$ then:

$$d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^k \omega \wedge d\eta \tag{1.1.17}$$

c) $d \circ d \equiv 0$

d) For $f \in \Omega^0(M) = C^\infty(M)$, df is the differential of f, given by $df(X) = Xf$

e) In a smooth coordinate chart, with $\omega = \omega_I dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, d is given by:

$$d\omega = \frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \tag{1.1.18}$$

Proof. We begin with existence, suppose $\omega \in \Omega^k(M)$, we wish to define $d\omega$ via (1.1.18) in each chart, thus e will hold trivially. Let (U, ϕ) be a smooth chart for M, we set:

$$d\omega = \phi^* d(\phi^{-1*}\omega)$$

First note that by (1.1.18), d commutes with the pullback of a smooth map F. Indeed:

$$\begin{aligned} F^*(d\omega) =& F^* \left(\frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \right) \\ =& F^* \left(d\omega_I \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \right) \\ =& d(\omega_I \circ F) \wedge d(x^{i_1} \circ F) \wedge \cdots \wedge d(x^{i_k} \circ F) \end{aligned}$$

while:

$$\begin{aligned} d(F^*\omega) =& d((\omega_I \circ F) d(x^{i_1} \circ F) \wedge \ldots d(x^{i_k} \circ F)) \\ =& d(\omega_I \circ F) \wedge d(x^{i_1} \circ F) \wedge \ldots d(x^{i_k} \circ F) \end{aligned}$$

thus:

$$d(F^*\omega) = F^*(d\omega)$$

For two charts (U, ϕ), and (V, ψ), we have that:

$$\begin{aligned} \psi^* \circ (\phi \circ \psi^{-1})^* (d\phi^{-1*}\omega) =& \psi^{-1*} d((\phi \circ \psi^{-1})^* \phi^{-1*}\omega) \\ =& \psi^* d(\psi^{-1*}\omega) \end{aligned}$$

however:

$$\begin{aligned} \psi^* \circ (\phi \circ \psi^{-1})^* (d\phi^{-1*}\omega) =& \psi^* \circ \psi^{-1*} \circ \phi^* d(\phi^{-1*}\omega) \\ =& \phi^* d(\phi^{-1*}\omega) \end{aligned}$$

hence:

$$\psi^* d(\psi^{-1*}\omega) = \phi^* d(\phi^{-1*}\omega)$$

therefore d is a well defined operation. To show *a)* we note that for two k forms ω, and η, and $a, b \in \mathbb{R}$, that $a\omega \in \Omega^k(M)$ and $b\eta \in \Omega^k(M)$. Furthermore, since $\Lambda^k(T_p^*M)$ is a vector space, $a\omega + b\eta \in \Omega^k(M)$, and can be written in a coordinate chart as:

$$a\omega + b\eta = (a\omega_I + b\eta_I) dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

Therefore:

$$\begin{aligned} d(a\omega + b\eta) =& d(a\omega_I + b\eta_I) dx^{i_1} \wedge \cdots \wedge dx^{i_k} \\ =& \frac{\partial (a\omega_I + b\eta_I)}{\partial x^j} dx^j \wedge \cdots \wedge dx^{i_k} \\ =& a \frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^I + b \frac{\partial \eta_I}{\partial x^j} dx^j \wedge dx^I \\ =& a d\omega + b d\eta \end{aligned}$$

thus d is linear over $\mathbb{R}$. To show *b)*, we recall:

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

we then have that:

$$\begin{aligned} d(\omega \wedge \eta) =& d(\omega_I \eta_J dx^I \wedge dx^J) \\ =& (\eta_J d\omega_I + \omega_I d\eta_J) \wedge dx^I \wedge dx^J \\ =& \eta_J d\omega_I \wedge dx^I \wedge dx^J + \omega_I d\eta_J \wedge dx^I \wedge dx^J \\ =& \eta_J d\omega_I \wedge dx^I \wedge dx^J + (-1)^k \omega_I dx^I \wedge d\eta_J \wedge dx^J \\ =& d\omega \wedge \eta + (-1)^k \omega \wedge d\eta \end{aligned}$$

since $d\eta_J$ is a one form, hence $k \cdot 1 = k$. For *c)* we note that for a 0-form f:

$$\begin{aligned} d(df) =& d\left(\frac{\partial f}{\partial x^i} dx^i\right) \\ =& \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \wedge dx^j \\ =& \sum_{i<j} \left(\frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 f}{\partial x^j \partial x^i}\right) dx^i \wedge dx^j \\ =& 0 \end{aligned}$$

as partial derivatives commute with one another. Now, by (1.1.17), we see that for a k form ω:

$$\begin{aligned} d(d\omega) =& d\left(d(\omega_I) \wedge dx^I\right) \\ =& d(d\omega_I) \wedge dx^I - d\omega_I \wedge d(dx^I) \\ =& -d\omega_I \wedge d(dx^{i_1} \wedge \cdots dx^{i_k}) \\ =& \sum_j^k (-1)^j d\omega_I \wedge dx^{i_1} \wedge \cdots \wedge d(dx^{i_j}) \wedge \cdots \wedge dx^{i_k} \\ =& 0 \end{aligned}$$

thus $d \circ d$ is identically zero. Finally, *d)* clearly follows from d being well defined, equation (1.1.18), and our earlier discussion on the differential of a smooth function f.

To show uniqueness, suppose that d is any operator satisfying *a), b),c),* and *d)*, and that ω_1 and ω_2 are two k forms on an open set $U \subset M$. We would like to show that d is determined locally, that is if ω_1 and ω_2 agree on U, then $d\omega_1 = d\omega_2$ on U as well. Let $\eta = \omega_1 - \omega_2$, and for an arbitrary point $p \in U$, let $\psi \in C^\infty(M)$ be a function that is identically 1 on a open neighborhood of p, and zero outside of U. Then, $\psi\eta = 0$ on all U, and we have that:

$$\begin{aligned} 0 =& d(\psi\eta) \\ =& d\psi \wedge \eta + \psi d\eta \\ =& \psi d\omega_1 - \psi d\omega_2 \end{aligned}$$

Evaluating at the point p, we find that $d\omega_1|_p - d\omega_2|_p = 0$, hence d is determined locally. Now let $\omega \in \Omega^k(M)$ be an arbitrary k form on M, then in some smooth coordinate chart (U, ϕ) we can write ω as:

$$\omega = \omega_I dx^I$$

For any $p \in U$,by means of functions like ψ, (i.e. identically 1 in a neighborhood of p, and identically 0 outside of U) we can then construct global smooth functions y^i and f_I on M such that they agree with ω_I and x^i in a neighborhood of p. By the preceding paragraph, it then suffices to show that d agrees with (1.1.18) at the point p. By *a), b),c),* and *d)* we then see that:

$$\begin{aligned} d\omega|_p =& d(\omega_I dx^I)|_p \\ =& d\omega_I \wedge dx^I|_p + \omega \wedge d(dx^I)|_p \\ =& \frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^I|_p \end{aligned}$$

Since p was arbitrary, d must be unique, as desired. □

As we shall see in the next example, the exterior derivative can be used to generalize the major vector calculus derivative operations.

Example 1.1.14. Let $M = \mathbb{R}^3$, then the following vector spaces $\Lambda^1(T_p^*\mathbb{R}^3)$, $\Lambda^2(T_p^*\mathbb{R}^3)$, and $\Lambda^2(T_p^*\mathbb{R}^3)$ are of dimension 3, 3, and 1, respectively. We see that for some function $f : \mathbb{R}^3 \to \mathbb{R}$ that:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

which under an isomorphism $F : \Omega^1(\mathbb{R}^3) \to \mathfrak{X}(M)$, can be written as the vector field:

$$F(df) = \frac{\partial f}{\partial x}\frac{\partial}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial}{\partial z}$$

which is exactly the result of the gradient operation in vector calculus. Furthermore, for a one a form ω given in coordinates by:

$$\omega = f dx + g dy + h dz$$

we see that:

$$d\omega = \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy \wedge dz + \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right) dx \wedge dz + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy \wedge dz$$

Define an isomorphism $G : \mathfrak{X}(M) \to \Omega^2(M)$ by:

$$X \longmapsto X \lrcorner (dx \wedge dy \wedge z)$$

which provides the following identification of a global frame of vector fields on $\mathbb{R}^3$, with a global frame of two form on $\mathbb{R}^3$:

$$\frac{\partial}{\partial x} \leftrightarrow dy \wedge dz$$
$$\frac{\partial}{\partial y} \leftrightarrow -dx \wedge dz$$
$$\frac{\partial}{\partial z} \leftrightarrow dx \wedge dy$$

Hence we have that:

$$G(d\omega) = \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)\frac{\partial}{\partial x} - \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right)\frac{\partial}{\partial y} + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)\frac{\partial}{\partial z}$$

which is exactly the curl operation $\nabla\times$ for vector fields on $\mathbb{R}^3$. Finally, for a two form given by:

$$\eta = f dy \wedge dz + g dz \wedge dx + h dx \wedge dy$$

we see that :

$$d\eta = \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}\right) dx \wedge dy \wedge dz \tag{1.1.19}$$

Under the isomorphism $H : C^\infty(M) \to \Omega^3(M)$ given by:

$$f \longmapsto f dx \wedge dy \wedge dz$$

we have that (1.1.19) maps to the function given by the divergence of the vector field corresponding to η, thus the exterior derivative of 2 forms on $\mathbb{R}^3$ corresponds to the divergence operation $\nabla\cdot$ for vector fields on $\mathbb{R}^3$.

Before we can discuss integration, we must briefly dive into the notion of assigning an orientation to M, much like we assign an orientation to a vector space V.

Definition 1.1.24. Let M be a smooth manifold. A **pointwise orientation** is a choice of orientation for each tangent space. Given a pointwise orientation, if a local frame $\{E_i\}$ for TM agrees with the orientation for all $p \in U$ then $\{E_i\}$ is an **oriented frame**. A pointwise orientation such that all $p \in M$ is contained in the domain of a local oriented frame is a **orientation** for M, and if M is said to be an **orientable smooth manifold** if such an orientation exists.

Proposition 1.1.13. *Let M be a smooth manifold, then M is orientable if and only there exists a covering of M by coordinate charts $\{U_i, \phi_i\}$ such that the Jacobian of each transition function is positive.*

Proof. Suppose that M is orientable, and is equipped with an orientation. Then for each point $p \in M$ there exists an orientation of T_pM. Every $p \in M$ is contained in some coordinate chart (U, ϕ), and if that chart is negatively oriented, i.e. the the orientation of the coordinate frame $\{\partial/\partial x^i\}$ differs from the orientation on M, we can define a new chart by:

$$\phi' = (-x^1, x^2, \dots, x^n)$$

which then agrees with the orientation on M. Repeating this process for all charts, we obtain a covering of M by coordinate charts such that each chart determines a positively oriented coordinate frame, thus the Jacobian of the transition functions must have positive determinant.

Conversely, assume that M admits a covering of coordinate charts $\{U_i, \phi_i\}$ such that Jacobian of each transition function is positive. Define a pointwise orientation on M such that the coordinate frame at each point is positively oriented. The Jacobian of the transition functions have positive determinant, so every chart determines the same pointwise orientation. Furthermore, every point in M is in the domain of a coordinate frame, which by construction is an oriented frame for TM, hence M is orientable. $\square$

Using this proposition we wish to prove the existence of a non-vanishing top form on M, corresponding to the orientation of M, but first we introduce a partition of unity. Let $f : X \to \mathbb{R}$ for a topological space X, we define the support of f, denoted supp f, as the closure of the set of points where $f \neq 0$.

Definition 1.1.25. Suppose M is a topological space, and let $X = (X_i)_{i\in I}$ be an arbitrary open cover of M, indexed by the set I. A **partition of unity subordinate to X** is an indexed family $\{\psi_i\}_{i\in I}$ of continuous functions $\psi_i : M \to \mathbb{R}$ such that:

a) $0 \leq \psi_i(x) \leq 1$ for all $x \in M$, $i \in I$.

b) supp $\psi_i \subset X_i$ for each $i \in I$.

c) The family of supports (supp $\psi_i)_{i\in I}$ is locally finite, meaning that that every point has a neighborhood that intersects supp ψ_i for only finitely many values of i.

d) $\sum_{i\in I} \psi_i(x) = 1$ for all $x \in M$.

It can be shown that for any indexed open cover of a smooth manifold M, there exists a smooth partition of unity subordinate to the aforementioned open cover. We omit the proof of this statement, and move onwards.

Theorem 1.1.4. *A smooth manifold M of dimension n is orientable if and only if there exists a non-vanishing top form on M; we call such a form an orientation form. Equivalently, M is orientable if and only if M has a trivial top form bundle.*

Proof. Suppose such a form ω exists and let (U, ϕ) be a smooth chart for M, then we have that on U:

$$\omega = f dx^1 \wedge \cdots \wedge dx^n$$

for some $f \in C^\infty(U)$, such that $f > 0$ for all $p \in U$. Let (V, ψ) be another chart on M, such that $U \cap V \neq \emptyset$, then by **Proposition 1.1.12**:

$$(\phi \circ \psi^{-1})^*\omega = (f \circ \phi \circ \psi^{-1}) \det\left(\frac{\partial x^j}{\partial y^i}\right) dy^1 \wedge \cdots \wedge dy^n$$

is a top form on the $U \cap V$ in the y^i coordinates determined by ψ. Since it is assumed that ω vanishes nowhere, we see that for all points $p \in U \cap V$, the determinant of the Jacobian of $\phi \circ \psi^{-1}$ must must always be positive. Furthermore, since ω vanishes nowhere, there must exist charts that cover M such that the Jacobian of the transition function is always positive, hence M is orientable.

Suppose M is orientable. We can cover M with countably many smooth charts, $\{(U_i, \phi_i)\}_{i\in I}$, such that the transition functions all of have positive determinant. This is an open cover of M.

Let $\{\psi_i\}_{i\in I}$ be a partition of unity subordinate to the aforementioned open cover. In a smooth chart (U_j, ϕ_j), we locally define a top form ω_j by:

$$\omega_j = dx^1 \wedge \cdots \wedge dx^n$$

Now consider the global top form η_j defined by:

$$\eta_j = \psi_j \omega_j$$

Then the the top form η determined by the sum:

$$\eta = \sum_{j\in I} \eta_j = \sum_{j\in I} \psi_j \omega_j \tag{1.1.20}$$

is also globally defined. To show this form is non-vanishing, we consider an arbitrary point $p \in M$, at this point p, by the definition of a partition of unity, there exists a neighborhood of p that intersects the support of ψ_i for only finitely many $i \in I$. Therefore, at this point p, (1.1.20) becomes a finite sum, thus we reindex I such that for some $n \in \mathbb{N}$:

$$\eta_p = \sum_i^n \psi_i(p)\omega_i$$

We now rewrite this in the coordinate basis for (U_1, ϕ_1):

$$\eta_p = \sum_i^n \psi_i(p) \det\big(D_p(\phi_i \circ \phi_1^{-1})\big) dx^1 \wedge \cdots \wedge dx^n$$

which cannot be 0 or negative as the determinant of the Jacobian of every transition function is positive. Thus, since p was arbitrary we have that η vanishes nowhere.

The final statement comes from the fact that the top form bundle over M is:

$$\Lambda^n(T^*M) = \coprod_{p\in M} \Lambda^n(T_p^*M),$$

hence since $\Lambda^n(T_p^*M)$ is a one dimensional vector space, we must show that $\Lambda^n(T^*M) \cong M \times \mathbb{R}$. If there exists a nowhere vanishing top form ω, then every $\eta \in \Lambda^n(T^*M)$, can be written as $a\omega_p \in \Lambda^n(T^*M)$ for some $p \in P$ and $a \in \mathbb{R}$; η is only the zero top form if $a = 0$. Thus we construct an isomorphism $\alpha : \Lambda^n(T^*M) \to M \times \mathbb{R}$ by:

$$\alpha : (p, a\omega_p) \longmapsto (p, a)$$

with inverse given by:

$$\alpha^{-1} : (p, a) \longmapsto (p, a\omega_p)$$

hence $\Lambda^n(T^*M)$ is trivial. If $\Lambda^n(T^*M)$ is trivial, then there exists an isomorphism $\alpha : \Lambda^n(T^*M) \to M \times \mathbb{R}$, thus for any $(p, a) \in M \times \mathbb{R}$, we have that:

$$\alpha^{-1}(p, a) \neq (p, 0)$$

unless $a = 0$. Thus there must exist a nowhere vanishing top form, namely the one given by $\omega_p = \alpha^{-1}(p, 1)$. Therefore, we have that if the top form bundle is trivial, there exists a nowhere vanishing top form, and thus M is orientable, and vice versa as desired. $\square$

We now see that differential forms are important for two reasons. First, a top form on an n dimensional manifold gives us a scaled determinant function on each tangent space to M, and can thus be thought of as providing us with a way to measure the volume of the parallepiped spanned by n tangent vectors. Secondly, a non-vanishing top form encodes an orientation of each tangent space to M, and can thus give us a signed volume of this parallepiped, depending on the orientation of the vectors. These two facts are key to allowing us to define integrals on a general orientable smooth manifold M, such that we are consistent with the case $M = \mathbb{R}^n$. Now, let M

be an oriented smooth manifold of dimension n, and let ω be an n form with compact support in a single chart (U, ϕ). We can then integrate ω over n via:

$$\int_M \omega = \pm \int_{\phi(U)} \phi^{-1*}\omega$$

where we have the plus sign for a positively oriented chart, and the minus sign for a negatively oriented chart. Furthermore, let ω be an n form with compact support finitely covered by charts $\{U_i\}$. The integral of ω over M is defined via a partition of unity subordinate to the aforementioned charts in the following way:

$$\int_M \omega = \sum_i \int_M \psi_i \omega$$

It can be shown that both integrals do not depend on choice of chart, open cover, or partition of unity. We now list some properties of the integral of an n form on M:

Theorem 1.1.5. *Suppose M and N are oriented smooth n manifolds, and ω, η are compactly supported n forms on M, then the following properties hold:*

a) If $a, b \in \mathbb{R}$ then:

$$\int_M a\omega + b\eta = a\int_M \omega + b\int_M \eta$$

b) If $-M$ denotes M with the opposite orientation then:

$$\int_{-M} \omega = -\int_M \omega$$

c) If ω is a positively oriented orientation form, then:

$$\int_M \omega > 0$$

d) If $F^+ : N \to M$ is an orientation preserving diffeomorphism and F^- is a orientation reversing diffeomorphism then:

$$\int_M \omega = \pm \int_M F^{\pm *}\omega$$

The proceeding definitions of the integral of a top form over a smooth manifold M are in general very difficult to actually carry out computation wise. Indeed, the extra factor due to the partition of unity often will cause trouble when calculating an integral. Instead, it is usually easier to integrate using a parameterization of M, that leaves out sets of measure zero[4], such as the half arc of the great circle that is not covered by the angle parameterization of $\mathbb{S}^2$. We describe this methodology in the following theorem, which we state without proof:

Theorem 1.1.6. *Let M be an oriented smooth manifold of dimension n, and ω a compactly supported top form on M. Suppose $D_1, \ldots, D_k$ are open domains of integration in $\mathbb{R}^n$, and for $i = 1, \ldots, k$, we have smooth maps $F_i = \bar{D}_i \to M$, where $\bar{D}_i$ is the closure of D_i, satisfying:*

a) F_i restricts to an orientation preserving diffeomorphism from D_i onto an open subset $W_i \subset M$

b) $W_i \cap W_j = \emptyset$ for $i \neq j$

c) $\operatorname{supp} \omega \subset \bar{W}_1 \cup \cdots \cup \bar{W}_k$

then:

$$\int_M \omega = \sum_{i=1}^k \int_{D_i} F_i^* \omega$$

[4]We define sets of measure zero on a smooth manifold M of dimension n, as sets that are covered by a smooth chart which restricts to some subset of $\mathbb{R}^n$ that has n dimensional Lebesgue measure zero in $\mathbb{R}^n$.

We now turn to an example:

Example 1.1.15. Let $\mathbb{S}^2$, oriented by the boundary of $\bar{B}^3$, be parameterized by the orientation preserving map $\phi^{-1} : \mathbb{R}^2 \to \mathbb{S}^2 \setminus (0,0,1)$:

$$\phi^{-1}(u,v) = \left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{u^2+v^2-1}{1+u^2+v^2}\right)$$

The singleton set $(0,0,1)$ is clearly of measure zero in $\mathbb{S}^2$, thus we may apply **Theorem 1.1.6** for the following 2-form. Let ω be the two form on $\mathbb{R}^3$ defined by:

$$\omega = xdy \wedge dz + ydz \wedge dx + zdx \wedge dy$$

then we have that:

$$\begin{aligned}
\phi^{-1*}(xdy \wedge dz) =& \frac{16(u^2+u^2v^2+u^4)}{(1+u^2+v^2)^5} du \wedge dv \\
\phi^{-1*}(ydx \wedge dz) =& \frac{16(v^2+u^2v^2+v^4)}{(1+u^2+v^2)^5} du \wedge dv \\
\phi^{-1*}(zdx \wedge dy) =& \frac{4(1-u^2-v^2)}{(1+u^2+v^2)^4} du \wedge dv
\end{aligned}$$

An algebraic manipulation then shows that:

$$\phi^{-1*}(\omega) = \frac{4}{(1+u^2+v^2)^2} du \wedge dv$$

hence, we obtain the improper integral:

$$\begin{aligned}
\int_{\mathbb{S}^2} \omega =& \int_{\mathbb{R}^2} \frac{4}{(1+u^2+v^2)^2} du \wedge dv \\
=& \int_0^\infty \int_0^\infty \frac{4}{(1+u^2+v^2)^2} dudv
\end{aligned}$$

The substitution given by $u = r\cos\theta$, $v = r\sin\theta$ gives:

$$\int_{\mathbb{S}^2} \omega = \int_0^{2\pi} \int_0^\infty \frac{4}{(1+r^2)^2} rdrd\theta \tag{1.1.21}$$

Again, applying a substitution $w = 1 + r^2$, we obtain that $dw = 2rdr$, thus (1.1.21) becomes:

$$\begin{aligned}
\int_{\mathbb{S}^2} \omega =& \int_0^{2\pi} \int_1^\infty \frac{2}{w^2} dwd\theta \\
=& 4\pi \int_1^\infty \frac{2}{w^2} dw \\
=& -4\pi \left[\frac{1}{w}\right]\Bigg|_1^\infty \\
=& 4\pi
\end{aligned}$$

We shall see later that ω actually corresponds to a canonical volume form on $\mathbb{S}^2$ given by the Riemannian metric induced on $\mathbb{S}^2$ by restricting the Euclidean metric to $\mathbb{S}^2$, thus this result is not coincidental.

We now turn to a highly important theorem in smooth manifold theory, that is **Stokes' theorem on manifolds**. Before though, we must briefly touch on what a smooth manifold with boundary is, which we define below:

Definition 1.1.26. Let M be a topological space; if M is:

a) M is second countable

b) M is Hausdorff

c) M is locally homeomorphic to open sets of $\mathbb{R}^n$ or the closed upper half plane $\mathbb{H}^n$ defined as the set:

$$\mathbb{H}^n\{(x^1,\ldots,x^n)\in\mathbb{R}^n : x^n \geq 0\}$$

then M is a **topological manifold with boundary** of dimension n. If M can be equipped with a maximal smooth atlas then M is a **smooth manifold with boundary** of dimension n.

Many of the previous results discussed on smooth manifolds carry over to smooth manifolds with boundary, but not all. For instance, the product of two smooth manifolds with boundary is not a smooth manifold with boundary, and **Proposition 1.1.13** only holds when M is without boundary, or $\dim M > 1$. The boundary of M, denoted ∂M, is the set of all points $p\in M$, such that there exists a coordinate chart (U,ϕ) containing p which satisfies:

$$\phi(p) = (x^1,\ldots,x^{n-1},0)$$

The interior of M, is the set of all points contained in a coordinate chart such that $\phi(U)$ is open in $\mathbb{R}^n$. One can show that these two sets are disjoint though we omit the proof of this fact. It should be clear that if M is a smooth manifold without boundary then $\partial M = \emptyset$. For an in depth discussion on smooth manifolds with boundary, we refer the reader to Lee's Smooth Manifolds. We now state, without proof, Stoke's theorem on manifolds:

Theorem 1.1.7. *If M is an orientable smooth manifold with boundary, and ω is an $n-1$ form with compact support on M, then:*

$$\int_M d\omega = \int_{\partial M}\omega$$

One should check that ∂M inherits an orientation from M if M is orientable, however we omit this proof, largely because outside of **Example 1.1.16**, we will only apply Stoke's theorem to manifolds with empty boundary. In fact, will not use Stokes theorem again until chapter 3, where we exploit the fact that if M is a smooth manifold without boundary, and $\omega\in\Omega^{n-1}(M)$ then:

$$\int_M d\omega = \int_\emptyset \omega = 0$$

We also note that the major integration theorems in vector calculus: divergence theorem , Stoke's theorem, and Green's theorem, can all be shown to be consequence of Stoke's theorem on manifolds, via the isomorphisms given in **Example 1.1.14 f**. We end with the following example:

Example 1.1.16. Let M be the closed unit disc in $\mathbb{R}^2$, and let ω be the two form defined by:

$$\omega = dx\wedge dy = -d(ydx) = d(\eta)$$

Then by Stoke's theorem on manifolds:

$$\begin{aligned}\int_M \omega &= \int_{\partial M}\eta\\ &= \int_{\mathbb{S}^1}\eta\end{aligned}$$

Applying **Theorem 1.1.6** with the orientation preserving map $\phi^{-1}(\theta) = (\cos\theta,\sin\theta)$ we have that:

$$\begin{aligned}\int_M \omega =& \int_0^{2\pi}\sin^2\theta d\theta\\ =&\pi\end{aligned}$$

1.1.4 The Lie Derivative

Suppose M is a smooth manifold, and I an interval, i.e. a connected open subset of $\mathbb{R}$. A smooth map, $\gamma : J\to M$, is called a *smooth curve* on M. In a coordinate chart we write:

$$\gamma(t) = (x^1(t),\ldots,x^n(t))$$

where $x^1, \ldots, x^n$ are coordinates on M. The velocity vector at $t = t_0$, $\gamma'(t_0)$, defined by:

$$D_{t_0}\gamma(1) = (\dot{x}^1(t_0), \ldots, \dot{x}^n(t_0))$$

is an element of $T_{\gamma(t_0)}M$.

Example 1.1.17. Let $M = \mathbb{R}^2$, and $\gamma : (0, 2\pi) \to \mathbb{R}^2$ be the curve defined:

$$\gamma(t) = (\cos t, \sin t)$$

Then we have that:

$$\gamma'(t) = (-\sin t, \cos t)$$

and the velocity vector at $t = \pi$ is given by:

$$\gamma'(\pi) = (0, -1)$$

Example 1.1.18. Let $M = \mathbb{S}^2$, and $\gamma : (0, \pi) \to \mathbb{S}^2$ be the curve given in the (θ, ϕ) coordinates by:

$$\gamma(t) = (\theta(t), \phi(t)) = (2t, t^2)$$

Then:

$$\gamma'(t) = (2, 2t)$$

and the velocity vector at $t_0 = \pi/4$ is given by:

$$\gamma'(t_0) = (2, \pi/2) = 2\frac{\partial}{\partial\theta}\bigg|_{(\pi/2, \pi^2/16)} + \frac{\pi}{2}\frac{\partial}{\partial\phi}\bigg|_{(\pi/2, \pi^2/16)}$$

From **Example 1.1.7**, in $\mathbb{R}^3$ this vector is in the tangent space at the point:

$$p = \left(\cos\frac{\pi^2}{16}, \sin\frac{\pi^2}{16}, 0\right)$$

which lies on the equator of $\mathbb{S}^2$. Furthermore, this vector is given by:

$$\gamma'(t_0) = -\frac{\pi}{2}\sin\frac{\pi^2}{16}\frac{\partial}{\partial x}\bigg|_p + \frac{\pi}{2}\cos\frac{\pi^2}{16}\frac{\partial}{\partial y}\bigg|_p - 2\frac{\partial}{\partial z}\bigg|_p$$

Using smooth curves we can calculate the differential of a smooth map in an extraordinarily convenient way. Let M and N be smooth manifolds, and F a smooth map between them. Furthermore, let $\gamma(t)$ be a smooth curve on M, starting at p, with velocity vector v at p. Then the differential of F is given by:

$$D_pF(\gamma'(0)) = D_0(F \circ \gamma)(1) = \lim_{t\to 0}\frac{d}{dt}F(\gamma(t)) \tag{1.1.22}$$

We will employ (1.1.22) when we discuss finding the Lie algebra to a Lie group.

Definition 1.1.27. If X is a vector field on M, an **integral curve of X** is a smooth curve $\gamma : I \to M$ such that for all points $t \in I$ we have:

$$\gamma'(t) = X_{\gamma(t)}$$

If $0 \in I$, then $\gamma(0)$ is the **starting point of $\boldsymbol{\gamma}$**.

Example 1.1.19. Let $M = \mathbb{R}^2$, and X be the vector field given in standard coordinates by:

$$X = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$$

Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be the curve given by:

$$\gamma(t) = (x(t), y(t))$$

then for γ to be an integral curve of X, we necessitate that:

$$\begin{aligned} x'(t) =& x(t) \\ y'(t) =& y(t) \end{aligned}$$

which has solutions:

$$\begin{aligned} x(t) =& ae^t \\ y(t) =& ce^t \end{aligned}$$

for $a, b, c, d \in \mathbb{R}$.

In principal, finding integral curves for some vector field X is equivalent to solving a system of ordinary differential equations in some smooth chart. Using the existence, uniqueness, and smoothness theorems of ODE's one can prove a great variety of properties regarding integral curves, such as:

Proposition 1.1.14. *Let V be a smooth vector on a smooth manifold M. For each point $p \in M$, there exist $\epsilon > 0$ and a smooth curve $\gamma(-\epsilon, \epsilon) \to M$ that is an integral curve of V starting at p*

Furthermore we have the following lemma:

Lemma 1.1.4. *Let V be a smooth vector field on a smooth manifold M, let $I \subset \mathbb{R}$ be an interval and $\gamma : I \to M$ be an integral curve of V. For any $s \in \mathbb{R}$, the curve $\gamma_s : I_s \to M$ defined by $\gamma_s(t) = \gamma(t+s)$ is also an integral curve of V, where $I_s = \{t : t+s \in I\}$*

Proof. Let $f \in C^\infty(M)$, and $t_0 \in I_s$, we then have that:

$$\begin{aligned} \gamma_s'(t_0)f =& \lim_{t \to t_0} \frac{d}{dt}(f \circ \gamma_s)(t) \\ =& \lim_{t \to t_0} \frac{d}{dt}(f \circ \gamma)(t+s) \\ =& (f \circ \gamma)'(t_0 + s) \\ =& \gamma'(t_0 + s)f \\ =& V_{\gamma_s(t_0)} f \end{aligned}$$

hence γ_s is an integral curve as desired. □

With **Lemma 1.1.4** in mind, we move to look at the family of integral curves of a vector field in a different manner. First, for a smooth manifold M, and a vector field $X \in \mathfrak{X}(M)$, assume that for each $p \in M$, X has a unique integral curve starting at p, defined for all $t \in \mathbb{R}$. We denote this integral curve by $\theta^{(p)} : \mathbb{R} \to M$. Furthermore, for each t in $\mathbb{R}$ we can define the map $\theta_t : M \to M$, by:

$$\theta_t(p) = \theta^{(p)}(t)$$

By **Lemma 1.1.4**, we have that $t \mapsto \theta^{(p)}(t+s)$ is also an integral curve of X, starting at $q = \theta^{(p)}(s)$, thus by the uniqueness assumption we also have $\theta^{(q)}(t) = \theta^{(p)}(t+s)$, hence:

$$\theta_t \circ \theta_s(p) = \theta_{t+s}(p)$$

Finally, we have that $\theta_0(p) = \theta^{(p)}(0)$. We then define

Definition 1.1.28. Let M be a smooth manifold, then a **smooth global flow** on M is a smooth map $\theta : \mathbb{R} \times M$ satisfying:

a) $\theta(0, p) = p$ for all $p \in M$

b) $\theta(t, \theta(s, p)) = \theta(t+s, p)$ for all $s, t \in \mathbb{R}$ and $p \in M$.

A **smooth global flow** is also an example of a smooth group action on M by the additive group $\mathbb{R}$.

For a smooth global flow θ, it can be shown that the vector field $X \in \mathfrak{X}$ defined pointwise by:

$$X_p = \lim_{t\to 0} \frac{d}{dt}\theta(t,p) := \theta^{(p)'}(0)$$

is indeed smooth a vector field, where the curve $\theta^{(p)}(t)$ is an integral curve of X. We call such a vector field the *infinitesimal generator of θ*. Via the assignment above, and some faith that X is smooth, it is clear that smooth global flow has a smooth vector field as an infinitesimal generator, however the converse need not be true; that is not every smooth vector field need be the infinitesimal generator of a smooth global flow. To somewhat correct this sad fact of life, we construct a local analogue to a smooth global flow in the following way:

Definition 1.1.29. Let $\mathcal{D}$ be an open subset of $\mathbb{R} \times M$, such that for each $p \in M$ we have that:

$$\mathcal{D}^{(p)} = \{t \in \mathbb{R} : (t,p) \in \mathcal{D}\}$$

is an open interval in $\mathbb{R}$. A **smooth flow** is then a smooth map $\theta : \mathcal{D} \to M$ that satisfies:

a) $\theta(0,p) = p$ for all $p \in M$

b) $\theta(t, \theta(s,p)) = \theta(t+s,p)$

At times we refer to this a **smooth local flow**. We also define the set M_t to be:

$$M_t = \{p \in M : (t,p) \in \mathcal{D}\}$$

It can also be shown that every smooth local flow θ has a smooth vector field X as it's infinitesimal generator. We say a flow, or integral curve, is *maximal* if it can not be extended to any larger open set containing $\mathcal{D}$, or any larger open interval respectively. We state the following theorem without proof:

Theorem 1.1.8. *Let V be a smooth vector field on a smooth manifold M. There is a unique smooth maximal flow $\theta : \mathcal{D} \to M$ whose infinitesimal generator is V. This flow has the following properties:*

a) For each $p \in M$, the curve $\theta^{(p)} : \mathcal{D}^{(p)} \to M$ is the unique maximal integral curve of V starting at p.

b) If $s \in \mathcal{D}^{(p)}$, then $\mathcal{D}^{(\theta(s,p))}$ is the interval $\mathcal{D}^{(p)} - s = \{t - s : t \in \mathcal{D}^{(p)}\}$

c) For each $t \in R$, the set M_t is open in M, and $\theta_t : M_t \to M_{-t}$ is a diffeomorphism with inverse θ_{-t}

Example 1.1.20. Following **Example 1.1.19** we have, we have that the smooth global flow of X is given by:

$$\theta(t,x,y) = (xe^t, ye^t)$$

which is a globally defined smooth flow.

Example 1.1.21. Let $M = \mathbb{S}^2$, and let U be the open subset of $\mathbb{S}^2$ covered by the chart given by the previously used angle parameterization. Let X be the smooth local vector field defined in coordinates by:

$$X = \frac{\partial}{\partial \theta}$$

We can find the smooth flow of X by finding the family of integral curves of X near a point p. Let γ be a smooth curve from $I \to \mathbb{S}^2$, then for it to be an integral it must satisfy the following ODE's in coordinates:

$$\theta'(t) = 1$$
$$\phi'(t) = 0$$

Thus we have that:

$$\gamma(t) = (t + a, b)$$

where a, b is the starting point of the curve. We then have that the local flow is given by:

$$\theta(t,\theta,\phi) = (\theta + t, \phi)$$

which is a family of integral curves that are arcs of a great circle going through the north pole, though the flow does not extend to the north or south pole. We also see that:

$$D^{(\theta,\phi)} = (0-\theta, \pi-\theta)$$

and that:

$$M_t = (0,\pi) \times (0, 2\pi)$$

which is the open subset of the two sphere, in the angle coordinates, covered by the chart.

Using the guarantee that every vector field has a unique flow in some open subset $\mathcal{D} \subset \mathbb{R} \times M$, we can push vector fields around an open neighborhood of some point $p \in M$ using the differential of the flow. For example, let θ be the flow of some vector field X, and Y be an other vector field. At the point $p \in M$, and for some $t \in D^{(p)}$, we push W forward to $\theta_t(p)$ via:

$$D_p(\theta_t)\,(X_p) \tag{1.1.23}$$

If (1.1.23) is equal to $X_{\theta_t(p)}$ for general $(t,p) \in \mathcal{D}$, then we say that Y is invariant under the flow of X, however, this is not generally the case. Using (1.1.23) we can develop a way of differentiating vector fields along one another:

Definition 1.1.30. Let $X, Y \in \mathfrak{X}(M)$, and $\theta : \mathcal{D} \to M$ be the flow of X. Then the **Lie Derivative** of Y along X at the point p, denoted $(\mathscr{L}_X Y)_p$, is given by:

$$(\mathscr{L}_X Y)_p = \lim_{t\to 0} \frac{d}{dt} D_{\theta_t(p)}\theta_{-t}(Y_{\theta_t(p)})$$

If $\mathscr{L}_X Y$ is identically zero on M_0, we say that Y is invariant under the flow of X.

Let us quickly check that our two notions of commuting flows are equivalent. If (1.1.23) is equal to $Y_{\theta_t(p)}$ for all $(t,p) \in \mathcal{D}$, then:

$$\begin{aligned}
(L_X Y)_p =& \lim_{t\to 0} \frac{d}{dt} D_{\theta_t(p)}\theta_{-t}(D_p\theta_t(Y_p)) \\
=& \lim_{t\to 0} \frac{d}{dt} D_p(\theta_{-t}\circ\theta_t)(Y_p) \\
=& \lim_{t\to 0} \frac{d}{dt} D_p \mathrm{Id}_M(Y_p) \\
=& 0
\end{aligned}$$

hence Y is invariant under the flow of X according to our definition. Further, let $f(t) : \mathcal{D}^p \to T_pM$ be the function defined by:

$$f(t) = D_{\theta_t(p)}\theta_{-t}(Y_{\theta_t(p)})$$

then, since D_p is independent of the time derivative we have that:

$$\begin{aligned}
f'(t_0) =& \lim_{s\to 0} \frac{d}{ds} f(t_0+s) \\
=& \lim_{s\to 0} \frac{d}{ds} D_{\theta_{-t_0-s}}(Y_{\theta_{s+t_0}(p)}) \\
=& \lim_{s\to 0} D_{\theta_{-t_0(p)}} \frac{d}{ds} D_{\theta_{-s}(p)}(Y_{\theta_s(\theta_{t_0}(p))}) \\
=& D_{\theta_{-t_0(p)}}(\mathscr{L}_X Y)_{\theta_{t_0}(p)}
\end{aligned}$$

Now, assuming that $\mathscr{L}_X Y$ is identically zero on M_0, we see that for all $t \in \mathcal{D}^p$:

$$f'(t) = 0$$

and since $f(0) = Y_p$, we have that $f(t) = Y_p$, hence:

$$D_{\theta_t(p)}\theta_{-t}(Y_{\theta_t(p)}) = Y_p$$

Applying the inverse of $D_{\theta_t(p)}\theta_{-t}$ to both sides gives:

$$Y_{\theta_t(p)} = D_p\theta_t(Y_p)$$

thus our definition agrees with our original discussion.

Theorem 1.1.9. *If M is a smooth manifold and $X, Y \in \mathfrak{X}(M)$, then $\mathscr{L}_X Y = [X, Y]$*

Proof. Let $S(X)$ denote the support of X. We proceed by cases; first, if we show that $(\mathscr{L}_X Y)_p = [X, Y]_p$ for all of p in the interior of $S(X)$, then the statement on the entirety of $S(X)$ follows by continuity. Let θ be the flow of X, and $p \in \mathrm{int}(S(X))$, then for $f \in C^\infty(M)$, we have:

$$\lim_{t\to 0} \frac{d}{dt} f(\theta_t(p)) = X_p f \tag{1.1.24}$$

Furthermore, from the definition of a tangent vector applied to a function we have:

$$\begin{aligned} \theta_{-t*}(Y)_{\theta_t(p)}(f \circ \theta_t)(p) =& D_p(f \circ \theta_t)(\theta_{-t*}Y_{\theta_t(p)}) \\ =& D_{\theta_t(p)}(f \circ \theta_t \circ \theta_{-t})(Y_{\theta_p}) \\ =& D_{\theta_t(p)}(f)(Y_{\theta_t(p)}) \\ =& (Yf)(\theta_t(p)) \end{aligned}$$

Differentiating both sides at $t = 0$, we see that from (1.1.24) the RHS is:

$$\lim_{t\to 0} \frac{d}{dt}(Yf)(\theta_t(p)) = (X(Yf))_p$$

and that from the product rule, the LHS is:

$$\begin{aligned} \lim_{t\to 0} \frac{d}{dt}\left(\theta_{-t*}(Y)_{\theta_t(p)}(f \circ \theta_t)(p)\right) =& \lim_{t\to 0}\left[\frac{d}{dt}\left(\theta_{-t*}(Y)_{\theta_t(p)}\right)(f \circ \theta_t)(p) + \theta_{-t*}(Y)_{\theta_t(p)}\frac{d}{dt}(f \circ \theta_t)(p)\right] \\ =& ((\mathscr{L}_X Y)f)_p + (Y(Xf))_p \end{aligned}$$

hence:

$$(\mathscr{L}_X Y f)_p = (X(Yf))_p - (Y(Xf))_p = [X, Y]_p f$$

Thus on the interior of $S(X)$ we have that the Lie derivative and the Lie bracket are equivalent, and by continuity it must hold on the boundary as well, therefore for all $p \in S(X)$, the two are equivalent. Suppose now that $p \notin S(X)$, then $X = 0$ on a neighborhood of p, and the flow is just the identity map for all t, hence the Lie derivative is zero. Furthermore, if $X_p = 0$ then the Lie bracket is also clearly zero from our earlier discussion on vector fields. Hence for $p \notin S(X)$ we have that the Lie bracket and the Lie derivative are continuous. Combining the two cases, we see that for all $p \in M$:

$$(\mathscr{L}_X Y)_p = [X, Y]_p$$

as desired. □

From **Theorem 1.1.9**, it is clear that our earlier discussion of of vector fields being invariant under the flow of another vector field leads to the following corollary:

Corollary 1.1.4. *For a smooth manifold M, if $X, Y \in \mathfrak{X}(M)$, then the following are equivalent:*

- $[X, Y] = 0$*, i.e. X and Y are commuting vector fields.*
- *X is invariant under the flow of Y.*
- *Y is invariant under the flow of X.*

We now let T^kM be the disjoint union:

$$T^kM = \coprod_{p\in M} T_p^{(0,k)}M$$

then T^kM has the natural topology and structure of a smooth manifold, and much like $\Lambda^K(M)$ is a vector bundle over M. Smooth sections of this bundle are then smooth *tensor fields* on M of type $(0,k)$, or covariant tensor fields. We end our discussion on the Lie derivative by generalizing the Lie derivative to covariant tensor fields in the following way:

Definition 1.1.31. Let M be a smooth manifold, then for $X \in \mathfrak{X}(M)$ with flow θ, and a smooth covariant k tensor field A, the Lie derivative of A is given by:

$$(\mathscr{L}_X A)_p = \lim_{t\to 0} \frac{d}{dt}(\theta_t^* A)_p$$

where for vector fields $Y_1, \dots, Y_k \in \mathfrak{X}(M)$:

$$\theta_t^*(A_{\theta_t(p)})(Y_{1p}, \dots, Y_{kp}) = A_{\theta_t(p)}(D_p\theta_t(Y_1), \dots, D_p\theta_t Y_k)$$

If we define the pullback of a vector field by Y by θ_t as the pushforward by θ_{-t}, we quickly see that this definition is analogous to the previous one for vector fields. Furthermore, regarding $f \in C^\infty(M)$ as a covariant tensor with $k = 0$, we see that:

$$\begin{aligned}\mathscr{L}_X(f)_p =& \lim_{t\to 0} t\frac{d}{dt}\theta_t^*(f)\\ =& \lim_{t\to 0}\frac{d}{dt} f(\theta_t(p))\\ =& X_p f\end{aligned}$$

This calculation leads us to our final proposition:

Proposition 1.1.15. *Let M be a smooth manifold and $X \in \mathfrak{X}(M)$. Suppose $f \in C^\infty(M)$, and A and B are smooth covariant k and l tensor fields respectively. Then the following are true:*

a) $\mathscr{L}_X(fA) = \mathscr{L}_X(f)A + f\mathscr{L}_X(A)$

b) $\mathscr{L}_X(A\otimes B) = (\mathscr{L}_X A)\otimes B + A\otimes(\mathscr{L}_X B)$

c) If $Y_1, \dots, Y_k$ are smooth vector fields then:

$$\mathscr{L}_X(A)(Y^1, \dots, Y^k) = \mathscr{L}_X(A(Y_1, \dots Y_k)) - A(\mathscr{L}_X Y_1, \dots, Y_k) - \dots - A(Y_1, \dots, \mathscr{L}_X Y_k)$$

Proof. We assume p is in the interior of $S(X)$, then by continuity the following hold for all $p \in S(X)$, and for $p \notin S(X)$ our previous argument in **Theroem 1.1.8** holds as well. For a) we see that:

$$\begin{aligned}\mathscr{L}_X(fA) =& \lim_{t\to 0}\frac{d}{dt}\theta_t^*(fA_{\theta_t(p)})\\ =& \lim_{t\to 0}\frac{d}{dt}((f\circ\theta_t)_p \cdot \theta_t^* A_{\theta_t(p)})\end{aligned}$$

which by product rule becomes:

$$\begin{aligned}\mathscr{L}_X(fA)_p =& \lim_{t\to 0}\left[\frac{d}{dt}(f\circ\theta_t)_p\theta_t^*A_{\theta_t(p)} + (f\circ\theta_t)_p\frac{d}{dt}(\theta_t^*A_{\theta_t(p)})\right]\\ =& \mathscr{L}_X(f)A_p + f\mathscr{L}_X(A)_p\end{aligned}$$

We also have that b) follows by the limit definition of the derivative:

$$\begin{aligned}\mathscr{L}_X(A\otimes B)_p =& \lim_{t\to 0}\frac{d}{dt}\left(\theta_t^*A_{\theta_t(p)}\otimes\theta_t^*B_{\theta_t(p)}\right)\\ =& \lim_{t\to 0}\frac{\theta_t^*A_{\theta_t(p)}\otimes\theta_t^*B_{\theta(p)} - A_p\otimes B_p}{t}\\ =& \lim_{t\to 0}\frac{\theta_t^*A_{\theta_t(p)}\otimes\theta_t^*B_{\theta(p)} - \theta_t^*A_{\theta_t(p)}\otimes B + \theta_t^*A_{\theta_t(p)}\otimes B_p - A_p\otimes B_p}{t}\\ =& \lim_{t\to 0}\left[(\theta_t^*A_{\theta_t(p)})\otimes\left(\frac{\theta *_t B_{\theta_t(p)} - B_p}{t}\right) + \left(\frac{\theta_t^*(A_{\theta_t(p)}) - A_p}{t}\right)\otimes B_p\right]\\ =& (A\otimes\mathscr{L}_X B)_p + (\mathscr{L}_X A\otimes B)_p\end{aligned}$$

as desired. For c) we define the function $C : T^{(k,k)}M \to \mathbb{R}$ for $A \in T^kM$ and $Y_1, \ldots, Y_k \in \mathfrak{X}(M)$ pointwise by:

$$[C(A \otimes Y_1 \otimes \cdots \otimes Y_k)]_p = A(Y_1, \ldots, Y_k)_p$$

Furthermore, with the pull back of $Y_j \in \mathfrak{X}(M)$ by θ_t defined by the pushforward as θ_{-t} we have that:

$$[\theta_t^* C(A \otimes Y_1 \otimes \cdots \otimes Y_k)]_p = C(\theta_t^* A_{\theta_t(p)} \otimes \theta_{-t*} Y_{1\theta_t(p)} \otimes \cdots \otimes \theta_{-t*} Y_{k\theta_t(p)})$$

By the linearity of the map C at the point p we have that:

$$\begin{aligned} \frac{d}{dt} C(\theta_t^* A_{\theta_t(p)} \otimes \theta_{-t*} Y_{1\theta_t(p)} \otimes \cdots \otimes \theta_{-t*} Y_{k\theta_t(p)}) &= \\ C\left(\frac{d}{dt}\left[\theta_t^* A_{\theta_t(p)} \otimes \theta_{-t*} Y_{1\theta_t(p)} \otimes \cdots \otimes \theta_{-t*} Y_{k\theta_t(p)}\right]\right) & \end{aligned} \tag{1.1.25}$$

Taking the limit as t goes to 0, we obtain:

$$\lim_{t\to 0} \frac{d}{dt} [\theta_t^* C(A \otimes Y_1 \otimes \cdots \otimes Y_k)]_p = \mathscr{L}_X(A(Y_1, \ldots, Y_k))$$

hence, from applying a product rule like b)[5] to (1.1.25) we obtain:

$$\mathscr{L}_X(A(Y_1, \ldots, Y_k)) = \mathscr{L}_X(A)(Y_1, \ldots, Y_k) + \sum_{i=1}^{k} A(Y_1, \ldots, \mathscr{L}_X(Y_i), \ldots, \mathscr{L}_X(Y_k))$$

therefore:

$$\mathscr{L}_X(A)(Y^1, \ldots, Y^k) = \mathscr{L}_X(A(Y_1, \ldots Y_k)) - A(\mathscr{L}_X Y_1, \ldots, Y_k) - \cdots - A(Y_1, \ldots, \mathscr{L}_X Y_k)$$

as desired. $\square$

Before moving onwards we list two formulas, which can be easily verified in a coordinate chart, for the the exterior derivative of a one form ω and two form β:

$$d\omega(X, Y) = \mathscr{L}_X(\omega(Y)) - \mathscr{L}_Y(\omega(X)) - \omega(\mathscr{L}_X Y) \tag{1.1.26}$$

$$\begin{aligned} d\beta(X, Y, Z) =& \mathscr{L}_X(\beta(Y, Z)) + \mathscr{L}_Y(\beta(Z, X)) + \mathscr{L}_Z(\beta(X, Y)) \\ &- \beta(\mathscr{L}_X Y, Z) - \beta(\mathscr{L}_Y Z, X) - \beta(\mathscr{L}_Z X, Y) \end{aligned} \tag{1.1.27}$$

The components of the aforementioned forms can be found in coordinates by replacing the vector fields X, Y, Z with coordinate vector fields.

1.1.5 (Pseudo)-Riemannian Metrics

A Riemannian metric is a smoothly varying inner product on each tangent space of a smooth manifold M. More precisely:

Definition 1.1.32. Let M be a smooth manifold, and g be a global smooth section of $T^2M = T^*M \otimes T^*M$. Then if g is symmetric, nondegenerate, and positive definite, we call g a **Riemannian metric** . Any smooth manifold M, with a Riemannian metric g, written as the ordered pair (M, g), is a called a **Riemannian manifold.**

By nondegenerate, we mean that for all $p \in M$ and for all $v \in T_pM$, there exists a $w \in T_pM$ such that $g_p(v, w) \neq 0$. An object of interest, particularly in General Relativity, is a mild generalization of the Riemannian metric:

Definition 1.1.33. Let M be a smooth manifold, and g be a global smooth section of T^2M. Then if g is symmetric, and nondegenerate, we call g a **Pseudo-Riemannian metric** . Any smooth manifold M, with a pseudo-Riemannian metric g, written as the ordered pair (M, g) is a called a **Pseudo Riemannian manifold**.

[5] The proof is essentially the same.

Let M be n dimensional. Choosing a basis for T_pM, we can make an orthonormal basis $\{e_i\}_p$ with respect to g_p. In the Riemannian case, it is clear that:

$$g(e_i, e_j) = \delta_{ij}$$

However, in the pseudo-Riemannian case this no longer holds, as g_p is not positive definite. Instead, we have that the orthonormal basis splits into $(e_1, \ldots, e_t)$ and (e_{t+1}, e_{s+t}), where $(s+t) = n$, such that:

$$\begin{aligned} g_p(e_i, e_i) =& -1 \quad \forall 0 \leq i \leq t \\ g_p(e_j, e_j) =& 1 \qquad \forall t+1 \leq j \leq s+t \end{aligned}$$

The *signature* of g_p is the ordered pair (t, s), and determines how many basis vectors have negative 'magnitude', and how many have positive 'magnitude'. More concretely, the signature of g_p determines the maximum dimension of a positive, or negative definite subspace. We now check that this is well defined for any vector space with a symmetric non-degenerate, bilinear form.

Proposition 1.1.16. *Let V be a $\mathbb{R}$-linear vector space, and η a symmetric, non-degenerate bilinear form on V. Then η has a well defined signature (t, s).*

Proof. Let $\{e_i\}$ be the standard basis, and denote it's dual basis by $\{e^i\}$ for V. We represent each e_i by the column vector:

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

Define the components η_{ij} by:

$$\eta_{ij} = \eta(e_i, e_j)$$

so that η can be written as:

$$\eta = \eta_{ij} e^i \otimes e^j$$

Clearly, as η is symmetric:

$$\eta_{ij} = \eta_{ji}$$

We see that for $v = v^i e_i$, and $w = w^j e_j$:

$$\eta(v, w) = \eta_{ij} v^i w^j$$

The matrix:

$$A = \begin{pmatrix} \eta_{11} & \cdots & \eta_{1n} \\ \vdots & \ddots & \vdots \\ \eta_{n1} & \cdots & \eta_{nn} \end{pmatrix}$$

is then symmetric, and satisfies:

$$\eta(v, w) = v^T A w$$

Since A is symmetric, and η is non-degenerate, A has an eigenbasis $\{u^i\}$, such that each u_i has non zero eigenvalue $\lambda_i \in \mathbb{R}$. In this eigenbasis, A is diagonal, i.e.:

$$A^u = \begin{pmatrix} \lambda_i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

We order this basis in such away that the first t elements down the diagonal are negative, and the last s elements are positive, where $s+t=n$. Furthermore, since A is symmetric, we have that:

$$A = P^T A^u P$$

where P is the linear map taking $e_i \mapsto u_i$, i.e. P writes u_i in the e_i basis. Then:

$$\begin{aligned}\eta(u_i, u_j) =&(Pe_i)^T A(Pe_j)\\ =&e_i^T(P^T AP)e_j\\ =&e_i^T A^u e_j\\ =&\lambda_j \delta_{ij}\end{aligned}$$

which is zero if $i \neq j$, hence the eigenbasis is orthogonal. Furthermore, we see that the new basis:

$$f_i = \frac{1}{\sqrt{|\lambda_i|}} u_i$$

is orthonormal, which in this case means $\eta(f_i, f_j) = \pm\delta_{ij}$. We see that η restricted to the subspace spanned by $\{f_1, \cdots, f_t\}$ is negative definite. Let W be any other negative definite subspace, and P be the positive definite subspace spanned by $\{f_{t+1}, \cdots, f_s\}$. Then $P \cap W$ is the zero subspace, hence $\dim W + s \leq n$, implying that $\dim W \leq t$. Similarly, if N is the negative definite subspace spanned by $\{f_1, \ldots, f_t\}$, and W is any positive definite subspace of V, then $W \cap N$ is the zero subspace, hence $\dim W + t \leq n$, implying that $\dim W \leq s$. This implies the claim. □

For reasons that will be clear in the **General Relativity** chapter, we characterize the following classes of vectors:

Definition 1.1.34. Let (M, g) be a pseudo Riemannian manifold, and $v \in T_pM$, then v is **light like** if:

$$g(v, v) = 0$$

is **space like** if:

$$g(v, v) > 0$$

and is **time like** if:

$$g(v, v) < 0$$

Let (M, g) be a (pseudo)-Riemannian manifold with coordinates x^i, then g can be written in coordinates as:

$$g = g_{ij} dx^i \otimes dx^j$$

where $g_{ij} = g_{ji}$ for all i, j. Let y^l be another coordinate system, and recall that for

$$v = \frac{\partial}{\partial y^k}$$

we have:

$$v = \frac{\partial x^j}{\partial y^k} \frac{\partial}{\partial x_j}$$

Furthermore, let:

$$w = \frac{\partial}{\partial y^l} = \frac{\partial x^i}{\partial y^l} \frac{\partial}{\partial x^i}$$

then:

$$g(v, w) = g_{ij} \frac{\partial x^i}{\partial y^l} \frac{\partial x^j}{\partial y^k}$$

Thus, g takes the following form in the y^l coordinate system:

$$g = g_{ij} \frac{\partial x^i}{\partial y^l} \frac{\partial x^j}{\partial y^k} dy^l \otimes dy^k \tag{1.1.28}$$

We clearly see that (1.1.28) aligns with our coordinate transformation formula for one forms. Furthermore, we see that for:

$$v = v^i \frac{\partial}{\partial x^i} \qquad w = w^j \frac{\partial}{\partial x^j}$$

the inner product is given by:

$$g(v, w) = v^i w^j g_{ij}$$

In the y^i coordinate system we then obtain by the chain rule:

$$\begin{aligned} g(v, w) =& v^i w^j \frac{\partial y^l}{\partial x^i} \frac{\partial y^k}{\partial x^j} \frac{\partial x^m}{\partial y^l} \frac{\partial x^n}{\partial y^k} g_{mn} \\ =& v^i w^j \delta_i^m \delta_j^n g_{mn} \\ =& v^i w^j g_{ij} \end{aligned}$$

so the inner product is independent of the coordinates chosen.

Example 1.1.22. Let $M = \mathbb{R}^n$, we can put the standard Euclidean inner product on $T_x\mathbb{R}^n$ for all $x \in \mathbb{R}^n$. Since $\mathbb{R}^n$ is covered by a single chart with coordinates x^i, we have that the Riemannian metric corresponding to this assignment can be written in coordinates as:

$$g = \delta_{ij} dx^i \otimes dx^j$$

Example 1.1.23. Let $M = \mathbb{S}^2$, and let (ϕ, U) be the chart that encodes the angle coordinates on $\mathbb{S}^2$. Equipping $\mathbb{R}^3$ with the standard Euclidean metric:

$$g = dx \otimes dx + dy \otimes dy + dz \otimes dz$$

we can restrict this metric to $\mathbb{S}^2$ via the pullback of g by ϕ^{-1}. Indeed:

$$\begin{aligned} \phi^{-1*} g =& d(x \circ \phi^{-1}) \otimes d(x \circ \phi^{-1}) + d(y \circ \phi^{-1}) \otimes d(y \circ \phi^{-1}) + d(z \circ \phi^{-1}) \otimes d(z \circ \phi^{-1}) \\ =& d(\sin\theta \cos\phi) \otimes d(\sin\theta\cos\phi) + d(\sin\theta\sin\phi) \otimes d(\sin\theta\sin\phi) + d(\cos\theta) \otimes d(\cos\theta) \\ =& \sin^2\theta d\phi \otimes d\phi + d\theta \otimes d\theta \end{aligned}$$

This is the metric induced on $\mathbb{S}^2$ from the Euclidean metric on $\mathbb{R}^3$; it is often referred to as the round metric, and can be extended globally to $\mathbb{S}^2$, by applying the same process to a set of charts which cover the sphere. The corresponding coordinate representations of the metric will be related on the overlap by (1.1.26).

For brevity, we will at times drop the $\otimes$ notation, and simply write:

$$g = g_{ij} dx^i dx^j$$

If the $g_{ij} = 0$ for all $i \neq j$, then we write:

$$g = g_{ii} (dx^i)^2$$

In this notation, the round metric for $\mathbb{S}^2$ can be written as:

$$g = \sin^2\theta d\phi^2 + d\theta^2$$

Theorem 1.1.10. *Every smooth n dimensional manifold M admits a Riemannian metric.*

Proof. Let $\{U_i, \phi_i\}_{i\in I}$ be a locally finite countable covering of smooth charts for M. Furthermore, let $\{\psi_i\}_{i\in I}$ a partition of unity subordinate to $\{U_i, \phi_i\}_{i\in I}$. We define a metric locally in coordinate of each chart by:

$$h_i = \phi_i^* g$$

where g is the standard Euclidean metric on $\mathbb{R}^n$. Now consider the following global section of T^2M:

$$g_i = \psi_i h_i$$

Then the smooth section of T^2M determined by the sum:

$$g_M = \sum_{i\in I} g_i$$

is also globally defined. Since each g_i is symmetric, positive definite, and nondegenerate by construction, and since for all $p \in M$, $\sum_{i\in I} \psi_i(p) = 1$, we have that g_M is also symmetric, positive definite, and nondegenerate, hence g_M is a Riemannian metric for M. □

It is important to note that there actually exists an uncountable number of Riemannian metrics on M, indeed by replacing g_M with some fg_M for some $f \in C^\infty(M)$, such that $f(p) > 0$ for all $p \in M$, we obtain a new Riemannian metric. Furthermore, it is not the case that every smooth manifold M admits a pseudo-Riemannian metric, as our partition of unity argument will no longer hold since each g_i will not be positive definite; further, there are certain topological restrictions which we will not delve into, but, as an example, $\mathbb{S}^2$ admits no pseudo-Riemannian metric.

Proposition 1.1.17. *Let (M, g) be a smooth (pseudo)-Riemannian manifold. For each $p \in M$, there is a smooth orthonormal frame on a neighborhood of p.*

Proof. Let $p \in M$ be arbitrary, and let the chart (U, ϕ) contain p, with coordinates x^i. There exists a v_1 in T_pM such that:

$$g(v_1, v_1) \neq 0$$

otherwise g would be degenerate. We can write $v \in T_pM$ as:

$$v = v^i \frac{\partial}{\partial x^i}\Big|_p$$

Let f_1^i be a a family of functions such that $f_1^i(p) = v^i$, then we can create a local vector field V_1 given by:

$$V_1 = f_1^i \frac{\partial}{\partial x^i}$$

which, by continuity, satisfies:

$$g(V_1, V_1) \neq 0$$

for a neighborhood around p. Let W be the subspace of T_pM defined by:

$$W = \{u \in V : g(u, v) = 0\}$$

Then there exists a $v_2 \in W$ such that:

$$g(v_2, v_2) \neq 0$$

otherwise V would be degenerate. In a similar fashion, we construct a local vector field V_2 such that at the point p, V_2 equals v_2. Furthermore, by continuity, in an open neighborhood of p we have that:

$$g(V_2, V_2) \neq 0$$

Proceeding inductively $n-2$ more times, we obtain a set of orthogonal vectors fields which span each tangent space on a small enough open neighborhood of p, hence on such a neighborhood of p we have the orthogonal frame $\{V_1, \ldots, V_n\}$. Defining E_i to be:

$$E_i = \frac{1}{\sqrt{|g(V_i, V_i)|}} V_i$$

we obtain an orthonormal frame $\{E_1, \ldots, E_n\}$ on a neighborhood of p, as desired. □

With **Proposition 1.1.17** we can show the following:

Theorem 1.1.11. *Let (M, g) be a connected pseudo Riemannian manifold. Then the signature (t, s) of g_p at T_pM is the same for all $p \in M$.*

Proof. First, we note that since M is connected there is only one connected component, and in particular M is path connected[6]. We now proceed by contradiction, suppose that the signature of g_p is not the same at all $p \in M$, then since M is connected we must have an open $U \subset M$ which contains a point p where the signature of g_p is (t, s), and a point q where the signature is $(t+a, s-a)$, for some $a \in \mathbb{Z}$, such that $-t \leq a \leq s$. Without loss of generality, we take a to be equal to 1. Since g must vary smoothly from point to point, we have that p must be in an open neighborhood of q and vice versa. By **Proposition 1.1.17** there exists an orthonormal in an open neighborhood $V \subset U$ of p, and since p and q can be taken to be arbitrarily close to each other we have that V must contain q as well. Then at p we have that:

$$g_p(E_{t+1}, E_{t+1}) = 1$$

and at q:

$$g_q(E_{t+1}, E_{t+1}) = -1$$

Let γ be a smooth curve such that $\gamma(0) = p$, and $\gamma(1) = q$. Let $f \in C^\infty(V)$ be defined by:

$$f(p) = g_p(E_{t+1}, E_{t+1})$$

Then:

$$(\gamma^* f)_t = f(\gamma(t))$$

is a smooth function on $I \to \mathbb{R}$. However, since $\{E_i\}$ is an orthornormal frame on V, f itself must be constant, therefore we clearly have that $f(\gamma(t))$ is not continuous, and hence not smooth, since there must be some $c \in (0, 1)$ where:

$$\lim_{t \to c^-} f(\gamma(t)) = 1 \neq -1 = \lim_{t \to c^+} f(\gamma(t)$$

Since γ is smooth by construction, and the composition of smooth maps is smooth, we must have that g is not smooth at all $p \in M$, and therefore not a global smooth section of T^2M. This a contradiction, hence the signature of g must be the same at all points in p. □

Clearly if M is disconnected then we could have different signatures on each connected component. In light of this, going forward we assume that either (M, g) is connected, or admits a pseudo Riemannian metric which has consistent signature on each connected component.

Definition 1.1.35. Let (M, g) be a pseudo Riemannian manifold of dimension n. If the signature of g is $(n-1, 1)$ or $(1, n-1)$, then (M, g) is called a **Lorentzian Manifold**.

We define the inverse of a (pseudo)-Riemannian metric g to be a symmetric, nondegenerate, global smooth section of $T^{(2,0)}M$, such that at each $p \in M$:

$$g_p^{-1} \lrcorner g_p = g_p \lrcorner g_p^{-1} = \mathrm{Id}$$

where Id is the identity on T_pM, or T_p^*M. The existence of such an inverse at each p is guaranteed by the nondegeneracy of g. In coordinates x^i, letting:

$$g = g_{ij} dx^i \otimes dx^j$$

[6]This comes fairly simply from the fact that M is locally Euclidean

and:

$$g^{-1} = g^{lk}\frac{\partial}{\partial x^l} \otimes \frac{\partial}{\partial x^k}$$

we see that:

$$\begin{aligned} g^{-1} \lrcorner g =& g^{lk} g_{ij} dx^i(\partial_{x^k}) \otimes dx^j \otimes \partial_{x^l} \\ =& g^{lk} g_{ij} \delta^i_k dx^j \otimes \partial_{x^l} \\ =& g^{li} g_{ij} dx^j \otimes \partial_{x^l} \end{aligned} \tag{1.1.29}$$

hence the component functions of (1.1.29) must satisfy:

$$g^{li} g_{ij} = \delta^l_j$$

We can now easily see in coordinates that the contraction of (1.1.27) with a one form ω, or a vector v, returns the same one form, or the same vector, hence it is the identity map on both T_pM and T_p^*M. Motivated by this, we define a coordinate free isomorphism $\alpha_p : T_pM \to T_p^*M$ as follows:

$$\alpha_p(v) = v \lrcorner g_p = g_p(v, \cdot) \tag{1.1.30}$$

$$\alpha_p^{-1}(\omega) = \omega \lrcorner g_p^{-1} = g_p^{-1}(\omega, \cdot) \tag{1.1.31}$$

where $v\lrcorner$ is another way of denoting contraction with v. Thus α_p takes a vector v to the one form λ that satisfies $\lambda(w) = g_p(v, w)$, for all $w \in T_pM$, and α_p^{-1} takes a one form ω to the vector u that satisfies $\eta(u) = g_p^{-1}(\eta, \omega)$ for all $\eta \in T_p^*m$. Furthermore, we note that α is injective as by the nondegeneracy of g:

$$\alpha_p(v) = 0 \Leftrightarrow g_p(v, w) = 0, \forall w \neq 0 \in T_pM \Leftrightarrow v = 0$$

Finally, from rank-nullity we see that α_p must be an isomorphism $T_pM \to T^*M$.

Proposition 1.1.18. *Let (M, g) be a (pseudo)-Riemannian manifold. The metric g, induces a bundle isomorphism $TM \to T^*M$. This isomorphism is commonly referred to as the musical isomorphism.*

Proof. Note that α_p, as defined in (1.1.30) is a pointwise isomorphism $T_pM \to T_p^*M$, with inverse given by α_p^{-1}, as defined in (1.1.31). Thus we have that:

$$\alpha(p, v) = \alpha_p(v)$$

is a bijection $TM \to T^*M$. All that is left then is to show that this map is smooth. Let $X, Y \in \mathfrak{X}(M)$, then:

$$\alpha(p, X)(Y) = g_p(X, Y) \tag{1.1.32}$$

Since $\alpha(p, X)(Y)$ is linear over $C^\infty(M)$ as a function of Y, and $f(p) = g_p(X, Y)$ is a smooth function on M, we have that $\alpha(p, X)$ must be a smooth section of T^*M, i.e a differential one form. Hence, α takes smooth sections to smooth sections and is thus a smooth map. Therefore, $\alpha(p, v)$ is a smooth bijection with smooth inverse[7], which is $\mathbb{R}$-linear, and thus a bundle isomorphism. □

In coordinates x^i, we see that:

$$\begin{aligned} X \lrcorner g =& g_{ij} dx^i(X^l \partial_{x^l}) \otimes dx^j \\ =& g_{ij} X^i dx^j \end{aligned}$$

and that:

$$\begin{aligned} \omega \lrcorner g^{-1} =& g^{ij} \omega_l dx^l(\partial_{x^i}) \otimes \partial_{x^j} \\ =& g^{ij} \omega_i \partial_{x^j} \end{aligned}$$

[7] This comes from the same argument

Contracting the metric and its inverse with vector fields and one forms is commonly referred to as lowering and raising indices, respectively.

On a Riemannian manifold (M, g) we can define the length of a smooth curve $\gamma : I \to M$ as follows:

$$L(\gamma) = \int_I \sqrt{g\left(\dot{\gamma}(t), \dot{\gamma}(t)\right)} dt \tag{1.1.33}$$

When $M = \mathbb{R}^n$, (1.1.33) agrees with the usual formula for the length of a curve encountered in multivariate calculus. Furthermore, in the (pseudo)-Riemannian case, we can calculate the 'length' of *space like curves*, or *time like curves*[8] by taking the absolute value of $g(\dot{\gamma}, \dot{\gamma})$. Clearly, *light like curves* have zero 'length'.

Theorem 1.1.12. *Let (M, g) be an oriented (pseudo) Riemannian n-manifold, then there exists a unique smooth orientation form $\omega_g \in \Omega^n(M)$, commonly referred to as a* ***volume form*** *that satisfies:*

$$\omega_g(E_1, \dots, E_n) = 1 \tag{1.1.34}$$

for every local oriented orthonormal frame (E_i) for M. In any oriented smooth coordinates, (x^i), the Riemannian volume form is given by:

$$\omega_g = \sqrt{|\det(g)|} dx^1 \wedge \cdots \wedge dx^n$$

where g has components of $g_{ij}d$ in these coordinates.

Proof. We first note that as M is oriented, we have that there exists a nowhere vanishing top form ω on M. Let $(E_1, \dots, E_n)$ be a local oriented orthonormal frame on an open set $U \subset M$, and denote its dual frame by $(E^1, \dots, E^n)$. Then ω is given by:

$$\omega = f E^1 \wedge \cdots \wedge E^n$$

Setting $f = 1$, we have that ω_g is then uniquely determined by:

$$\omega_g = E^1 \wedge \cdots \wedge E^n \tag{1.1.35}$$

since:

$$E^1 \wedge \cdots \wedge E^n \left(E_1, \dots, E_n\right) = 1 = \det(E_1, \dots, E_n)$$

To prove existence, we define ω_g in a neighborhood of each point by (1.1.35); we must show that this definition is independent of choice of oriented orthonormal frame. Let $(F_1, \dots, F_n)$ be another oriented orthonormal frame, with dual frame $(F^1, \dots, F^n)$. In this frame we set:

$$\tilde{\omega}_g = F^1 \wedge \cdots \wedge F^n$$

Let A be the matrix of functions such that:

$$F_i = A_i^j E_j$$

Since A takes an oriented orthonormal basis to another oriented orthonormal basis, we have $\det(A) = 1$.

$$\begin{aligned} \omega_g(F_1, \dots, F_n) &= \det\left(A_1^j E_j, \dots, A_n^j E_j\right) \\ &= \det(A) \det(E_1, \dots, E_n) \\ &= 1 = \tilde{\omega}_g(F_1, \dots, F_n) \end{aligned}$$

Thus $\omega_g = \tilde{\omega}_g$, hence defining ω_g with respect to some smooth oriented orthonormal frame gives a global n form satisfying (1.1.31). Now let B be the smooth matrix of functions such that:

$$\frac{\partial}{\partial x^i} = B_i^j E_j$$

[8] A space like curve is just a smooth curve $\gamma : I \to M$ such that $g(\dot{\gamma}, \dot{\gamma}) > 0$ for all $t \in I$, while a light curve is one such that $g(\dot{\gamma}, \dot{\gamma}) < 0$ for all $t \in I$

for some oriented coordinate frame. Then we see that:

$$\begin{aligned}\omega_g\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^n}\right) &= \det(B)\det\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^n}\right)\\ &= \det(B)dx^1\wedge\cdots\wedge dx^n\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^n}\right)\end{aligned}$$

However,

$$\begin{aligned}g_{ij} = g\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right) =& g(B^k_i E_k, B^l_j E_l)\\ =& B^k_i B^l_j \delta_{kl}\\ =& B^k_i B^k_j\end{aligned} \tag{1.1.36}$$

However, if the signature of g is (t,s), such that $s+t=n$, we know that there is actually a minus sign needed in (1.1.36) for $1\leq k\leq t$, hence (1.1.36) should be written as:

$$g_{ij} = -\sum_{k=1}^{t} B^k_i B^k_j + \sum_{k=t+1}^{s+t} B^k_i B^k_j \tag{1.1.37}$$

This can further be reworked if we introduce a matrix η, with entries:

$$\eta_{ij} = -\delta_{ij}\ \forall 1\leq i,j\leq t \qquad \text{and} \qquad \eta_{ij} = \delta_{ij}\ \forall t+1\leq i,j\leq s+t$$

then (1.1.36) can be rewritten as:

$$\begin{aligned}g_{ij} =& B^l_i \eta_{lk} B^k_j\\ \Rightarrow g =& B\eta B^T\end{aligned}$$

Therefore we have:

$$\det(g) = \det\big(B\eta B^T\big) = \det(B)^2\det(\eta) = \det(B)^2(-1)^t$$

B has positive determinant as it takes an oriented basis to an oriented basis, hence:

$$\det(B) = \sqrt{|\det(g)|}$$

Hence, in these coordinates, we obtain that:

$$\omega_g = \sqrt{|\det(g)|}dx^1\wedge\cdots\wedge dx^n$$

as desired. □

In the Riemannian case, the determinant of g is always positive, since it's signature is $(t,0)$, hence there is no need for the absolute value under the square root. Furthermore, when (M,g) is a Lorentzian manifold, since the the determinant of g is always negative, we write:

$$\omega_g = \sqrt{-\det(g)}$$

We end with three examples:

Example 1.1.24. Let $M=\mathbb{R}^3$, and let $\gamma: I\to\mathbb{R}^3$ be a smooth curve. Further, let g be the standard Euclidean metric:

$$g = \delta_{ij}dx^i\wedge dx^j = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$$

Note that γ can be written as:

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

where each $\gamma^i \in C^\infty(\mathbb{R})$. We can pull this metric back to the curve parameterized by γ in the following way:

$$\begin{aligned} g_\gamma =& \gamma^* \left((dx^1)^2 + (dx^2)^2 + (dx^3)^2\right) \\ =& \left(d(x^1 \circ \gamma)\right)^2 + \left(d(x^2 \circ \gamma)\right)^2 + \left(d(x^3 \circ \gamma)\right)^2 \\ =& \left(\dot\gamma^1\right)^2 dt^2 + \left(\dot\gamma^2\right)^2 dt^2 + \left(\dot\gamma^3\right)^2 dt^2 \end{aligned}$$

The determinant of g_γ is then:

$$\det(g_\gamma) = \left(\dot\gamma^1\right)^2 + \left(\dot\gamma^2\right)^2 + \left(\dot\gamma^3\right)^2$$

Hence the volume form on γ is given in these coordinates by:

$$\omega_{g_\gamma} = \sqrt{(\dot\gamma^1)^2 + (\dot\gamma^2)^2 + (\dot\gamma^3)^2}dt$$

Thus the arc length of this curve over some interval $[a, b] \in I$ is given by:

$$L(\gamma) = \int_a^b \sqrt{(\dot\gamma^1)^2 + (\dot\gamma^2)^2 + (\dot\gamma^3)^2}dt$$

Furthermore, for some $f \in C^\infty(\mathbb{R}^3)$ we obtain the formula for the integral of a scalar function along a curve, encountered in a standard multivariate calculus course:

$$\int_\gamma f dt = \int_I f \sqrt{(\dot\gamma^1)^2 + (\dot\gamma^2)^2 + (\dot\gamma^3)^2}dt$$

Example 1.1.25. Let $M = \mathbb{R}^3 \setminus \mathbf{0}$, again equipped with the standard Euclidean metric. We wish to calculate the metric in spherical coordinates, that is under the identification:

$$\begin{aligned} x =& r \sin\theta \cos\phi \\ y =& r \sin\theta \sin\theta \\ z =& r \cos\theta \end{aligned}$$

for $r \in (0, \infty)$, $\theta \in [0, \pi]$, and $\phi \in [0, 2\pi]$. We calculate term by term:

$$\begin{aligned} (d(r\sin\theta\cos\phi))^2 =& (\sin\theta\cos\phi dr + r\cos\theta\cos\phi d\theta - r\sin\theta\sin\phi d\phi)^2 \\ (d(r\sin\theta\sin\phi))^2 =& (\sin\theta\sin\phi dr + r\cos\theta\sin\phi d\theta + r\sin\theta\cos\phi d\phi)^2 \\ (d(r\cos\theta))^2 =& (\cos\theta dr - r\sin\theta d\theta)^2 \end{aligned}$$

Expanding each term, and adding them together, we find that all cross terms cancel, and we obtain:

$$\begin{aligned} g =& \left(\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta\right) dr^2 + (r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta)d\theta^2 \\ &+ (r^2\sin^2\theta\sin^2\phi + r^2\sin^2\theta\cos^2\phi)d\phi^2 \\ =& dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \end{aligned}$$

hence the volume form in these coordinates is given by:

$$\omega_g = r^2 \sin\theta dr \wedge d\theta \wedge d\phi \tag{1.1.38}$$

We see that (1.1.38) aligns with the volume element derived for integration in spherical coordinates, commonly encountered in a multivariate calculus course, or an upper level physics course.

Example 1.1.26. Let $M = \mathbb{S}^2$, and g be the round metric encountered in **Example 1.1.23**. Then:

$$g = d\theta^2 + \sin^2\theta d\phi^2$$

in the standard angle coordinates, which we have encountered before. Thus, the volume form is given by:

$$\omega_g = \sin\theta d\theta \wedge d\phi$$

We would like to check that this form corresponds to the ω encountered in **Example 1.1.15**. Letting:

$$(x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\theta, \cos\theta)$$

We see that that:

$$\psi(x, y, z) = \left(\arccos\frac{z}{\sqrt{x^2+y^2+z^2}}, \arctan\frac{y}{x}\right)$$

where the first entry is the θ coordinate, and the second is the ϕ coordinate. Pulling ω_g back to $\mathbb{S}^2 \in \mathbb{R}^3$, we calculate the following:

$$\sin(\theta) = \sqrt{1 - \frac{z^2}{x^2+y^2+z^2}} = \sqrt{1-z^2}$$

Furthermore:

$$d\theta = \frac{xz}{\sqrt{x^2+y^2}}dx + \frac{yz}{\sqrt{x^2+y^2}}dy + \frac{-x^2-y^2}{\sqrt{x^2+y^2}}dz$$

and:

$$d\phi = \frac{-y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$$

Note that:

$$x^2+y^2+z^2 = 1 \Rightarrow \sqrt{1-z^2} = \sqrt{x^2+y^2}$$

hence:

$$\sin\theta d\theta = xzdx + yzdy - (x^2+y^2)dz$$

Therefore we see that:

$$\sin\theta d\theta \wedge d\phi = \frac{x^2z}{x^2+y^2}dx \wedge dy - \frac{y^2z}{x^2+y^2}dy \wedge dx + ydz \wedge dx + xdy \wedge dz$$

Grouping the first two terms via the relation $dy \wedge dx = -dx \wedge dy$, we see that:

$$\sin\theta d\theta \wedge d\phi = zdx \wedge dy + ydz \wedge dx + xdy \wedge dz$$

which is exactly the two form ω we encountered in **Example 1.1.15**, thus it is no coincidence that:

$$\int_{\mathbb{S}^2} \omega = 4\pi$$

1.2 Lie Theory and Representation Theory

In this chapter, we briefly delve into the field of Lie theory, a broad and interesting field of mathematics in it's own right. We, however, limit ourselves to the topics necessary for understanding the role that Lie theory plays in gauge theory. We first introduce Lie groups, which are ubiquitous in theoretical physics, and, in our case, can be thought of as encoding smooth symmetries in our field theories, i.e., as mentioned in the introduction, the *action* of these groups on fields leaves the Lagrangian invariant. We then move on to discuss Lie algebras, and the Maurer Cartan form, which will equip us with the tools to study the smooth manifold properties of Lie Groups. Finally, and perhaps most importantly for our work in Gauge theory, we move on to quotient manifolds, and the representation theory of Lie groups and algebras. Much of this section is drawn from the text Hamilton's *Mathematical Gauge Theory*, and many important theorems are presented without proof. To the interested reader, we recommend the aforementioned text.

1.2.1 Lie Groups

We first recall the definition of a group:

Definition 1.2.1. A **group** is a set G, endowed with a multiplicative action $\cdot$ such that the following axioms are satisfied:

a) G is closed under multiplication, i.e. $\forall g, h \in G \; g \cdot h \in G$

b) G has an identity element e such that $\forall g \in G \; e \cdot g = g$

c) Every element $g \in G$ has an inverse $g^{-1} \in G$ such that $g^{-1} \cdot g = e$.

d) Multiplication is associative, i.e. $\forall g, h, i \in G$:

$$(g \cdot h) \cdot i = g \cdot (h \cdot i)$$

In general, groups arise naturally by looking for symmetry preserving actions on sets or objects. For example, the dihedral groups arise by finding the rotations and reflections which leave a regular n-gon invariant. If we want to look at *smooth* symmetries however, we must have some notion of smoothness in addition to a group structure, motivating our next definition:

Definition 1.2.2. A **Lie Group** is a group, G, which is also a smooth manifold, such that the map:

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\longmapsto g \cdot h^{-1} \end{aligned}$$

is smooth.

Equivalently, as the next lemma shows, we can check that the multiplication, and inversion are smooth maps independently.

Lemma 1.2.1. *A group G is a a Lie group if and only if it is at the same time a smooth manifold so that both of the maps:*

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\longmapsto g \cdot h \\ G &\longrightarrow G \\ g &\longmapsto g^{-1} \end{aligned}$$

Proof. Suppose that multiplication and inversion are smooth maps, then their composition is a smooth map, as the composition of any two maps is smooth, hence G is a Lie group by **Definition 1.2.2**.

Conversely, suppose that G is a Lie group, then the map:

$$\begin{aligned} G \longrightarrow G \times G &\longrightarrow G \\ g \longmapsto (e, g) &\longmapsto e \cdot g^{-1} = g^{-1} \end{aligned}$$

is smooth, thus inversion is a smooth map. Furthermore, since inversion is smooth, the map:

$$\begin{aligned} G \times G &\longrightarrow G \times G \longrightarrow G \\ (g,h) &\longmapsto (g,h^{-1}) \longmapsto g \cdot \left(h^{-1}\right)^{-1} = g \cdot h \end{aligned}$$

is smooth, hence multiplication is a smooth map. □

We would also like to discuss maps from Lie groups to other Lie groups which preserve the group structure:

Definition 1.2.3. Let G and H be Lie groups, and $\phi : G \to H$ a smooth map between them. Then, ϕ is a Lie group homomorphism if for every $g_1, g_2 \in G$:

$$\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2)$$

With this definition, and the preceding lemma, we obtain the following corollary:

Corollary 1.2.1. *Let G be a Lie group, then the maps $L_g : G \to G$ and $R_g : G \to G$, which denote left and right multiplication by a fixed element $g \in G$, are diffeomorphisms. Furthermore, the inversion map is a diffeomorphism, and a Lie group isomorphism if and only if G is abelian.*

Furthermore, we can construct new Lie groups out of given ones in the following way:

Proposition 1.2.1. *Let G and H be Lie groups. Then the product smooth manifold $G \times H$ with the direct product structure of a group is a Lie group, called the* ***Product Lie group***

Proof. With the product smooth structure note that inversion is smooth on G and H, so the map:

$$\begin{aligned} G \times H &\longrightarrow G \times H \\ (g,h) &\longmapsto (g^{-1}, h^{-1}) \end{aligned}$$

is also smooth. Furthermore, as multiplication is smooth in both G and H, the map:

$$\begin{aligned} (G \times H) \times (G \times H) &\longrightarrow G \times H \\ (g_1, h_1, g_2, h_2) &\longmapsto (g_1 \cdot g_2, h_1 \cdot h_2) \end{aligned}$$

is also smooth by the same argument as before. Thus, by **Lemma 1.2.1**, $G \times H$ is a Lie group, as desired. □

We now turn to our first example of a Lie group:

Example 1.2.1. Let $GL_n(\mathbb{R})$ be the set of linear transformation of $\mathbb{R}^n$ which are invertible, i.e.:

$$GL_n(\mathbb{R}) = \{A \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : \det(A) \neq 0\}$$

This set has a multiplicative action given by ordinary matrix multiplication, and is further closed under said multiplicative action as for $A, B \in GL_n(\mathbb{R})$ we have that:

$$\det(A \cdot B) = \det(A) \cdot \det(B) \neq 0$$

hence $GL_n(\mathbb{R})$ is closed under multiplication. Furthermore, the identity matrix, I, is in $GL_n(\mathbb{R})$ as $\det(I) = 1$, and each A has an inverse in $GL_n(\mathbb{R})$ as:

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)} \neq 0$$

Finally, multiplication is associative since matrix multiplication is associative, so $GL_n(\mathbb{R})$ is indeed a group.

Furthermore, $\mathrm{Mat}_{n\times n}(\mathbb{R})$ is a vector space isomorphic to $\mathbb{R}^{n^2}$ hence a smooth manifold. As the determinant map is a multilinear map, we have that $\det \in C^\infty(\mathrm{Mat}_{n\times n}(\mathbb{R}))$ smooth, and thus continuous. We see that the inverse image, $\det^{-1}(0)$, is a closed set in $\mathrm{Mat}_{n\times n}(\mathbb{R})$, as $\{0\}$ is closed in $\mathbb{R}$. Note that the complement is given by:

$$\det{}^{-1}(0)^c = \{A \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : \det(A) \neq 0\} = GL_n(\mathbb{R})$$

so $GL_n(\mathbb{R})$ is a open subset of $\mathrm{Mat}_{n\times n}(\mathbb{R})$, and hence an open submanifold of $\mathrm{Mat}_{n\times n}(\mathbb{R})$.

We now have that $GL_n(\mathbb{R})$ is both a group and submanifold, so we now need only show that multiplication and inversion are smooth maps. Multiplication by any two elements is a bilinear map, and thus a smooth map $\mathrm{Mat}_{n\times n}(\mathbb{R}) \times \mathrm{Mat}_{n\times n}(\mathbb{R}) \to \mathrm{Mat}_{n\times n}(\mathbb{R})$, so the restriction of the multiplication map to $GL_n(\mathbb{R}) \times GL_n(\mathbb{R})$ must be smooth as well. Furthermore, for $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, if the $\det(A) \neq 0$, the coordinates of A^{-1} are rational functions of the coordinates of A, which are smooth exactly when $\det A \neq 0$, and undefined otherwise. This implies that inversion is smooth on $GL_n(\mathbb{R})$, and hence $GL_n(\mathbb{R})$ is a Lie group by **Definition 1.2.2**.

The example above is the quintessential Lie group, and called the general linear group. In fact, all of the Lie groups we will examine in this paper, save the Spin and Pin groups, will be subgroups of the general linear group over some field, usually $\mathbb{C}$ or $\mathbb{R}$. We will spend the remainder of this subsection defining and examining common Lie groups, but first state the following theorem, known as **Cartan's Closed Subgroup Theorem**:

Theorem 1.2.1. *Let G be a Lie group, and let $H \subset G$ be a subgroup of G. Then H is an embedded Lie subgroup of G if and only if H is closed in the topology of G.*

Example 1.2.2. We examine the set:

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) : \det(A) = 1\}$$

often referred to as the **Special Linear Group**, which is the set of linear transformations which preserve the volume of the parallepiped spanned by n vectors in $\mathbb{R}^n$. This is clearly a subgroup as $I \in SL_n(\mathbb{R})$, $SL_n(\mathbb{R})$ is closed under matrix multiplication, and $SL_n(\mathbb{R})$ contains all it's inverses. Furthermore, this is easily seen to a Lie subgroup of $GL_n(\mathbb{R})$ by **Theorem 1.2.1**, however we would also like to know the dimension of $SL_n(\mathbb{R})$. To do this we note that $SL_n(\mathbb{R})$ is the inverse image the determinant map:

$$\det : \mathrm{Mat}_{n\times n} \to \mathbb{R}$$

at 1, we will show that this is a regular value. Let γ be the curve going through the identity, such that:

$$\gamma(t) = I + tX$$

for some $X \in T_{\gamma(0)}\mathrm{Mat}_{n\times n}(\mathbb{R})$. As $\mathrm{Mat}_{n\times n}(\mathbb{R})$ is the vector space isomorphic to $\mathbb{R}^{n^2}$, we identify it's tangent space with $\mathrm{Mat}_{n\times n}(\mathbb{R})$. Let X have eigenvalues λ_i, which may or may not be complex, and may repeat, then we see that:

$$\begin{aligned} D_I \det(X) &= \lim_{t\to 0} \frac{d}{dt} \det(I + tX) \\ &= \lim_{t\to 0} \frac{d}{dt} \prod_{i=1}^{n} (1 + t\lambda_i) \\ &= \mathrm{Tr}(X) \end{aligned}$$

Note that even though the eigenvalues may be complex, the trace is still real, as it must be equal to the sum:

$$\sum X_{ii}$$

which must be real, as our vector space is over $\mathbb{R}$. Furthermore, if $\gamma(t)$ is a curve in $\mathrm{Mat}_{n\times n}(\mathbb{R})$, such that $\dot\gamma(0) = X$, and $\gamma(0) = I$, we see that:

$$\begin{aligned} D_I \left(\det \circ L_A\right)(X) &= \lim_{t\to 0} \frac{d}{dt} \det(A\gamma(t)) \\ &= \det(A) \lim_{t\to 0} \frac{d}{dt} \det(\gamma(t)) \\ &= \det(A) D_e \det(X) \\ &= \det(A)\, \mathrm{Tr}(X) \end{aligned}$$

Furthermore, for $A \in GL_n(\mathbb{R})$, we have that:

$$\begin{aligned} D_I L_A(X) =& \lim_{t\to 0} \frac{d}{dt} A\gamma(t) \\ =& AX \in T_A \mathrm{Mat}_{n\times n}(\mathbb{R}) \end{aligned}$$

Combining these two facts we see that:

$$\begin{aligned} D_A \det(X) =& D_A \det\left(AA^{-1}X\right) \\ =& D_I \left(\det \circ L_A\right) \circ D_A L_{A^{-1}}(X) \\ =& \det(A)\operatorname{Tr}\left(A^{-1}X\right) \end{aligned} \tag{1.2.1}$$

Suppose $\det(A) = 1$, and $X = xn^{-1}A \in T_A\mathrm{Mat}_{n\times n}(\mathbb{R})$ for any $x \in \mathbb{R}$, then:

$$\begin{aligned} D_A \det(X) =& \operatorname{Tr}\left(xn^{-1}A^{-1}A\right) \\ =& x \end{aligned}$$

Hence for any $A \in SL_n(\mathbb{R})$ we see that det is a smooth submersion, so 1 is a regular value and by **Theorem 1.1.1** $SL_n(\mathbb{R})$ is a closed submanifold of $\mathrm{Mat}_{n\times n}(\mathbb{R})$ of dimension $n^2 - 1$.

Example 1.2.3. Let $O(n)$ be the set:

$$O(n) = \{A \in GL_n(\mathbb{R}) : \forall v, w \in \mathbb{R}^n, \langle Av, Aw\rangle = \langle v, w\rangle\}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product. This is often referred to as the **Orthogonal Group**, and is the set of linear transformations which preserve the Euclidean inner product on $\mathbb{R}^n$. Note that:

$$\langle v, w\rangle = v^T w$$

so:

$$\langle Av, Aw\rangle = \langle v, w\rangle \Rightarrow v^T A^T Aw = v^T w \Rightarrow A^T A = I$$

Let $F : \mathrm{Mat}_{n\times n}(\mathbb{R}) \to Sym_n(\mathbb{R})$[9] be the map:

$$F(A) = A^T A$$

which is clearly continuous. Then $O(n)$ is the inverse image of $\{I\}$, and hence closed in $GL_n(\mathbb{R})$. Furthermore, $O(n)$ is a subgroup as for $A, B \in O(n)$ we have that:

$$(AB)^T(AB) = A^T B^T BA = A^T A = I$$

so $O(n)$ is closed under multiplication. We also have that $O(n)$ contains it's inverses as for $A \in O(n)$, we have $A^{-1} = A^T$, hence:

$$AA^T = (A^T)^T = (I)^T = I$$

Finally, $O(n)$ clearly contains the identity, so $O(n)$ is a closed subgroup of $GL_n(\mathbb{R})$, and thus an embedded Lie subgroup by **Theorem 1.2.1**. To determine the dimension of $O(n)$ we see that I is actually a regular value of F. Let γ be a curve through $GL_n(\mathbb{R})$ such that $\gamma(0) = A \in O(n)$, and $\dot\gamma(0) = X \in T_A GL_n(\mathbb{R}) \cong \mathrm{Mat}_{n\times n}(\mathbb{R})$, then:

$$\begin{aligned} D_A F(X) =& \lim_{t\to 0} \frac{d}{dt}\left(\gamma^T\gamma\right) \\ =& \lim_{t\to 0} \dot\gamma^T\gamma + \gamma^T\dot\gamma \\ =& X^T A + A^T X \end{aligned}$$

[9] $Sym_n(\mathbb{R})$ is the vector space of symmetric matrices, which has dimension $n(n+1)/2$

Then for $B \in Sym_n(\mathbb{R})$ we see that $X = \frac{1}{2}AB$ gives:

$$\begin{aligned} D_A F(X) =& \frac{1}{2}(AB)^T A + \frac{1}{2}A^T AB \\ =& \frac{1}{2}B^T + \frac{1}{2}B \\ =& B \end{aligned}$$

hence $D_A F$ is a surjection for all $A \in F^{-1}(I)$. Therefore, we see that by **Theorem 1.1.1** $O(n)$ is a smooth sub manifold of dimension:

$$\dim O(n) = n^2 - \frac{n^2+n}{2} = \frac{n^2-n}{2}$$

Additionally, we see by the Hein-Borell Theorem that $O(n)$ is compact as for $A \in O(n)$:

$$(A^T A)_{ii} = \sum_{j=1}^{n} (A_{ij})^2 = 1$$

so we have that $|A_{ij}| < 1$ for all $1 \leq i, j \leq n$, hence $O(n)$ is a closed and bounded subset of $\mathrm{Mat}_{n\times n}(\mathbb{R})$, and thus compact.

Example 1.2.4. For $A \in O(n)$ we see that:

$$\det\left(A^T A\right) = \det(A)^2 = 1 \Rightarrow \det(A) = \pm 1$$

so $O(n)$ consists of two connected component, one corresponding to elements of $O(n)$ satisfying $\det(A) = -1$ and another corresponding to elements of $O(n)$ satisfying $\det(A) = 1$. Note that the connected component of $O(n)$ corresponding to $\det(A) = -1$ cannot be a group, as it does not contain the identity, however, we see that the set:

$$SO(n) = \{A \in O(n) : \det(A) = 1\}$$

is clearly a subgroup of $O(n)$. Furthermore, we see that under the map:

$$\det : O(n) \to \{-1, 1\}$$

$SO(n)$ is the inverse image of 1, and hence closed in $O(n)$ as det is continuous. So, by **Theorem 1.2.1**, $SO(n)$ is an embedded Lie subgroup of $O(n)$. In particular, $SO(n)$ is thought of the group of orthogonal transformations which preserve the orientation of $\mathbb{R}^n$, and is often called the **Special Orthogonal Group**. We also see that the complement of $SO(n)$ in $O(n)$ is the other connected component of $O(n)$, which is closed in $O(n)$ as the inverse image of -1 under the determinant map. This implies that $SO(n)$ is also open in $O(n)$, and thus an open submanifold of $O(n)$, so we obtain $\dim SO(n) = \dim O(n)$. Finally, since $SO(n)$ is closed in $O(n)$, and $O(n)$ is compact, we also have that $SO(n)$ is compact.

Example 1.2.5. Recall from our earlier work on pseudo-Riemannian metrics that we can define a symmetric indefinite, non-degenerate bilinear form of signature (t, s) on $\mathbb{R}^n$. For the standard basis $\{e_i\}$ of $\mathbb{R}^n$, we define such a form η as:

$$\begin{aligned} \eta(e_i, e_j) =& -\delta_{ij} \qquad \text{for } 1 \leq i, j \leq t \\ \eta(e_i, e_j) =& \delta_{ij} \qquad \text{for } t+1 \leq i, j \leq s+t \end{aligned}$$

The set:

$$O(t,s) = \{A \in GL_n(\mathbb{R}) : \forall v, w \in \mathbb{R}^n, \eta\,(Av, Aw) = \eta(v,w)\}$$

referred to as the **Pseudo-Orthogonal Group** is then a subgroup of $GL_n(\mathbb{R})$. We see this by first defining a matrix η as:

$$\eta = \begin{pmatrix} -I_t & 0 \\ 0 & I_s \end{pmatrix}$$

where I_t is the identity on the subspace spanned by the first t basis vectors, and I_s is the identity on the subspace spanned by complimentary set of basis vectors complimentary to t. Then, in the standard basis:

$$\eta(v, w) = v^T \eta w$$

so:

$$\eta(Av, Aw) = v^T A^T \eta A w = v^T \eta w \Rightarrow A^T \eta A = \eta$$

The set clearly contains the identity matrix. Further, it is then closed under multiplication, as for $A, B \in O(t, s)$ we have:

$$(AB)^T \eta (AB) = B^T A^T \eta A B = B^T \eta B = \eta$$

Also, $O(t, s)$ contains it's inverses, since $\eta \in O(t, s)$ as:

$$\eta^T \eta \eta = \eta^3 = \eta$$

so for $A \in O(t, s)$:

$$\eta A^T \eta A = \eta^2 = I$$

so $A^{-1} = \eta A^T \eta$. Now examine the map:

$$\begin{aligned} F : GL_n(\mathbb{R}) &\longrightarrow Sym_n(\mathbb{R}) \\ A &\longmapsto A^T \eta A \end{aligned}$$

then $F^{-1}(\eta)$ is closed in $GL_n(\mathbb{R})$, so by **Theorem 1.2.1** $O(t, s)$ is an embedded Lie subgroup of $GL_n(\mathbb{R})$. The differential of this map is:

$$D_A F(X) = X^T \eta A + A^T \eta X$$

For any $B \in SyM_n(\mathbb{R})$, and any $A \in O(t, s)$, we see that for $X = \frac{1}{2} A \eta B \in T_A GL_n(\mathbb{R})$:

$$\begin{aligned} D_A F(X) =& \frac{1}{2} B^T \eta A^T \eta A + \frac{1}{2} A^T \eta A \eta B \\ =& \frac{1}{2} B^T \eta^2 + \frac{1}{2} B \eta^2 \\ =& B \end{aligned}$$

So η is a regular value of F, implying that:

$$\dim O(t, s) = \dim O(n)$$

where $n = s + t$. It is important to note that unless $s = 0$ or $t = 0$, $O(t, s)$ is not compact.

Before moving on to our final example, we must briefly discuss the notion of time orientability. Let $V = \mathbb{R}^n$, and η be a symmetric, indefinite, non-degenerate, bilinear form on V of signature (t, s). We see that V splits into two vector subspaces:

$$V_- = \mathbb{R}^t = \text{span}\{e_1, \ldots, e_t\} \qquad \text{and} \qquad V_+ = \mathbb{R}^s = \text{span}\{e_{t+1}, \cdots, e_{t+s}\}$$

Clearly, η is negative definite on V_- and positive definite on V_+. Let π denote the projection map:

$$\pi : V \to V_-$$

and let $W \subset V$, be any maximally negative definite vector subspace. Then, by rank nullity, we have that the restriction of π to W:

$$\pi|_W : W \to V_-$$

is a vector space isomorphism. If we fix an orientation on V_-, then there exists a unique orientation on W such that $\pi|_W$ is an orientation preserving isomorphism. Furthermore, we see that if $A \in O(t, s)$, the image of the map:

$$A|_W : W \to A(W)$$

is also a maximally negative definite subspace of V, as A restricted to any subspace is an isomorphism of vector subspaces, and A preserves η, so every vector in the image of A is also negative definite. We employ the following definition:

Definition 1.2.4. Let $A \in O(t,s)$, then A has **time orientability** $+1$ if:

$$A|_{V_-} : V_- \to A(V_-)$$

preserves the orientation of V_-, and -1 otherwise.

With this definition at hand, we wish to prove the following lemma:

Lemma 1.2.2. *Let W be an arbitrary maximally negative definite subspace of V, and $A \in O(t,s)$. A has time orientability $+1$ if and only if:*

$$A|_W : W \to A(W)$$

preserves orientation with the orientation on W and $A(W)$ determined by the projection π.

Proof. Given that $\pi|_W$ is an isomorphism $W \to V_-$, we can find a unique basis for W, $\{w_i\}$ such that:

$$w_i = e_i + v_i$$

for $v_i \in V_+$, and $i \in \{1, \dots, t\}$. Since W is a negative definite subspace, for any non zero $w = a^i w_i$, we have:

$$\eta(w,w) = \sum_{i=1}^{t} -a_i^2 + a_i^2 \eta(v_i, v_i) < 0$$

We construct a family of maximally negative definite subspaces, parameterized by $\tau \in [0,1]$, as follows:

$$W_\tau = \{e_1 + \tau v_i, \dots, e_t + \tau v_t\}$$

We see that $W_0 = V_-$, and $W_1 = W$, and that for any $\tau \in [0,1]$, any non zero vector $w_\tau = a^i(e_i + \tau v_i)$ satisfies:

$$\eta(w_\tau, w_\tau) = \sum_{i=1}^{t} -a_i^2 + \tau^2 a_i^2 \eta(v_i, v_i) \leq \sum_{i=1}^{t} -a_i^2 + a_i^2 \eta(v_i, v_i) < 0$$

so W_τ is indeed negative definite. We also define the linear transformation A_τ by:

$$A_\tau = A|_{W_\tau}$$

Since A restricted to any subspace is an isomorphism of subspaces, and π restricted to any maximally negative definite subspace is an isomorphism, we can write the following commutative diagram:

$$\begin{tikzcd}
W_\tau \arrow[r, "A_\tau"] \arrow[d, "\pi"'] & A_\tau(W_\tau) \arrow[d, "\pi"] \\
V_- \arrow[r, "B_\tau"'] & V_-
\end{tikzcd}$$

where B_τ is a continuous curve in $GL_t(\mathbb{R})$ given by the composition of isomorphisms:

$$B_\tau = \pi|_{A_\tau(W_\tau)} \circ A_\tau \circ (\pi|_{W_\tau})^{-1} \tag{1.2.2}$$

We see that if A has time orientability $+1$, then $A_0 : V_- \to V_-$ is an orientation preserving isomorphism, so, by (1.2.2), B_0 is also an orientation preserving isomorphism, implying that $\det(B_0) > 0$. We define a function:

$$\begin{aligned} \gamma(\tau) : [0,1] &\longrightarrow \mathbb{R} \\ \tau &\longmapsto \det(B_\tau) \end{aligned}$$

Since det is a continuous function, and B_τ is a continuous curve, we see that $\gamma(\tau)$ is also a continuous function. Suppose then that at some point τ_0 we have $\gamma(\tau_0) < 0$, by the intermediate value theorem this implies that there exists a $c \in (0, \tau_0)$ such that:

$$\gamma(c) = 0 \Rightarrow \det(B_c) = 0$$

but $B_\tau \in GL_t(\mathbb{R})$ for all $\tau \in [0,1]$, so no such τ_0 can exist. Thus, for all $\tau \in [0,1]$ we see that $\det(B_\tau) > 0$, so A_1 must be an orientation preserving isomorphism $W \to A(W)$.

For the other direction, assume that $A|_W$ is an orientation preserving isomorphism $W \to A(W)$. By an argument similar to the one above, we then have that B_1 is an orientation preserving isomorphism. By the continuity of γ, B_0 is then also an orientation preserving isomorphism, thus $A_0 : V_- \to A(V_-)$ also preserves the orientation of V_-, which completes the proof. □

Importantly, **Lemma 1.2.2** gives us the following proposition:

Proposition 1.2.2. *Suppose $A, B \in O(t,s)$ both have time orientability $+1$, then AB and A^{-1} also have time orientability $+1$. Furthermore, if A is written as the block matrix:*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{1.2.3}$$

then A is time orientable if and only if $\det(A_{11}) > 0$.

Proof. Let $W = A(V_-)$, then A^{-1} is an orientation preserving isomorphism from $W \to A^{-1}(W) = V_-$, where W is a maximally negative definite subspace of V, then by **Lemma 1.2.2**, A^{-1} has time orientability 1.

Let $W = B(V_-)$, then we see that AB is the composition:

$$AB : V_- \xrightarrow{B|_{V_-}} W \xrightarrow{A|_W} A(W)$$

Since A has time orientability $+1$, by **Lemma 1.2.2**, $A|_W$ is an orientation preserving isomorphism, so AB is an orientation preserving isomorphism, and thus has time orientability $+1$.

Suppose A has time orientability $+1$, then the define the linear map $F : V_- \to V_-$ by $F = \pi|_{A(V_-)} \circ A|_{V_-}$. We see that F is given by:

$$F = A_{11}$$

where A_{11} is the first block in (1.2.3), and since it must preserve the orientation of V_- by construction, we see that $\det(A_{11}) > 0$. If $\det(A_{11}) > 0$, then we have that $\pi|_{A(V_-)} \circ A|_{V_-} : V_- \to V_-$, is an orientation preserving isomorphism. So, since $\pi|_{A(V_-)}$ is an orientation preserving isomorphism, $A|_{V_-} : V_- \to A(V_-)$ must also be an orientation preserving isomorphism, hence A has time orientability $+1$. □

With **Proposition 1.2.2**, we are now in a position to move on to our penultimate example of the section.

Example 1.2.6. We write $A \in O(t,s)$ as the block matrix:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where $A_{11} \in \mathrm{Mat}_{t\times t}(\mathbb{R})$, $A_{22} \in M_{s\times s}(\mathbb{R})$, $A_{21} \in \mathrm{Mat}_{s\times t}(\mathbb{R})$ and $A_{12} \in \mathrm{Mat}_{t\times s}(\mathbb{R})$. **Proposition 1.2.2** then implies that the set:

$$O^+(t,s) = \{A \in O(t,s) : \det(A_{11}) > 0\}$$

is a subgroup of $O(t,s)$. Since $\det\left(\pi|_{A(V_-)} \circ A|_{V_-}\right)$ is continuous, we see that $O^+(t,s)$ is open in $O(t,s)$, and thus an open submanifold of $O(t,s)$, implying that $O^+(t,s)$ is Lie subgroup of $O(t,s)$. Since $O^+(t,s)$ preserves the orientation of any subspace of time like vectors of $\mathbb{R}^n$, we call this group the **Orthochronus Pseudo Orthogonal Group**.

We also have the **Special Pseudo Orthogonal Group**, given by:

$$SO(t,s) = \{A \in O(t,s) : \det(A) = 1\}$$

and the **Proper Orthochronus Group**:

$$SO^+(t,s) = \{A \in O(t,s) : \det(A) = 1, \det(A_{11}) > 0\}$$

which can both be shown to be Lie subgroups of $O(t,s)$ via our previous arguments. None of these groups are compact, and all have dimension equal to $\dim O(t,s)$. Finally, one can also verify directly that $O(t,s)$ has four connected components, and, by the use of homogenous spaces, that $SO^+(t,s)$ is the connected component of the identity.

Example 1.2.7. For the vector space $\mathbb{C}^n$, equipped with the standard Hermitian norm, we can generalize **Example 1.2.3** and **Example 1.2.4** with the following Lie groups:

$$\begin{aligned} U(n) =& \{A \in GL_n(\mathbb{C}) : A^\dagger A = I\} \\ SU(n) =& \{A \in U(n) : \det(A) = 1\} \end{aligned}$$

where $A^\dagger$ denotes the Hermitian transpose of A, obtained by taking the complex conjugate of the entries A, and then applying the transpose operation. $U(n)$ is called the **Unitary Group**, and $SU(n)$ is called the **Special Unitary Group**. It is easy to verify that $U(1) \cong S^1$ and $SU(2) \cong S^3$.

It turns out the definition of a Lie group is redundant. Indeed if multiplication is smooth then inversion is smooth, as the next lemma shows.

Lemma 1.2.3. *Let G be a group which is at the same time a smooth manifolds such that the multiplication map:*

$$\begin{aligned} \mu : G \times G &\longrightarrow G \\ (g,h) &\longmapsto gh \end{aligned}$$

is smooth. Then, the inversion map:

$$\begin{aligned} i : G &\longrightarrow G \\ g &\longmapsto g^{-1} \end{aligned}$$

is smooth.

Proof. Let $\gamma : I \to G$ be a smooth curve such that $\gamma(0) = h$ and $\dot{\gamma}(0) = X \in T_h G$, then:

$$\begin{aligned} D_h L_g(X) =& \lim_{t \to 0} \frac{d}{dt} L_g(\gamma(t)) \\ =& \lim_{t \to 0} g \cdot \frac{d}{dt}\gamma(t) \\ =& g \cdot X \end{aligned}$$

Note that the final line is a mild abuse of notation, and is really a stand in for $D_h L_g(X)$. In the case where the Lie group is is a matrix Lie group, then, as will show, tangent vectors are also matrices, and the final line truly is multiplication. For right multiplication we have a similar result:

$$D_h R_g(X) = X \cdot g$$

We now show that μ is a submersion. Let $(g,h) \in G \times G$, and let $\gamma_1, \gamma_2 : I \to G$ be smooth curves such that $\gamma_1(0) = g$, $\gamma_2(0) = h$, $\dot{\gamma}_1(0) = X \in T_g G$, and $\dot{\gamma}_2(0) = Y \in T_h G$. We see that:

$$\begin{aligned} D_{(g,h)}\mu(X,Y) =& \lim_{t \to 0} \frac{d}{dt}(\gamma_1(t)\gamma_2(t)) \\ =& \lim_{t \to 0} \left(\frac{d}{dt}(\gamma_1(t))\gamma_2(t) + \gamma_1(t)\frac{d}{dt}(\gamma_2(t)) \right) \\ =& X \cdot h + g \cdot Y \\ =& D_g R_h(X) + D_h L_g(Y) \end{aligned}$$

We note that the map L_g and R_g are a diffeomorphisms, as they are smooth and have smooth inverses given by $L_{g^{-1}}$ and $R_{g^{-1}}$ respectively. Let $Z \in T_{gh}G$, then since L_g is a diffeomorphism, there exists an $Y \in T_gG$ such that $D_hL_g(Y) = Z$. It follows that:

$$D_{(g,h)}\mu(0, Y) = Z \in T_{gh}Z$$

Since $(g, h) \in G \times G$, and $Z \in T_{gh}G$ were arbitrarily, we have that μ is a submersion by **Definition 1.1.7**.

By **Theorem 1.1.1** we have that the set:

$$\mu^{-1}(e) = \{(g, g^{-1}) \in G \times G\}$$

is an embedded submanifold of $G \times G$ of dimension $\dim \mu^{-1}(e) = \dim G$. Let f be the embedding $\mu^{-1}(e) \to G \times G$, and π_i the projection on the to ith copy of G in $G \times G$. Furthermore, let ϕ be the map:

$$\begin{aligned} \phi : G &\longrightarrow \mu^{-1}(e) \\ g &\longmapsto (g, g^{-1}) \end{aligned}$$

We want to show that ϕ is smooth. First note that the composition $\pi_1 \circ f : \mu^{-1}(e) \to G$ is smooth and satisfies:

$$\pi_1 \circ f(g, g^{-1}) = g$$

It is also clearly a bijection, we want to show this map is a diffeomorphism. By **Proposition 1.1.2**, we need only show that $D_p(\pi_1 \circ f)$ an isomorphism for all $p \in \mu^{-1}(e)$. By rank nullity, we need only show the map is injective. Note that for any $p \in \mu^{-1}(e)$, we have that:

$$\mu(g, g^{-1}) = e$$

Let $\gamma : I \to \mu^{-1}(e)$ be some smooth curve in $\mu^{-1}(e) \subset G \times G$, satisfying $\gamma(0) = p$, and $\dot{\gamma}(0) = X \in T_p\mu^{-1}(e)$. The differential of the embedding is injective, thus $D_pf(X) \neq (0, 0)$ for all $X \in T_p\mu^{-1}(e)$, and $T_{f(p)}(G \times G) \cong T_p\mu^{-1}(e)$. Furthermore, we have that $\mu \circ f(p) = e$ for all $p \in \mu^{-1}(e)$, so:

$$D_p(\mu \circ f)(X) = 0$$

hence $D_pf(X) \in \ker D_{f(p)}\mu$, implying that $T_{f(p)}(G \times G) \subset \ker D_{f(p)}\mu$. However, by rank nullity:

$$2 \dim G = \dim G + \dim \ker D_{f(p)}\mu \Longrightarrow \dim \ker D_{f(p)} = \dim G$$

hence $T_{f(p)}(G \times G) = \ker D_{f(p)}\mu \cong T_p\mu^{-1}(e)$. Let $D_pf(X) = (Y, Z)$, then, if $f(p) = (g, g^{-1})$:

$$D_{(g,g^{-1})}(Y, Z) = D_gR_{g^{-1}}(Y) + D_{g^{-1}}L_g(Z) = 0$$

Note that if Y is zero, then we must have that Z is zero as well, and vice versa, as both maps are isomorphisms. Thus, we must have that:

$$D_pf(X) \neq (0, Z)\text{or}(Y, 0)$$

This implies that:

$$D_p(\pi_1 \circ f)(X) = 0$$

only when X is zero, hence $D_p(\pi_1 circ f)$ is injective, implying $\pi_1 \circ f$ is a diffeomorphism. However for all $(g, g^{-1}) \in \mu^{-1}$:

$$\phi \circ (\pi_1 \circ f)(g, g^{-1}) = \phi(g) = (g, g^{-1})$$

while:

$$(\pi_1 \circ f) \circ \phi(g) = (\pi_1 \circ f)(g, g^{-1}) = g$$

so ϕ is the inverse of $\pi_1 \circ f$, and thus must be smooth. It follows that the inversion map i is the composition of smooth maps:

$$i = \pi_2 \circ \phi$$

and thus smooth as desired.

□

1.2.2 Lie Algebras

Let G be a Lie group, then since G is a smooth manifold, by **Proposition 1.1.8**, $\mathfrak{X}(G)$ is a Lie algebra over $\mathbb{R}$. In this section we wish to study a specific Lie subalgebra of $\mathfrak{X}(G)$, which we will call the Lie algebra of G. As we shall shortly, this Lie subalgebra will be isomorphic to the tangent space of G at the identity. We begin with the following definition:

Definition 1.2.5. Let G be a Lie group. A vector field $X \in \mathfrak{X}(G)$ is called **left invariant** if for all $g \in G$:

$$L_{g*}X = X$$

The preceding definition implies that a left invariant vector field X transforms pointwise by:

$$L_{g*}(X_h) = X_{gh} \tag{1.2.4}$$

We need the following proposition:

Proposition 1.2.3. *The set:*

$$\mathfrak{g} = \{X \in \mathfrak{X}(G) : \forall g \in G,\ L_{g*}X = X\}$$

is a Lie subalgebra of $\mathfrak{X}(G)$.

Proof. Let X and Y be left invariant vector fields, and $a, b \in \mathbb{R}$, then the vector field:

$$W = aX + bY$$

is also left invariant as for any $g \in G$:

$$\begin{aligned} L_{g*}W =& L_{g*}(aX + bY) \\ =& L_{g*}(aX) + L_{g*}(bY) \\ =& aL_{g*}X + bL_{g*}Y \\ =& aX + bY = W \end{aligned}$$

This show thats that $\mathfrak{g}$ is vector subspace of $\mathfrak{X}(G)$. To show that it is a Lie subalgebra, we must also show that it is closed under the bracket operation. Let $X, Y \in \mathfrak{g}$, then by **Proposition 1.1.9**

$$\begin{aligned} L_{g*}[X, Y] =& [L_{g*}X, L_{g*}Y] \\ =& [X, Y] \end{aligned}$$

so $[X, Y] \in \mathfrak{g}$, thus $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{X}(M)$. □

Definition 1.2.6. Let G be a Lie group, then the Lie algebra of G, denoted $\mathfrak{g}$, is the Lie subalgebra of left invariant vector fields on G.

Proposition 1.2.4. *Let G be a Lie group, and $\mathfrak{g}$ be the corresponding Lie algebra. Then, as Lie algebras:*

$$\mathfrak{g} \cong T_eG$$

Proof. Suppose G is an n dimensional Lie group, and let $v \in T_eG$, then we define a smooth left invariant vector field by:

$$\omega(v)_g = D_eL_g(v)$$

We first check that this is indeed a smooth vector field. The smooth map:

$$\begin{aligned} \mu : G \times G &\longrightarrow G \\ (g, h) &\longmapsto gh \end{aligned}$$

has differential:

$$\begin{aligned} D\mu : TG \times TG &\longrightarrow TG \\ ((g, x), (h, y)) &\longmapsto D_gR_h(x) + D_hL_g(y) \end{aligned}$$

which is itself a smooth map. So $\omega(v)$ is just the composition:

$$\begin{aligned}\omega(v) : &G \longrightarrow TG\\ &g \longmapsto D_{\mu}((g,0),(e,v)) = D_e L_g(v)x\end{aligned}$$

and is thus a smooth vector field. Furthermore, $\omega(v)$ is left invariant as for any $h, g \in G$:

$$\begin{aligned}L_{h*}\omega(v)_g =&D_g L_H \circ D_e L_g(v)\\ =&D_e L_h \circ L_g(v)\\ =&D_e L_{hg}(v)\\ =&\omega(v)_{hg}\end{aligned}$$

Finally, for any $a, b \in \mathbb{R}$, and any $v, w \in T_e G$, $\omega(av + bw)$ is the smooth vector field defined by:

$$\begin{aligned}\omega(av+bw)_g =&D_e L_g(av+bw)\\ =&D_e L_g(av) + D_e L_g(bw)\\ =&aD_e L_g(v) + bD_e L_g(w)\\ =&a\omega(v)_g + b\omega(w)_g\end{aligned}$$

Hence $\omega(av + bw)$ is the linear combination of the vector fields:

$$\omega(av + bw) = a\omega(v) + b\omega(w)$$

so ω is a linear map. This map is an isomorphism since it has inverse given by:

$$\begin{aligned}\lambda : \mathfrak{g} &\longrightarrow T_e G\\ X &\longmapsto X_e\end{aligned}$$

We first show the map is linear For any $a, b \in \mathbb{R}$, and $X, Y \in \mathfrak{g}$, $aX + bY \in \mathfrak{g}$, so:

$$\begin{aligned}\lambda(aX+bY) =&aX_e + bY_e\\ =&a\lambda(X) + b\lambda(Y)\end{aligned}$$

implies that λ is linear. For any $X \in \mathfrak{g}$:

$$\omega \circ \lambda(X) = \omega(X_e)$$

Furthermore, for any $g \in G$, we see that by (1.2.4):

$$\omega(X_e)_g = L_{g*}X_e = X_g$$

so $\omega(X_e) = X$, and λ is indeed the inverse of ω. Thus, ω is an isomorphism of vector spaces, and:

$$\mathfrak{g} \cong T_e G$$

as desired. We now equip $T_e G$ with the the Lie bracket defined by:

$$[v, w] = \lambda([\omega(v), \omega(w)])$$

and see that:

$$\omega([v, w]) = \omega(\lambda[\omega(v), \omega(w)]) = [\omega(v), \omega(w)]$$

so $\mathfrak{g} \cong T_e G$ as Lie algebras. □

Note that in the preceding proof we have that by definition:

$$[v, w] = [\omega(v), \omega(w)]_e$$

Moreover, **Proposition 1.2.4** gives the following corollary:

Corollary 1.2.2. *The Lie algebra of a Lie group G is finite dimensional, satisfying* $\dim \mathfrak{g} = \dim G$. *Furthermore, a left invariant vector field is entirely determined by it's value at a point.*

Note that by choosing a basis for T_eG, we can construct a global frame for TG using the map ω. Indeed, if $\{T_i\}$ is a basis for T_eG then for each T_i, $\omega(T_i)$ is a globally defined smooth vector which vanishes nowhere, as for any $g \in G$:

$$\omega(T_i)_g = D_eL_g(T_i) \tag{1.2.5}$$

which can't be the zero vector as D_eL_g is an isomorphism $T_eG \to T_gG$. Furthermore, this implies that if $\dim(G) = n$, then $TG \cong G \times \mathbb{R}^n$ via the diffeomorphism:

$$(g, x_1, \dots, x_n) \longmapsto \sum_{i=1}^{n} x_i\omega(T_i)_g$$

We now determine the Lie algebras for the Lie groups discussed in the previous sections.

Example 1.2.8. Let $G = GL_n(\mathbb{R})$, as an open submanifold of $\mathrm{Mat}_{n\times n}(\mathbb{R})$, we see that:

$$T_IG = T_I\mathrm{Mat}_{n\times n}(\mathbb{R}) = \mathrm{Mat}_{n\times}(\mathbb{R})$$

where I is the identity matrix. So the Lie algebra of $GL_n(\mathbb{R})$, denoted $\mathfrak{gl}_n(\mathbb{R})$, is the space of all real valued n by n matrices. By writing an $A \in GL_n(\mathbb{R})$ as the matrix $(A)^i_j$, and and a curve $\gamma : I \to GL_n(\mathbb{R})$ as the matrix $(\gamma)^j_k$, where $(\dot\gamma(0))^j_k = (X)^j_k$ for some $X \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ we see that:

$$\begin{aligned}(D_IL_A(X))^i_k =& \frac{d}{dt}\Big|_{t=0}(A)^i_j(\gamma)^j_k \\ =&(A)^ij(X)^j_k\end{aligned}$$

so:

$$D_IL_A(X) = A \cdot X$$

Moreover, since left and right multiplication by g is a diffeomorphism, we have that $T_gGL_n(\mathbb{R}) = g \cdot \mathrm{Mat}_{n\times n}(\mathbb{R}) \cong \mathrm{Mat}_{n\times n}(\mathbb{R})$, so any $X \in T_gGL_n(\mathbb{R})$ can be written as $g \cdot Y$ or $Z \cdot g$ for some $Y, Z \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. It follows that $D_hg(X)$ for any $h, g \in Gl_n(\mathbb{R})$ and any $X \in T_hGL_n(\mathbb{R})$ is truly given by matrix multiplication. If $X, Y \in \mathfrak{gl}_n(\mathbb{R})$, then we have that:

$$[X, Y]_I = (\mathscr{L}_XY)_I = \lim_{t\to 0}\frac{d}{dt}D_{\theta_t(I)}\theta_{-t}(Y_{\theta_t(I)})$$

As we shall see shortly, the flow $\theta(t, g)$ of any left invariant vector field X is given by:

$$\theta(t, g) = R_{\exp(tX)}g$$

where $\exp(tX) : I \to G$ is the soon to be defined exponential map. For now we will just assume that $\exp(tX)$ is the unique integral curve of $X \in \mathfrak{gl}_n(\mathbb{R})$ starting at I, and that $\theta_{-t} = R_{-\exp(tX)}$. With this we have that:

$$[X, Y]_I = \lim_{t\to 0}\frac{d}{dt}D_{\exp(tX)}R_{\exp(-tX)}(Y_{\exp(tX)})$$

Note that:

$$\begin{aligned}D_{\exp(tX)}R_{\exp(-tX)}(Y_{\exp(tX)}) =& D_I(R_{\exp(-tX)} \circ L_{\exp(tX)})(Y_I) \\ =& \exp(tX) \cdot Y_I \cdot \exp(-tX)\end{aligned}$$

So:

$$\begin{aligned}[X, Y]_I =& \lim_{t\to 0}\frac{d}{dt}\exp(tX) \cdot Y_I \cdot \exp(-tX) \\ =& X_I \cdot Y_I - Y_I \cdot X_I\end{aligned}$$

so the Lie bracket is the standard commutator on $\mathrm{Mat}_{n\times n}(\mathbb{R})$. This same argument then shows that the Lie bracket of the Lie algebra of each linear group amounts to standard commutator.

Example 1.2.9. Let $G = SL_n(\mathbb{R})$, and let $\gamma : J \to SL_n(\mathbb{R})$ be a smooth curve passing through the identity at $t = 0$, such that $\dot\gamma(0) = X$ for some arbitrary $X \in T_eSL_n(\mathbb{R})$. We see that:

$$\det(\gamma(t)) = 1$$

for all $t \in J$. Differentiating at $t = 0$ we obtain:

$$\begin{aligned} 0 &= \lim_{t\to 0} \frac{d}{dt} \det(\gamma(t)) \\ &= \mathrm{Tr}(X) \end{aligned}$$

We see that $\mathfrak{sl}_n(\mathbb{R})$ is n^2-1 dimensional, so since the kernel of the linear map $\mathrm{Tr} : \mathrm{Mat}_{n\times n}(\mathbb{R}) \to \mathbb{R}$ is $n^2 - 1$ we have that:

$$\mathfrak{sl}_n(\mathbb{R}) = \{X \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : \mathrm{Tr}(X) = 0\}$$

Example 1.2.10. Leg $G = O(n)$, and let $\gamma : J \to SL_n(\mathbb{R})$ be a smooth curve passing through the identity at $t = 0$ such that $\dot\gamma(0) = X$ for some arbitrary $X \in T_IO(n)$. Then:

$$\gamma^T(t)\gamma(t) = I$$

Differentiating at $t = 0$ we obtain:

$$\begin{aligned} 0 &= \lim_{t\to 0} \dot\gamma^T(t)\gamma(t) + \gamma(t)^T\dot\gamma(t) \\ &= X^T + X \end{aligned}$$

So the Lie algebra of $O(n)$, denoted $\mathfrak{o}(n)$ satisfies:

$$\mathfrak{o}(n) \subset \{X \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : X^T + X = 0\} = V$$

We note that any $X \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ can be written as:

$$(X + X^T)/2 + (X - X^T)/2$$

where the left term is a symmetric matrix, and right term is antisymmetric, and that $V \cap Sym_n(\mathbb{R}) = \{0\}$, so:

$$\mathrm{Mat}_{n\times n}(\mathbb{R}) = V \oplus Sym_n(\mathbb{R})$$

It follows that $\dim_{\mathbb{R}} V = n^2 - (n^2 + n)/2 = (n^2 - n)/2$, hence:

$$\mathfrak{o}(n) = \{X \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : X^T + X = 0\}$$

Furthermore, since $SO(n)$ is an open submanifold of $O(n)$, we have that $T_ISO(n) \cong T_IO(n)$, hence the Lie algebra of $SO(n)$, denoted $\mathfrak{so}(n)$ is isomorphic to $\mathfrak{o}(n)$.

Example 1.2.11. Let $G = O(t, s)$, and $\gamma : J \to O(t, s)$ a smooth curve passing through the identity at $t = 0$, such that $\dot\gamma(0) = X$ for some arbitrary $X \in T_IO(t, s)$. Then:

$$\gamma^T(t)\eta\gamma(t) = \eta$$

Differentiating at $t = 0$, we obtain:

$$X^T\eta + \eta X = 0$$

So the Lie algebra of $O(t, s)$, denoted $\mathfrak{o}(t, s)$ is:

$$\mathfrak{o}(t, s) \subset \{X \in \mathrm{Mat}_{n\times n}(\mathbb{R}) : X^T\eta + \eta X = 0\}$$

and a similar dimension argument demonstrates equality. By the same argument in **Example 1.2.10**, we see that the Lie algebras of $O^+(t, s)$, $SO(t, s)$, and $SO^+(t, s)$, respectively denoted $\mathfrak{o}^+(t, s)$, $\mathfrak{so}(t, s)$ and $\mathfrak{so}^+(t, s)$, satisfy:

$$\mathfrak{o}^+(t, s) \cong \mathfrak{so}(t, s) \cong \mathfrak{so}^+(t, s) \cong \mathfrak{o}(t, s)$$

We now move on to a brief a discussion on induced Lie algebra homomorphisms. Let $\phi : G \to H$ be a Lie group homomorphism, and $X \in \mathfrak{g}$, then, since $\phi(e) = e$, we can define a unique a left invariant vector field $\phi_* X$, by:

$$(\phi_* X)_h = D_e L_h \circ D_e \phi(X_e) \tag{1.2.6}$$

where $h \in H$. Note that $\phi_* X$ is smooth since it is the the composition of smooth maps:

$$H \xrightarrow{\quad e \quad} \{e\} \xrightarrow{\quad X \quad} T_e G \xrightarrow{\quad D_e\phi \quad} T_e H \xrightarrow{\quad \omega \quad} \mathfrak{h}$$

where e is the trivial homomorphism. Essentially, we have used the group structure of G and H, namely the fact they both have a preferred element e satisfying $\phi(e) = e$ for any homomorphism, and **Corollary 1.2.2** to bypass the requirement of a smooth inverse, which, in the general smooth manifold case, guarantees that the push forward is well defined. Note that this clearly only holds for left invariant vector fields.

Definition 1.2.7. Let G and H be Lie groups, and ϕ a Lie group homomorphism between them. The map:

$$\begin{aligned} \phi_* : &\mathfrak{g} \longrightarrow \mathfrak{h} \\ &X \longmapsto \phi_* X \end{aligned}$$

is called the **induced homomorphism**

We now must show that this induced homomorphism is actually a Lie algebra homomorphism.

Proposition 1.2.5. *Let $\phi : G \to H$ be a Lie group homomorphism, then the induced homomorphism, ϕ_*, is a Lie algebra homomorphism.*

Proof. We have to show that:

$$\phi_*[X_1, X_2] = [\phi_* X_1, \phi_* X_2]$$

So, by **Proposition 1.1.9**, we need only show that $\phi_* X_i$ is ϕ related to X_i. From (1.2.6), we see that for any $g \in G$:

$$\begin{aligned} (\phi_* X)_{\phi(g)} =& D_e L_{\phi(g)} \circ D_e \phi(X_e) \\ =& D_e (L_{\phi(g)} \circ \phi)(X_e) \\ =& D_e(\phi \circ L_g)(X_e) \\ =& D_g \phi(X_g) \end{aligned}$$

So $\phi_* X$ is ϕ related to X, and thus ϕ_* is a Lie algebra homomorphism. $\square$

So far, the group structure of the Lie group has allowed us to construct a variety convenient global properties which are not necessarily available in the general smooth manifold case. As we shall see shortly, this motif carries over to integral curves as well, and will eventually lead us to the famed exponential map, a way of locally identifying small enough open sets of G with its Lie algebra.

Theorem 1.2.2. *Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. Let:*

$$\phi_X : \mathbb{R} \supset I \longrightarrow G \tag{1.2.7}$$

$$t \longmapsto \phi_X(t) \tag{1.2.8}$$

denote the maximal integral curve of a left invariant vector field $X \in \mathfrak{g}$ through the identity element e. The following then hold:

a) $\phi_X(s+t) = \phi_X(s) \cdot \phi_X(t)$ $\forall s, t \in \mathbb{R}$

b) $\phi_X(t)$ *is defined for all* $t \in \mathbb{R}$

c) $\phi_{sX}(t) = \phi_X(st)$ $\forall s, t \in R$.

Proof. We begin with a weaker version of a). Assume ϕ_X is only defined on some open interval $I = (t_{\min}, t_{\max})$, and that $s \in I$, then we want to show that:

$$\phi_X(s) \cdot \phi_X(t) = \phi_X(s+t)$$

so long as $s+t \in I$. By **Lemma 1.1.4**, $\phi_X(s+t)$ is then an integral curve on $I_s = (t_{\min}-s, t_{\max}-s)$, starting at $\phi_X(s)$. Let $\phi_X(s) = g \in G$, then:

$$\begin{aligned}\alpha : I &\longrightarrow G \\ t &\longmapsto g \cdot \phi_X(t)\end{aligned}$$

is an integral curve starting at g. Indeed α starts at g since:

$$\alpha(0) = g$$

and α is an integral curve as:

$$\begin{aligned}\frac{d}{dt}\alpha(t) =& gX_{\phi_X(t)} \\ =& D_{\phi_X(t)}L_g X_{\phi_X(t)} \\ =& X_{g\phi_X(t)} \\ =& X_{\alpha(t)}\end{aligned}$$

By the uniqueness of integral curves, we then have that for all $t \in I \cap I_s$:

$$\phi_X(s) \cdot \phi_X(t) = \phi_X(s+t)$$

as desired. Proving *b*) will now imply *a*), we would like to show that no upper or lower bound exists for I. Proceeding by way of contradiction, suppose that there exist a $t_{\max}$ and $t_{\min}$ such that ϕ_X is defined only on $I = (t_{\min}, t_{\max})$. Let $\xi = \min(|t_{\min}|, t_{\max})$, and consider the curve:

$$\begin{aligned}\gamma : \left(t_{\min} + \frac{\xi}{2}, t_{\max} + \frac{\xi}{2}\right) &\longrightarrow G \\ t &\longmapsto \phi_X\left(\frac{\xi}{2}\right)\phi_X\left(t - \frac{\xi}{2}\right)\end{aligned}$$

γ then starts at e as:

$$\begin{aligned}\gamma(0) =& \phi_X\left(\frac{\xi}{2}\right)\phi_X\left(-\frac{\xi}{2}\right) \\ =& \phi(0) \\ =& e\end{aligned}$$

Furthermore this is an integral curve as:

$$\begin{aligned}\frac{d}{dt}\gamma(t) =& \phi_X\left(\frac{\xi}{2}\right)X_{\phi(t-\xi/2)} \\ =& X_{\gamma(t)}\end{aligned}$$

If $t_{\min} = -\infty$ then $\gamma : (-\infty, t_{\max} + \xi/2) \to G$ is a smooth extension of ϕ_X, and if $t_{\min} > -\infty$, then we have that:

$$\psi(t) = \begin{cases} \phi_X(t) & \text{if } t \in (t_{\min}, t_{\max}) \\ \gamma(t) & \text{if } t \in (t_{\min} + \xi/2, t_{\max} + \xi/2) \end{cases}$$

is also an extension of ϕ_X, as $\gamma(t) = \phi_X(t)$ for all $t \in (t_{\min} + \xi/2, t_{\max})$, so no such $t_{\max}$ can exist. Via a similar argument, we can prove that no such $t_{\min}$ exists, hence $\phi_X(t)$ is defined on all of $\mathbb{R}$, so for all $s, t \in \mathbb{R}$:

$$\phi_X(s+t) = \phi_X(s) \cdot \phi_X(t)$$

To show *c*), we recall that for any $s \in \mathbb{R}$, ϕ_{sX} is the integral curve of the left invariant vector field sX, starting at e. Fixing s, we define a curve:

$$\begin{aligned}\beta : \mathbb{R} &\longrightarrow G \\ t &\longmapsto \phi_X(st)\end{aligned}$$

which is a smooth curve in G by b)). We see that this curve clearly starts at e, and is also an integral curve of sX as, by the chain rule:

$$\begin{aligned}\frac{d}{dt}\beta(t) =& sX_{\phi_X(st)} \\ =& sX_{\beta(t)}\end{aligned}$$

Hence, by the uniqueness of integral curves, for any $s \in R$:

$$\phi_{sX}(t) = \phi_X(st)$$

as desired. □

We can now define the exponential map:

Definition 1.2.8. Let $\phi_X : \mathbb{R} \to G$ denote the integral curve of $X \in \mathfrak{g}$ starting at e. Then we define the **exponential map** as:

$$\begin{aligned}\exp : \mathfrak{g} &\longrightarrow G \\ X &\longmapsto \exp(X) = \phi_X(1)\end{aligned}$$

If G is a matrix Lie group, then it can be shown that the exponential map is literally the matrix exponential encountered in ODE's or physics, i.e.

$$\exp(X) = \sum_{n=0}^{\infty} \frac{X^n}{n!}$$

Proposition 1.2.6. *Let G be a Lie group, and $\mathfrak{g}$ its Lie algebra. Then, for all $s, t \in \mathbb{R}$, and $X \in \mathfrak{g}$, the exponential map satisfies:*

a) $\exp(0) = e$

b) $\exp((s+t)X) = \exp(sX)\exp(tX)$

c) $\exp(-X) = \exp(X)^{-1}$

Proof. We begin with a)). Let $\phi_0(t)$ be the integral curve of the left invariant 0 vector field. Then we see that:

$$\begin{aligned}\phi_0(1) =& \phi_0(-1)\phi_0(2) \\ =& \phi_0(-1)\phi_{(2\cdot 0)}(1) \\ =& \phi_0(-1)\phi_0(1) \\ =& \phi_0(0) \\ =& e\end{aligned}$$

so:

$$\exp(0) = \phi_0(1) = e$$

To prove b)), we note that for any $s, t \in \mathbb{R}$, $(s+t)X \in \mathfrak{g}$, hence:

$$\begin{aligned}\exp((s+t)X) =& \phi_{(s+t)X}(1) \\ =& \phi_X(s+t) \\ =& \phi_X(s)\phi_X(t) \\ =& \phi_{sX}(1)\phi_{tX}(1) \\ =& \exp(sX)\exp(tX)\end{aligned}$$

Finally, for c)),

$$\exp((-1+1)X) = e \Rightarrow \exp(-X)\exp(X) = e \Rightarrow \exp(-X) = \exp(X)^{-1}$$

as desired. □

As the next proposition shows, the flow of a left invariant vector field is intimately related to the exponential map.

Proposition 1.2.7. *Let G be a Lie group, and denote it's Lie algebra by $\mathfrak{g}$. The flow of $X \in \mathfrak{g}$:*

$$\begin{aligned}\theta :&\mathbb{R} \times G \longrightarrow G\\ &(t, g) \longmapsto \theta(t, g)\end{aligned}$$

is defined for all $\mathbb{R}$, and satisfies:

$$\theta(t, g) = g \cdot \exp(tX) = R_{\exp(tX)} g = L_g \exp(tX) \tag{1.2.9}$$

Proof. We define $\theta(t, g)$ for all $t \in \mathbb{R}$ by (1.2.9), then by the definition of the exponential map, and **Theorem 1.2.2**, $\theta(t, g)$ is defined for all of $\mathbb{R}$. We will show this definition is actually the smooth global flow of X. We see that for all $g \in G$:

$$\theta(0, g) = g \cdot \exp(0) = g$$

and that for all $s, t \in \mathbb{R}$, and all $g \in G$:

$$\begin{aligned}\theta(t, \theta(s, g)) =&\theta(t, g \cdot \exp(sX))\\ =&(g \cdot \exp(sX)) \cdot \exp(tX)\\ =&g \exp((s + t)X)\\ =&\theta(s + t, g)\end{aligned}$$

so θ as defined is a smooth global flow. Finally, note that:

$$\exp(tX) = \phi_{tX}(1) = \phi_X(t)$$

where ϕ_X is the integral curve of X. Then:

$$\begin{aligned}\frac{d}{dt}\bigg|_{t=0} \theta(t, g) =&\lim_{t \to 0} \frac{d}{dt} g \cdot \exp(tX)\\ =&\lim_{t \to 0} g \cdot \frac{d}{dt} \phi_X(t)\\ =&D_e L_g X_e\\ =&X_g\end{aligned}$$

so $\theta(t, g)$ is the flow of X as desired. □

The preceding proposition allows one to easily define curves starting at any point in G, with any initial velocity $X \in T_g G$. Furthermore, the integral curves $\theta(t, e) = \exp(tX)$ are one dimensional embedded abelian Lie subgroups of G. In particular, the map $\exp(tX)$ is a Lie group homomorphism $\mathbb{R} \to G$, and is a Lie group isomorphism onto its image.

Proposition 1.2.8. *Under the canonical identifications:*

$$T_0 \mathfrak{g} \cong \mathfrak{g}, \qquad T_e G \cong \mathfrak{g}$$

The differential of the exponential map at 0 is the identity:

$$D_0 \exp : \mathfrak{g} \longrightarrow T_0 G$$

is the identity map. In particular, there exists open neighborhoods V of 0 in $\mathfrak{g}$ and U of e in G such that:

$$\exp|_V : V \longrightarrow U$$

is a diffeomorphism.

Proof. Let $X \in \mathfrak{g}$ and $\gamma(t) = tX$ be the curve of constant velocity X in $\mathfrak{g}$. Then:

$$\begin{aligned} D_0 \exp(X) =& \frac{d}{dt}\Big|_{t=0} \exp(tX) \\ =& X \end{aligned}$$

hence $D_0 \exp$ is the identity map. The claim then follows by applying the inverse function theorem in some small enough coordinate charts around 0 and e. □

In this sense, exp is a local diffeomorphism, and provides us with a convenient description of the local behavior of G around the identity. In rare cases, as the following example shows, the exponential map can give 'most' of G.

Example 1.2.12. Let $G = U(1)$, then:

$$U(1) = \{z \in \mathbb{C} : z\bar{z} = 1\}$$

Let $X \in T_1U(1)$, and let $\gamma : I \to U(1)$ be the curve satisfying $\gamma(0) = 1$, $\dot{\gamma}(0) = X$. Then we see that:

$$\begin{aligned} 0 =& \frac{d}{dt}\Big|_{t=0} \gamma(t)\bar{\gamma}(t) \\ =& X + \bar{X} \end{aligned}$$

hence $X \in i\mathbb{R}$. The exponential map is then:

$$\begin{aligned} \exp : i\mathbb{R} &\longrightarrow U(1) \\ it &\longmapsto e^{it} \end{aligned}$$

We see that for any $\epsilon \in (-\pi, \pi)$, exp restricts to a diffeomorphism:

$$(-\pi + \epsilon, \pi - \epsilon) \longrightarrow U(1) \setminus \{e^{i(i\pi - \epsilon)}\}$$

So the exponential map gives us the entirety of $U(1)$ minus a point.

We end our discussion on Lie algebras with the following proposition:

Proposition 1.2.9. *Let $\phi : G \to H$ be a Lie group homomorphism, then for all $X \in \mathfrak{g}$:*

$$\phi(\exp(X)) = \exp(\phi_* X)$$

where $\phi_* : \mathfrak{g} \to \mathfrak{h}$ is the induced Lie algebra homomorphism

Proof. Let $\gamma(t)$ be the curve in H:

$$\gamma(t) = \phi(\exp(tX))$$

for some $X \in \mathfrak{g}$. We then see that via the chain rule:

$$\begin{aligned} \dot{\gamma}(t) =& D_{\exp tX}\phi \left(\frac{d}{ds}\Big|_{s=t} \exp(sX) \right) \\ =& D_{\exp tX}\phi \left(X_{\exp(tX)} \right) \\ =& D_e(\phi \circ L_{\exp(tX)})(X_e) \\ =& D_e(L_{\phi(\exp(tX))} \circ \phi)(X_e) \\ =& (\phi_* X)_{\gamma(t)} \end{aligned}$$

Thus γ is the unique integral of the left invariant vector field $\phi_* X$ starting at e. By our prior work the integral curve for $\phi_* X$ is given by $\exp(t\phi_* X)$, thus we obtain that:

$$\gamma(t) = \phi(\exp(tX)) = \exp(t\phi_* X)$$

Setting $t = 1$ then proves the claim. □

1.2.3 The Maurer Cartan Form

In this section we introduce the Maurer-Cartan form, a special type of one form on G that will prove of a great importance for our work in gauge theory. It is the first appearance in this paper of what we call a vector valued, or twisted k-form, a generalization of k-forms we now make precise.

Definition 1.2.9. Let M be a smooth manifold and V be a vector space. Furthermore, let $C^\infty(M, W)$ be the set of all smooth maps $M \to V$. A **k-form on M with values in V** is an alternating map:

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k \text{ copies}} \longrightarrow C^\infty(M, V)$$

The set of all k-forms with values V, or k-forms twisted in V, is denoted by $\Omega(M, V)$ and can be identified with:

$$\Omega(M, V) \cong \Omega(M) \otimes_{\mathbb{R}} V$$

Now, let G be a Lie group, and let $X \in \mathfrak{g}$, then at any point $g \in G$, we can pull X_g back to the identity via left multiplication by g^{-1}, i.e.

$$L_{g^{-1}*}X_g = \tilde{X} \in T_eG \cong \mathfrak{g}$$

It is important to note that X is, in general, not left invariant, so it would be incorrect to assert that $\tilde{X} = X_e$. With this in mind, we provide the following the definition:

Definition 1.2.10. The **Maurer-Cartan Form**, $\mu_G \in \Omega(G, \mathfrak{g})$, is a one form on G with values in the Lie algebra defined by:

$$\mu_G(X)_g = D_g L_{g^{-1}}(X)$$

for all $g \in G$, and $X \in T_gG$.

So μ_G associates to a vector $X \in T_gG$ a left invariant vector field $\tilde{X}$ satisfying $\tilde{X}_g = X$. In the context of matrix Lie groups, one also finds the following notation employed:

$$\mu_G = g^{-1}dg$$

where dg can be thought of some summand of one forms dual to some global frame. In particular, this notation proves rather convenient for explicit computation, as the following example shows.

Example 1.2.13. Let $G = SO(2)$, then the dimension of $SO(2) = 1$, and we can parameterize $SO(2)$ via:

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Then, in these coordinates:

$$\begin{aligned} \mu_G =& g^{-1}dg \\ =& \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -\sin\theta d\theta & -\cos\theta d\theta \\ \cos\theta d\theta & -\sin\theta d\theta \end{pmatrix} \\ =& \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\theta \end{aligned}$$

In this notation, it is clear that μ_G is an element of $\Omega(SO(2)) \otimes_{\mathbb{R}} \mathfrak{so}(2)$.

The Maurer-Cartan form also transforms nicely under pullbacks by right or left multiplication.

Proposition 1.2.10. *Let G be a Lie group, and μ_G the Maurer cartan from on G. Then, for all $g \in G$:*

$$L_g^*\mu_G = \mu_G \qquad \text{and} \qquad R_g^*\mu_G = c_{g^{-1}*} \circ \mu_G \tag{1.2.10}$$

where $c_{g^{-1}*}$ is the induced homomorphism of the conjugation map:

$$\begin{aligned} c_{g^{-1}} : G &\longrightarrow G \\ h &\longmapsto g^{-1}hg \end{aligned}$$

Proof. We prove this for all $g, h \in G$, and all $X \in T_h G$:

$$\begin{aligned}(L_g^*\mu_G)(X)_h =& \mu_G(D_h L_g X)_{gh}\\ =& D_{gh}L_{(gh)^{-1}} \circ D_h L_g X\\ =& D_h L_{h^{-1}} \circ L_{g^{-1}} \circ L_g(X)\\ =& D_h L_{h^{-1}}(X)\\ =& \mu_G(X)_h\end{aligned}$$

and further that:

$$\begin{aligned}(R_g^*\mu_G)(X)_h =& \mu_G(D_h R_g(X))_{hg}\\ =& D_{hg}L_{(hg)^{-1}} \circ D_h R_g(X)\\ =& D_h L_{g^{-1}} \circ L_{h^{-1}} \circ R_g(X)\\ =& D_e(L_{g^{-1}} \circ R_g) \circ D_g L_g^{-1}(X)\\ =& D_e c_g^{-1} \circ \mu_G(X)_h\\ =& c_{g^{-1}*} \circ \mu_G(X)_h\end{aligned}$$

Since this holds for all $h, g \in G$, and all $X \in T_h G$, we have proven the claim. □

1.2.4 Group Actions on Manifolds

In this section, we finally see how Lie groups bring about smooth symmetries on our manifolds. Recall that all of the Lie groups we discussed were first defined as subsets of the group of isomorphisms on some real or complex vector space V, so each of those Lie groups has a multiplicative action on that vector space given by the inclusion homomorphism $\phi : G \to GL(V)$. We can extend this to any real or complex vector space W by specifying a homomorphism $\rho : G \to GL(W)$, which is called a representation of G on W. We will tackle the specific case of representations in a subsequent section, but for now we wish to discuss the general case of a G action on a smooth manifold M. We first need the following proposition:

Proposition 1.2.11. *Let M be a smooth manifold, then the set of the diffeomorphisms on M, denoted Diff(M), is a group under composition.*

Proof. First note that the identity map, $\mathrm{Id}_M : M \to M$ is a diffeomorphism and hence contained in $\mathrm{Diff}(M)$. This is clearly the neutral element under composition, as for any $F \in \mathrm{Diff}(M)$:

$$\mathrm{Id}_M \circ F = F = F \circ \mathrm{Id}_M$$

Furthermore, for $F, G \in \mathrm{Diff}(M)$, we also have that $F \circ G$ and $G \circ F$ are in $\mathrm{Diff}(M)$. Indeed, since F and G are smooth homeomorphisms, $F \circ G$ and $G \circ F$ must be also be a smooth homeomorphisms; they must also have smooth inverses as both F^{-1} and G^{-1} are smooth, so their composition must be smooth as well. Finally, as the composition of maps is associative, and each $F \in \mathrm{Diff}(M)$ has an inverse by construction, we see that $\mathrm{Diff}(M)$ is indeed a group[10]. □

We would now like group actions on M to be homomorphisms $\phi : G \to \mathrm{Diff}(M)$, motivating our next definition.

Definition 1.2.11. Let G be a Lie group, and M a smooth manifold. A **left action** of G on M is a smooth map:

$$\begin{aligned}\Phi : G \times M &\longrightarrow M\\ (g, p) &\longmapsto \Phi(g, p) = g \cdot p\end{aligned}$$

such that the following hold for all $p \in M$, and $g, h \in G$:

- $(h \cdot g) \cdot p = h \cdot (g \cdot p)$

[10]With more technical machinery at hand, one can put a topology, and smooth structure on Diff(M) such that Diff(M) is a Lie group, and it can be further shown that the corresponding Lie algebra of Diff(M) is isomorphic to $\mathfrak{X}(M)$.

- $e \cdot p = p$

Furthermore, a **right action** is a smooth map:

$$\begin{aligned}\Phi : M \times G &\longrightarrow M\\ (p, g) &\longmapsto \Phi(p, g) = p \cdot g\end{aligned}$$

such that the following hold for all $p \in M$, and $g, h \in G$:

- $p \cdot (g \cdot h) = (p \cdot g) \cdot h$
- $p \cdot e = p$

Given a left action of G on M, and fixing a $g \in G$, we see that map:

$$\begin{aligned}L_g : M &\longrightarrow M\\ p &\longmapsto g \cdot p\end{aligned}$$

is a diffeomorphism, since it is smooth and has smooth inverse given by $L_{g^{-1}}$. Furthermore, the map sending $g \to L_g$, satisfies $e \longmapsto \mathrm{Id}_M$, and:

$$L_h \circ L_g(p) = L_h \circ (g \cdot p) = h \cdot (g \cdot p) = (h \cdot g) \cdot p = L_{hg}(p)$$

Thus, we see that a left action on M corresponds to a homomorphism $\phi : G \to \mathrm{Diff}(M)$. If we were given a right action of G on M, then we instead obtain a homomorphism $\phi : G^{op} \to \mathrm{Diff}(M)$, where G^{op} is the Lie group G equipped with the multiplicative action $\bullet$ defined for all $g, h \in G$ by:

$$g \bullet h = h \cdot g$$

Proposition 1.2.12. *Let G be a Lie group, M a smooth manifold, and Φ a left group action on M. Then, the map ψ:*

$$\begin{aligned}\Psi : M \times G &\longrightarrow M\\ (p, g) &\longmapsto p \bullet g = \Phi(g^{-1}, p) = g^{-1} \cdot p\end{aligned}$$

is a right action on M.

Proof. We first note that ψ is smooth, since if $i \times \mathrm{Id}_M : M \times G \to G \times M$ is the smooth map:

$$i \times \mathrm{Id}_M(p, g) = (g^{-1}, p)$$

then Ψ is the composition $\Phi \circ (i \times \mathrm{Id}_M)$. Furthermore, we see that for all $p \in M$:

$$p \bullet e = e^{-1} \cdot p = p$$

Finally, for all $g, h \in G$, and all $p \in M$,

$$p \bullet (g \cdot h) = (g \cdot h)^{-1} \bullet p = h^{-1} \cdot (g^{-1} \cdot p) = (p \bullet g) \bullet h$$

thus Ψ is a right action on M by definition. □

We will also need the following constructions to better describe group actions.

Definition 1.2.12. Let G be a Lie group, M a smooth manifold, and Φ and Ψ be left and right actions of G on M respectively.

a) For a left action, the **orbit** of G through a point p is the set:

$$\mathcal{O}_p = \{q \in M : \exists g \in G, q = \Phi(g, p) = g \cdot p\}$$

For a right action, the definition is similar:

$$\mathcal{O}_p = \{q \in M : \exists g \in G, q = \Psi(p, g) = p \cdot g\}$$

Both sets can be thought of as the images of an **orbit map**, denoted $\phi_p : G \to M$, or $\psi_p : G \to M$ respectively. The maps are defined via:

$$\phi_p(g) = g \cdot p \qquad \psi_p(g) = p \cdot g \tag{1.2.11}$$

b) For a left action, the **fixed point set** of an element $g \in G$ is given by:

$$M^g = \{p \in M : g \cdot p = p\}$$

For a right action we have:

$$M^g = \{p \in M : p \cdot g = p\}$$

c) For a left action, the **isotropy group** of a point $p \in G$ is:

$$G_p = \{g \in G : g \cdot p = p\}$$

For a right action:

$$G_p = \{g \in G : p \cdot g = p\}$$

Example 1.2.14. Let $M = \mathbb{S}^2$, and let $G = SO(3)$. Thinking of each point in $\mathbb{S}^2$ as a unit vector in $\mathbb{R}^3$, we let $SO(3)$ act on the left of M via matrix multiplication. Note that for any $\theta \in [0, 2\pi]$, the matrices:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \qquad B = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \qquad C = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are all in $SO(3)$. Furthermore, in the standard basis of $\mathbb{R}^3$, these matrices correspond to rotations around the x axis, y axis, and z axis respectively. For any $x \in \mathbb{S}^2$, we can obtain any other $y \in \mathbb{S}^2$ via some composition of A, B, and C acting on x, thus the orbit of any point in $\mathbb{S}^2$ is the entirety of $\mathbb{S}^2$. Furthermore, if $x \in \mathbb{S}^2$ is written in the standard basis as (x_1, x_2, x_3), then, via the Gram Schmidt algorithm, we can find some other basis where $x = (1, 0, 0)$. We see that in this basis, any matrix written like A leaves x unchanged, thus the isotropy group of any x is given by the set of all matrices $A \in SO(3)$. We note that this isotropy group is clearly isomorphic to $SO(2)$.

Proposition 1.2.13. *Let G be a Lie group, M a smooth manifold, and let Φ and Ψ be left and right actions of G on M respectively. For any points $p, q \in M$, the orbits of p and q are either disjoint or identical.*

Proof. We prove this for a right action Ψ, and note that the proof for a left action Φ is entirely analogous. Let $\mathcal{O}_p$ and $\mathcal{O}_q$ be the orbits of G through p and q respectively. Suppose for the sake of contradiction that these orbits are neither disjoint nor identical, then there exists an $n \in M$ such that $n \in \mathcal{O}_p \cap \mathcal{O}_q$, thus there exists $g, h \in G$ such that:

$$p \cdot g = n \qquad \text{and} \qquad q \cdot h = n$$

Let $q'\mathcal{O}_q$ be arbitrary, then we see that $q' = q \cdot l$ for some $i \in G$, but $q' \in \mathcal{O}_p$ as:

$$p \cdot (g \cdot h^{-1} \cdot i) = n \cdot (h^{-1} \cdot i) = q \cdot i = q'$$

so $\mathcal{O}_q \subset \mathcal{O}_p$. Furthermore, if $p' \in \mathcal{O}_p$ is arbitrary, then $p' = p \cdot j$ for some $j \in G$, but $p' \in \mathcal{O}_q$ as:

$$q \cdot (h \cdot g^{-1} \cdot j) = n \cdot (g^{-1} \cdot j) = p \cdot j = p'$$

so $\mathcal{O}_p \subset \mathcal{O}_q$. Therefore, $\mathcal{O}_p = \mathcal{O}_q$, a contradiction, so if $\mathcal{O}_p \cap \mathcal{O}_q \neq \emptyset$ we are forced into equality, thus two orbits are either disjoint or identical, as desired. $\square$

The preceding proposition allows us to obtain a partition of M via the disjoint union of the orbits of G. Furthermore, this partition is the determined by the equivalence relation:

$$p \sim q \iff \exists g \in G, p = q \cdot g$$

i.e. the two equivalence classes $[p]$ and $[q]$, which are the orbits of p and q respectively, are equal if and only if their orbits are identical. This partition will be of grave importance, so we give it a formal definition below.

Definition 1.2.13. Let Φ be a left or right action of G on M. Then the set:

$$M/G = \{\mathcal{O}_p \subset M : p \in M\}$$

is the called the **space of orbits**, or the **quotient space** of the action. Furthermore, the map:

$$\begin{aligned} \pi : M &\longrightarrow M/G \\ p &\longmapsto [p] \end{aligned}$$

is called the **canonical projection**. If $x \in M/G$, and if for some $p \in M$ we have $[p] = x$, then p is called a **representative** of x.

In particular, we are most interested in the cases when the space M/G is a smooth manifold. Furthermore, the convention when dealing with quotients in gauge theory is to use a right action of G on M, so going forward we will only consider right actions, though the corresponding case for left actions will usually be similar. That being said, for now we state the following definitions, and leave our discussion of quotients for later.

Definition 1.2.14. Let Φ be a right action of G on M, then:

a) The action is **transitive** if the orbit map ϕ_p is surjective for all $p \in M$. In other words, M consists of one orbit.

b) The action is **free** if the orbit map ϕ_p is injective for all $p \in M$.

c) The action is **simply transitive** if the orbit map ϕ_p is injective and surjective for all $p \in M$.

As we saw earlier, a homomorphism between Lie groups $\phi : G \to H$ gives rise to an induced homomorphism between Lie algebras. Since group actions can be viewed as a a homomorphism $\phi : G \to \text{Diff}(M)$, we wish to understand the induced homomorphism[11]

$$\phi_* : \mathfrak{g} \longrightarrow \mathfrak{X}(M)$$

We need the following definition:

Definition 1.2.15. Let Φ be a right action of G on M, and let $X \in \mathfrak{g}$. The **fundamental vector field** $\tilde{X}$, associated to X is then defined by:

$$\tilde{X}_p = \frac{d}{dt}\bigg|_{t=0} p \cdot \exp(tX)$$

If ϕ_p is the orbit map through the point p, then the construction of $\tilde{X}$ can be equivalently defined as:

$$\tilde{X}_p = D_{(e,p)}\Phi(X_e, 0) = D_e\phi_p(X_e)$$

This construction is just the restriction of the global differential $D\Phi : TG \times TM \to TM$ to the Cartesian product of the singleton set $\{X\} \subset T_eG$ with the zero section of TM, so $\tilde{X}$ is indeed a smooth vector field on M.

Proposition 1.2.14. *Let Φ be a smooth free right action of G on M. Then the map:*

$$\begin{aligned} \phi_* : \mathfrak{g} &\longrightarrow \mathfrak{X}(M) \\ X &\longmapsto \tilde{X} \end{aligned}$$

is injective.

Proof. We only need show that the kernel of ϕ_* is trivial, i.e. only $0 \in \mathfrak{g}$ maps to the zero vector field. Suppose the contrary, that for some non zero $X \in \mathfrak{g}$ we have that:

$$\phi_*(X) = 0$$

This then implies that for all $p \in M$:

$$\begin{aligned} \tilde{X}_p =& \frac{d}{dt}\bigg|_{t=0} p \cdot \exp(tX) \\ =& 0 \end{aligned}$$

[11] Note that for a right action the domain of the homomorphism is G^{op}. Despite working with right actions, we elect to ignore this as nothing truly depends on it.

So the curve $\gamma(t) = p \cdot \exp(tX)$ has zero velocity at $t = 0$, however, we then obtain that for arbitrary $s \in \mathbb{R}$:

$$\begin{aligned}\gamma'(s) =& \frac{d}{dt}\Big|_{t=s} p \cdot \exp(tX)\\ =& \frac{d}{dt}\Big|_{t=0} p \cdot \exp((t+s)X)\\ =& \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tX) \cdot \exp(sX)\\ =& D_p R_{\exp(sX)} \left(\frac{d}{dt}\Big|_{t=0} p \cdot \exp(tX) \right)\\ =& 0\end{aligned}$$

So $\gamma'(s)$ is zero for all s, and thus γ is the constant curve equal to p for all time. Therefore, $\phi_p(e) = \phi_p(\exp(tX))$ for all $t \in \mathbb{R}$, hence ϕ_p is not injective, a contradiction, so ϕ_* is injective as desired. □

Proposition 1.2.15. *Let Φ be a right action of G on M. The map:*

$$\begin{aligned}\phi_* : \mathfrak{g} &\longrightarrow \mathfrak{X}(M)\\ X &\longmapsto \tilde{X}\end{aligned}$$

is then a Lie algebra homomorphism. In particular, the fundamental vector fields form a Lie subalgebra of $\mathfrak{X}(M)$.

Proof. Fix a point $p \in M$ and let ϕ_p be the orbit map through said point. Then, for some $X, Y \in \mathfrak{g}$ and $a, b \in \mathbb{R}$:

$$\begin{aligned}\phi_*(aX + bY)_p =& D_e\phi_p(aX + bY)\\ =& aD_e\phi_p(X) + bD_e\phi_p(Y)\\ =& a\phi_*(X)_p + b\phi_*(Y)\end{aligned}$$

so the map is linear. Furthermore, we need to show that the left invariant vector field $X \in \mathfrak{g}$ is ϕ related to $\tilde{X}$, i.e. that:

$$\tilde{X}_{\phi_p(g)} = D_g\phi_p(X_g)$$

for all $g \in G$. Let $g \in G$ be arbitrary, then:

$$\begin{aligned}\tilde{X}_{\phi_p(g)} =& D_e\phi_{\phi_p(g)}(X_e)\\ =& D_e\phi_{\phi_p(g)} \circ D_g L_{g^{-1}}(X_g)\\ =& D_g\left(\phi_{\phi_p(g)} \circ L_{g^{-1}}\right)(X_g)\end{aligned}$$

Note that for any $h \in G$:

$$\begin{aligned}\phi_{\phi_p(g)} \circ L_{g^{-1}}(h) =& \phi_{\phi_p(g)}(g^{-1} \cdot h)\\ =& \phi_{\phi_p(g)}(e) \cdot g^{-1} \cdot h\\ =& p \cdot g \cdot g^{-1} \cdot h\\ =& p \cdot h\\ =& \phi_p(h)\end{aligned}$$

So we see that $\phi_{\phi_p(g)} \circ L_{g^{-1}} = \phi_p$, hence:

$$\tilde{X}_{\phi_p(g)} = D_g\phi_p(X_g)$$

so $\tilde{X}$ and X are ϕ_p related. From **Proposition 1.1.9**, this then implies that:

$$[\tilde{X}, \tilde{Y}] = [\phi_* X, \phi_* Y] = \phi_*[X, Y] = \widetilde{[X, Y]}$$

so ϕ_* is a Lie algebra homomorphism, and the set of fundamental vector fields in $\mathfrak{X}(M)$ is a vector subspace that closed under the Lie bracket, and hence a Lie subalgebra. □

Though fundamental vector fields are ϕ_p related to left invariant vector fields on G, as the next proposition shows, fundamental vector fields are not in general invariant under a group action of G on M.

Proposition 1.2.16. *Let Φ be a right action of G on M, and let $\tilde{X}$ be a fundamental vector field associated to $X \in \mathfrak{g}$. Then:*

$$R_{g*}\tilde{X} = \tilde{Y}$$

where $\tilde{Y}$ is the fundamental vector field associated to $Y = c_{g^{-1}*}X$.

Proof. We see that at any point $p \in M$:

$$\begin{aligned}(R_{g*}\tilde{X})_p =& D_{p\cdot g^{-1}}R_g(\tilde{X}_{pg^{-1}})\\ =& D_{p\cdot g^{-1}}R_g \circ D_e\phi_{p\cdot g^{-1}}(X_e)\\ =& D_e(R_g \circ \phi_{p\cdot g^{-1}}) \circ D_{g^{-1}}L_g(X_{g^{-1}})\\ =& D_{g^{-1}}(R_g \circ \phi_{p\cdot g^{-1}} \circ L_g)(X_{g^{-1}})\end{aligned}$$

Note that for any $h \in G$, we have that:

$$\begin{aligned}\phi_{p\cdot g^{-1}} \circ L_g(h) =& \phi_{p\cdot g^{-1}}(gh)\\ =& (p \cdot g^{-1}) \cdot (g \cdot h)\\ =& p \cdot h\\ =& \phi_p(h)\end{aligned}$$

so $\phi_{p\cdot g^{-1}} \circ L_g = \phi_p$, hence:

$$\begin{aligned}(R_{g*}\tilde{X})_p =& D_{g^{-1}}(R_g \circ \phi_p)(X_{g^{-1}})\\ =& D_e(R_g \circ \phi_p \circ L_{g^{-1}})(X_e)\end{aligned}$$

Furthermore, we see that for any $h \in G$:

$$R_g \circ \phi_p(h) = p \cdot h \cdot g$$

while[12]:

$$\phi_p \circ R_g(h) = \phi_p(h \cdot g) = p \cdot h \cdot g$$

so:

$$\phi_p \circ R_g = R_g \circ \phi_p$$

Therefore we obtain:

$$\begin{aligned}(R_{g*}\tilde{X})_p =& D_e(\phi_p \circ L_{g^{-1}} \circ R_g)(X_e)\\ =& D_e\phi_p \circ D_e c_{g^{-1}}(X_e)\\ =& D_e\phi_p(Y_e)\\ =& \tilde{Y}_p\end{aligned}$$

$\square$

We end our discussion on group actions by using the Maurer-Cartan form to calculate the differential of a right action Φ.

Proposition 1.2.17. *Let Φ be a right action of G on M, then the the differential of Φ at a point $(p, g) \in G \times M$ is the map:*

$$\begin{aligned}D_{(p,g)}\Phi : T_pM \oplus T_gG &\longrightarrow T_{p\cdot g}M\\ (X, Y) &\longmapsto D_pR_g(X) + \widetilde{\mu_G(Y)}_{p\cdot g}\end{aligned}$$

Where $\widetilde{\mu_G(Y)}$ is the fundamental vector field associated to the Lie algebra element $\mu_G(Y)$.

[12]Here we employ a mild abuse of notation; clearly these two R_g maps are not the same.

Proof. Let γ and ψ be curves in M and G tangent to $X \in T_pM$ and $Y \in T_gG$ at $t = 0$ respectively.

$$\begin{aligned} D_{(p,g)}\Phi(X,Y) =& \frac{d}{dt}\Big|_{t=0} \gamma(t) \cdot \psi(t) \\ =& \dot{\gamma}(0) \cdot \psi(0) + \gamma(0) \cdot \dot{\psi}(0) \\ =& D_pR_g(X) + D_g\phi_p(Y) \end{aligned}$$

In **Proposition 1.2.15** we saw that for a left invariant vector field $Z \in \mathfrak{g}$, that the fundamental vector field $\tilde{Z}$ associated to Z is ϕ_p related to Z, so:

$$\tilde{Z}_{\phi_p(g)} = D_g\phi_p(Z_g)$$

for all $g \in G$. Letting $Z = \mu_G(Y)$, at the point $g \in G$ we have that $Z_g = Y$, so:

$$D_g\phi_p(Y) = D_g\phi_p(Z_g) = \tilde{Z}_{p\cdot g}$$

Therefore the differential of Φ is given by:

$$D_{(p,g)}\Phi(X,Y) = D_pR_g(X) + \widetilde{\mu_G(Y)}_{p\cdot g}$$

as desired. $\square$

1.2.5 Quotient Manifolds

In the previous section, for a right action Φ of G on M, we defined the set M/G as the space of orbits:

$$M/G = \{\mathcal{O}_p \subset M : p \in M\}$$

We also showed that each $\mathcal{O}_p$ is either disjoint or identical, implying that M/G is a partition of M via the equivalence relation:

$$p \sim q \iff \exists g \in G,\ p \cdot g = q$$

As a set, M/G is then the quotient of M via the aforementioned equivalence relation. As discussed in the last section, we would like to know when M/G is a smooth manifold. We will first begin more generally, with any partition of M due to an equivalence relation $\sim$, and then take a theorem from Godement as god given to develop the M/G case. For more complete discussions on quotient manifolds, see Lee's *Smooth Manifolds* or Hamilton's *Mathematical Gauge Theory.*

We first take M to be a set, with an equivalence relation $\sim$, then we define the set R as

$$R = \{(p,q) \in M \times M : p \sim q\}$$

We also define *equivalence classes* as subsets of M:

$$[p] = \{q \in M : q \sim p\}$$

and finally the quotient set M/R as:

$$M/R = \{[p] : p \in M\}$$

There also exists the canonical projection $\pi : M \to M/R$ taking each element to it's equivalence class. We now examine the case where M is a a topological space; we need the following definition:

Definition 1.2.16. Let M be a topological space, and $\sim$ an equivalence relation on M. Then we define a topology on M/R, called the **quotient topology**, such that $\pi : M \to M/R$ is continuous. In other words, a set U in M/R are open if and only if $\pi^{-1}(U)$ is open in M.

With this quotient topology on M/R, we would first to like to determine when M/R is a Hausdorff topological space.

Lemma 1.2.4. *Let M be a topological space:*

a) If M/R is Hausdorff, then $R \subset M \times M$ is closed.

b) If $\pi : M \to M/R$ is open, and $R \subset M \times M$ is closed, then M/R is Hausdorff.

Proof. We use the fact from topology that a topological space N is Hausdorff if and only if the diagonal:

$$\Delta = \{(p,p) \in N \times N : p \in N\}$$

is a closed subset in $N \times N$; let Δ be the diagonal of M/R. We begin with a), by first noting that the map:

$$\begin{aligned} \pi \times \pi : M \times M &\longrightarrow M/R \times M/R \\ (p,q) &\longmapsto ([p],[q]) \end{aligned}$$

is continuous. Therefore, since Δ is closed in $M/R \times M/R$, $(\pi \times \pi)^{-1}(\Delta)$ is closed in $M \times M$. Furthermore, if $p \sim q$ we have $[p] = [q]$, hence:

$$\pi \times \pi(p,q) = ([p],[p]) \in \Delta$$

implying that

$$(\pi \times \pi)^{-1}(\Delta) = \{(p,q) \in M \times M : p \sim q\} = R$$

So R is closed in $M \times M$.

For b) assume that $\pi : M \to M/R$ is open, and that R is closed in $M \times M$. The map $\pi \times \pi$ is then an open map, and so the image of $(M \times M) \setminus R$ is open in $M/R \times M/R$. Furthermore, an element $([p],[q])$ is in $\pi \times \pi((M \times M) \setminus R)$ if and only if $[p] \neq [q]$, implying that:

$$([p],[q]) \in \pi \times \pi(M \times M \setminus R) \Longleftrightarrow ([p],[q]) \in (M/R \times M/R) \setminus \Delta$$

So the compliment of the image is Δ, and thus closed, so M/R is Hausdorff. ☐

We now site the aforementioned theorem from Godement, and proceed with the case where R is determined by a group action of G on a topological space M.

Theorem 1.2.3. *Let R be an equivalence relation on a smooth manifold M. Suppose that R is a closed embedded submanifold of $M \times M$ and $\pi_1|_R : R \to M$ a surjective submersion. Then M/R has a unique smooth structure of a smooth manifold such that the canonical projection $\pi : M \to M/R$ is a surjective submersion.*

With this theorem, we will determine when M/G is a smooth manifold. We first show that the canonical projection $\pi : M \to M/G$ is open.

Lemma 1.2.5. *Let M be a topological space, and Φ a right group action of G on M.*[13] *Then the canonical projection $\pi : M \to M/G$ is open.*

Proof. Our earlier work implies that each orbit through a point $p \in M$ is an equivalence class of the equivalence relation:

$$p \sim q \Longleftrightarrow \exists g \in G, \; p \cdot g = q$$

Let U be an open set of M, if π is open then $\pi(U)$ is open, so, by the definition of the quotient topology on M/G, we need to show that $\pi^{-1}(\pi(U))$ is open in M. We see that:

$$\pi(U) = \{[p] : p \in U\}$$

and that for any $[p] \in \pi(U)$, we have that:

$$\pi^{-1}([p]) = \mathcal{O}_p = \bigcup_{g \in G} R_g(\{p\})$$

[13] We defined right group actions to be smooth as a priori, but in this case we take Φ to be continuous. Similarly, G need only be a topological group for this lemma to hold. Going forward, when we make statements about right actions on topological spaces, we will take Φ to be continuous and G to be a topological group. If instead M is at any point stated to be a smooth manifold, it should be assumed that G is a Lie group, and Φ is smooth.

So the inverse image of $\pi(U)$ will be the union of the orbits through each $p \in U$, i.e.

$$\pi^{-1}(\pi(U)) = \bigcup_{p \in U} \mathcal{O}_p = \bigcup_{g \in G} R_g(U)$$

However, each R_g is a homeomorphism,[14] so $\pi^{-1}(\pi(U))$ is open in M, hence π is open as desired. $\square$

Again, working only topologically, **Lemma 1.2.4** and **Lemma 1.2.5** imply the following corollary:

Corollary 1.2.3. *The quotient space M/G is Hausdorff if and only if the map:*

$$\begin{aligned} \Psi : M \times G &\longrightarrow M \times M \\ (p, g) &\longmapsto (p, p \cdot g) \end{aligned}$$

has closed image.

Proof. Note that the image of Ψ is R, since if $(p, q) \in \Psi(M \times G)$, we have that for some $g \in G$, $q = p \cdot g$, so $(p, q) \in R$, and if $(p, q) \in R$ we know that $q = p \cdot g$, so $(p, q) \in \Psi(M \times G)$. Then, by **Lemma 1.2.5**, the map $\pi : M \to M/G$ is open, so if $\Psi(M \times G)$ is closed, **Lemma 1.2.4** implies that M/G is Hausdorff. Furthermore, if M/G is Hausdorff then by **Lemma 1.2.4**, $\Psi(M \times G)$ is closed. $\square$

We need to do develop some more topological results before moving onwards to the smooth manifold case.

Definition 1.2.17. A topological space M is **locally compact** if every point in M has a compact neighborhood.

Lemma 1.2.6. *Let M be a locally compact Hausdorff space. Then a subset $A \subset M$ is closed if and only if the intersection of A with any subset of M is compact.*

Proof. First let A be closed, and K be any compact subset of M. Since M is Hausdorff, K is closed, and thus $A \cap K$ is closed. Furthermore, $A \cap K$ is compact since it is a closed subset of K, and K is compact.

Now assume that $A \cap K$ is compact for any compact $K \subset M$. Let $p \in M \setminus A$, then, since M is locally compact, there exists an open neighborhood of p, $U \subset M$, contained in some compact $K \subset M$. We have that $A \cap K$ is compact, and thus closed in M since M is Hausdorff. Then:

$$U \setminus (A \cap K) = U \cap (A \cap K)^c = U \cap (M \setminus (A \cap K))$$

is a an open subset of $M \setminus A$, since it is the intersection of two of open sets in M. Furthermore, $p \in U \setminus (A \cap K)$, so it is an open neighborhood of p, thus, since p was arbitrary, every point in $M \setminus A$ has an open neighborhood, and thus $M \setminus A$ is open M. Therefore A is closed as desired. $\square$

We continue with the following definition:

Definition 1.2.18. Let $f : M \to N$ be a continuous map. We say that f is **proper** if for every compact $K \subset N$, $f^{-1}(A)$ is compact in M.

We can then use **Lemma 1.2.6** to prove the following:

Lemma 1.2.7. *Let $f : M \to N$ be a proper map between topological spaces, with N locally compact, and Hausdorff. Then f is closed.*

Proof. By **Lemma 1.2.6**, it suffices to show that for every closed $A \subset M$, and all compact $K \subset N$, $f(A) \subset K$ is compact. We see that for any compact $K \subset M$, since f is proper, $f^{-1}(K)$ is compact, so $A \cap f^{-1}(K)$ is compact, and thus, since proper maps are continuous, $f(A \cap f^{-1}(K)) = f(A) \cap K$ is compact. $\square$

Lemma 1.2.8. *If $f : M \to N$ is a closed continuous map, between topological spaces such that $f^{-1}(q)$ is compact for all $q \in N$, then f is proper.*

[14]Diffeomorphism if M is a smooth manifold, G is a Lie group, and Φ a right group action as defined earlier.

Proof. We first want to show that if f is closed, then for all $q \in N$, and open subset $O \subset M$ such that $f^{-1}(q) \subset O$, there exists an open neighborhood U of q such that $f^{-1}(U) \subset O$. Let $U = N \setminus f(M \setminus O)$; note that since $M \setminus O$ is closed, U is open. Furthermore:

$$f(M \setminus O) = \{p \in N : \exists x \in M \setminus O,\ p = f(x)\}$$

so $q \notin f(M \setminus O)$ as $f^{-1}(q) \subset O$, thus $q \in U$, so U is an open neighborhood of q. Finally, suppose $x \in f^{-1}(U)$ but not in O, then that implies that $f(x) \in f(M \setminus O)$, but $x \in f^{-1}(U) \subset N \cap f(M \setminus O)^c$, so $f(x) \in f(M \setminus O)^c$, a contradiction, thus $f^{-1}(U) \subset O$.

Now, let $K \subset N$ be compact, and suppose that $\mathcal{U}$ is any open covering of $f^{-1}(K)$. For all $y \in K$, we have that $f^{-1}(y) \subset f^{-1}(K)$ is compact, so finitely many $U_i \in \mathcal{U}$ cover $f^{-1}(y)$, i.e. for some $n \in \mathbb{N}$ there exist n subsets $U_i \in \mathcal{U}$ such that:

$$f^{-1}(y) \subseteq \bigcup_{i=1}^{n} U_i$$

By our earlier work, there then exists a $V_y \subset N$, such that:

$$f^{-1}(V_y) \subseteq \bigcup_{i=1}^{n} U_i$$

Construct a V_y for all $y \in K$, then:

$$\mathcal{V} = \{V_y : y \in Y\}$$

is an open cover of K. Since K is compact, there exists an $m \in \mathbb{N}$ such that for some $Y = \{y_1, \ldots, y_m\} \subset N$:

$$\mathcal{V}_m = \{V_{y_i} : y_i \in Y\}$$

We see that:

$$K \subset \bigcup_{i=1}^{m} V_{y_i} \Rightarrow f^{-1}(K) \subset f^{-1}\left(\bigcup_{i=1}^{m} V_{y_i}\right) \Rightarrow f^{-1}(K) \subset \bigcup_{i=1}^{m} f^{-1}(V_{y_i})$$

So:

$$f^{-1}(\mathcal{V}_m) = \{f^{-1}(V_{y_i}) : y_i \in Y\}$$

is a finite refinement of $\mathcal{U}$. Thus, $f-1(K)$ is compact, which implies the claim. □

We now move back to our discussion on group actions with the following definition.

Definition 1.2.19. A right group action Φ of G on a topological space M is called **proper** if the map:

$$\begin{aligned} \Psi : M \times G &\longrightarrow M \times M \\ (p, g) &\longmapsto (p, p \cdot g) \end{aligned}$$

is proper.

We then have the following two corollaries

Corollary 1.2.4. *Let Φ be a proper right group action of G on a topological space M, where M is locally compact and Hausdorff. Then the aforementioned map Ψ is closed, and M/G is Hausdorff.*

Proof. **Lemma 1.2.7** shows that Ψ is closed, then, **Corollary 1.2.3** shows that M/G is Hausdorff. □

Corollary 1.2.5. *Let Φ be a right group action of G on a Hausdorff space M. If G is compact, then Φ is proper.*

Proof. Let $K \subset M \times M$ be compact, then $\pi_1(K)$ is a compact subset of M. Furthermore, since M is Hausdorff, K is closed in $M \times M$, so by continuity $\Psi^{-1}(K)$ is closed and $M \times G$. Finally, since $\Psi(x, g) \in K$ implies that $x \in \pi_1(K)$, we have that $\Psi^{-1}(K) \subset L \times G$. Since $L \times G$ is compact because G is compact, and $\Psi^{-1}(K)$ is closed, $\Psi^{-1}(K)$ is a compact as well, hence Φ is a proper map. $\square$

We are now finally in the position to determine when M/G is a smooth manifold. We need the following definition:

Definition 1.2.20. Let M be a smooth manifold, and G a Lie group. A right action Φ of G on M is then called **principal** if the action is free and the map Ψ, as defined earlier, is closed.

We also need the following facts about submersions:

Lemma 1.2.9. *Let $F : M \to N$ be a surjective submersion between smooth manifolds. Then F admits smooth local sections, i.e. for each $x \in N$ there exists an open neighborhood $V \subset N$ of x, and a smooth map $s : V \to M$, called a local section of F, such that:*

$$F \circ s = Id_V$$

Proof. Let $x \in N$ be arbitrary, and choose a point $p \in F^{-1}(x)$. Then using the coordinate charts from **Theorem 1.1.2**, we have that the coordinate representation of F, F^c is given by:

$$F^c(x^1, \ldots, x^{n+k}) = (x^1, \ldots, x^n)$$

Any smooth map $s : U \to M$, with the coordinate representation:

$$s^c(x^1, \ldots, x^n) = (x^1, \ldots, x^n, s_1(x^1, \ldots, x^n), \ldots, s_k(x^1, \ldots x^n))$$

where $s_i \in C^\infty(U)$, then satisfies the criteria. $\square$

Lemma 1.2.10. *Let $F : M \to N$ be a surjective submersion. Then a map $G : N \to Q$ is smooth if and only if $G \circ F : M \to Q$ is smooth. Furthermore, G is a submersion if and only if $G \circ F$ is a submersion, and G is surjective if and only if $G \circ F$ is surjective.*

Proof. If G is smooth, then $G \circ F$ is smooth. Conversely, assume that $G \circ F$ is smooth. By **Lemma 1.2.9**, for any $x \in N$ there exists an open neighborhood $V \subset N$ and a smooth local section of F, $s : V \to M$. On V we have that $F \circ s = \mathrm{Id}_V$, so:

$$(G \circ F) \circ s = G$$

Hence G must be smooth on V as it is the composition of smooth maps. Since smoothness is a local criterion, and this holds for arbitrary $x \in N$, we have that G is smooth on all of N.

G being surjective and a submersion is clear from the properties of F. $\square$

Theorem 1.2.4. *Let Φ be a principal right action of G on M. Then M/G has the unique structure of a smooth manifold such that $\pi : M \to M/G$ is a smooth submersion.*

Proof. Since the action is free, we have that Ψ is injective. Indeed, let $(q, h), (p, g) \in M \times G$, such that:

$$(p, p \cdot g) = (q, q \cdot h)$$

then $p = q$, so:

$$(p, p \cdot g) = (p, p \cdot h)$$

This then implies that $\phi_p(g) = \phi_p(h)$, where ϕ_p is the orbit map. However, the orbit map is injective for all $p \in M$ so $g = h$, thus Ψ is injective. We want to show that this map is an immersion; by **Proposition 1.2.17**, the differential of Ψ is given by:

$$D_{(p,g)}\Psi(X, Y) = (X, D_pR_g(X) + \widetilde{\mu_G(Y)}_{p \cdot g})$$

for $(p,g) \in M \times G$, and $(X,Y) \in T_pM \oplus T_gG$. If $D_{(p,g)}(X,Y) = 0$, then $X = 0$, which implies that $\mu_G(Y)$ must be zero, hence $Y = 0$, thus the kernel of the differential is trivial for all $(p,g) \in M \times G$, so Ψ is an immersion.

Since Ψ is injective, it's also a bijection onto it's image, $R \subset M \times M$. Let $\Psi^{-1} : R \to M \times G$ be the inverse of Ψ. This map is continuous, since for any closed subset $U \subset M \times G$, we have that $\Psi(U)$ is closed in $M \times M$. Since $\Psi : M \times G \to R$ is continuous with continuous inverse, it is a homeomorphism, and thus and embedding. R is then a closed embedded manifold of $M \times M$.

Finally, we see that Ψ is a submersion onto it's image R, and that the composition:

$$\pi_1|_R \circ \Psi : M \times G \to M$$

is just the map $\pi_1 : M \times G \to M$, and thus a submersion. From **Lemma 1.2.10** we see that $\pi_1|_R$ must then be submersion as well. Godement's Theorem, **Theorem 1.2.3**, then proves the claim. □

Corollary 1.2.6. *Suppose that Φ is a free and principal right action of G on M. Then:*

$$\dim(M/G) = \dim M - \dim G$$

In particular, the kernel of the differential:

$$D_p\pi : T_pM \longrightarrow T_{[p]}M/G$$

at a point $p \in M$ is equal to the tangent space $T_p\mathcal{O}_p$ of the G-orbit through p.

Proof. Since π is a submersion, every $[p] \in M/G$ is a regular value value of π. By **Theorem 1.1.1**, this implies that $\pi^{-1}([p]) = \mathcal{O}_p$ is an embedded submanifold of M. Therefore, $\phi_p : G \to M$, is an embedding, and hence a homeomorphism onto it's image. This implies that $\dim \mathcal{O}_p = \dim G$, so:

$$\dim G = \dim M - \dim(M/G) \Rightarrow \dim(M/G) = \dim M - \dim G$$

as desired. Furthermore, take any curve in $\mathcal{O}_p$ such that $\gamma(0) = p$, and $\dot\gamma(0) = X \in T_p\mathcal{O}_p \subset T_pM$, then we see that:

$$\begin{aligned} D_p\pi(X) =& \frac{d}{dt}\Big|_{t=0} \pi(\gamma(t)) \\ =& \frac{d}{dt}\Big|_{t=0} [p] \\ =& 0 \end{aligned}$$

hence $T_p\mathcal{O}_p \subset \ker(D_p\pi)$, but these two vector spaces have the same dimension, so $T_p\mathcal{O}_p = \ker(D_p\pi)$, as desired. □

Finally, we see that for a free right action Φ of G on M, the preimage of the map $\Psi^{-1}(p,q)$ is either empty or a singleton set, and thus compact. **Lemma 1.2.8** then implies that principal right actions are proper right actions, and then **Corollary 1.2.4** shows that free proper actions are principal right actions, hence the two are equivalent. This gives us our final corollaries:

Corollary 1.2.7. *Suppose that Φ is right action of G on a smooth manifold M. The action is principal if and only if it is free and proper.*

Corollary 1.2.8. *Suppose that Φ is a free right action of G on a smooth manifold M. If G compact, Φ is principal.*

The preceding corollary demonstrates that if G is a compact Lie group acting freely on a smooth manifold M, then M/G is a smooth manifold as well.

Example 1.2.15. Denote by $\mathbb{S}^3$ the three sphere. It is easy to see that $\mathbb{S}^1 = U(1)$ and is thus a Lie group. We would like to show that $\mathbb{S}^3/\mathbb{S}^1$ is a quotient manifold diffeomorphic $\mathbb{S}^2$. This specific example is often called the Hopf fibration. First note that:

$$\mathbb{S}^3 = \{(z_1, z_2) \in \mathbb{C}^2 : z_1\bar{z}_1 + z_2\bar{z}_2 = |z_1|^2 + |z_2|^2 = 1\}$$

and that:

$$\mathbb{S}^1 = \{w \in \mathbb{C} : w\bar{w} = 1\}$$

implying that $w = e^{i\theta}$ for some $\theta \in [0, 2\pi)$. Hence, we define a right action on $\mathbb{S}^3$ by:

$$\begin{aligned} \mathbb{S}^3 \times \mathbb{S}^1 &\longrightarrow \mathbb{S}^3 \\ (z_1, z_2, e^{i\theta}) &\longmapsto (z_1 e^{i\theta}, z_2 e^{i\theta}) \end{aligned}$$

The action above is smooth as it is just multiplication in $\mathbb{C}$ in both coordinates, and has image in $\mathbb{S}^3$ as:

$$\begin{aligned} z_1 e^{i\theta}(z_1 e^{i\theta})^* + z_2 e^{i\theta}(z_2 e^{i\theta})^* =& z_1 e^{i\theta}\bar{z}_1 e^{-i\theta} + z_2 e^{i\theta}\bar{z}_2 e^{-i\theta} \\ =& z_1\bar{z}_1 + z_2\bar{z}_2 \\ =& 1 \end{aligned}$$

Since multiplication in $\mathbb{C}$ is abelian, and $\mathbb{S}^1$ is abelian, it follows that the action can be viewed as either right or left. Furthermore, the action is free, as for arbitrary $w, u \in S^1$, and $z \in \mathbb{C}$ we have that if:

$$z \cdot w = z \cdot u$$

then:

$$\bar{z}z \cdot w = \bar{z}z \cdot u$$

Since $\bar{z}z$ is just a real number, we can divide out on both sides to obtain:

$$w = u$$

Hence if ϕ_p is the orbit map for some $p = (z_1, z_2) \in \mathbb{S}^3$ we have that if:

$$\phi_p(u) = \phi_p(v)$$

for some $u, w \in \mathbb{S}^1$, then:

$$(z_1 \cdot u, z_2 \cdot u) = (z_1 \cdot w, z_2 \cdot w)$$

so $u = w$, and the action is free. Since $\mathbb{S}^1$ is compact, it then follows from **Corollary 1.2.8** that the action of $\mathbb{S}^1$ on $\mathbb{S}^3$ is principal, so $\mathbb{S}^3/\mathbb{S}^1$ is a smooth manifold. We need to show that $\mathbb{S}^3/\mathbb{S}^1$ is diffeomorphic to $\mathbb{S}^2$. By identifying $\mathbb{C}$ with $\mathbb{R}^2$ we see that:

$$\mathbb{S}^2 = \{(w, x) \in \mathbb{C} \times \mathbb{R} : |w|^2 + x^2 = 1\}$$

thus motivating the map:

$$\begin{aligned} F : \mathbb{S}^3/\mathbb{S}^1 &\longrightarrow \mathbb{S}^2 \\ [(z_1, z_2)] &\longmapsto (2z_1\bar{z}_2, 2|z_1|^2 - 1) \in \mathbb{C} \times \mathbb{R} \end{aligned}$$

The map is clearly well defined, and has image in $\mathbb{S}^2$ as:

$$\begin{aligned} |2z_1\bar{z}_2|^2 + \left(2|z_1|^2 - 1\right)^2 =& 4|z_1|^2|z_2|^2 + 4|z_1|^4 - 4|z_1|^2 + 1 \\ =& 4|z_1|^2(1 - |z_1|^2) + 4|z_1|^4 - 4|z_1|^2 + 1 \\ =& 1 \end{aligned}$$

Furthermore, the map is smooth as it consists of conjugation and multiplication in $\mathbb{C}$. We check that the map is injective. Let $[z_1, z_2], [w_1, w_2] \in \mathbb{S}^3/\mathbb{S}^1$, then:

$$2|z_1|^2 - 1 = 2|w_1|^2 - 1 \Longrightarrow |z_1|^2 = |w_1|^2 \Longrightarrow z_1 = w_1 e^{i\theta}$$

for some $\theta \in \mathbb{R}$. Therefore:

$$z_1\bar{z}_2 = w_1\bar{w}_2 \Longrightarrow z_1\bar{z}_2 = e^{-i\theta}z_1\bar{w}_2 \Longrightarrow \bar{z}_2 = e^{-i\theta}\bar{w}_2 \Longrightarrow z_2 = e^{i\theta}w_2$$

Thus if $F([z_1, z_2]) = F([w_1, w_2])$ we have that $[z_1, z_2] = [w_1, w_2]$ as $(z_1, z_2) = (w_1 e^{i\theta}, w_2 e^{i\theta})$ so F is injective. Furthermore, if $(w, y) \in \mathbb{S}^2 \subset \mathbb{C} \times \mathbb{R}$, we see that[15]:

$$(z_1, z_2) = \left(i\sqrt{\frac{1}{2}(1-y)}, \frac{1}{2} \frac{i\bar{w}}{\sqrt{\frac{1}{2}(1-y)}} \right)$$

satisfies:

$$F([z_1, z_2]) = (w, y)$$

as:

$$2|z_1|^2 - 1 = (1+y) - 1 = y$$

and:

$$2z_1\bar{z}_2 = \sqrt{\frac{1}{2}(1-y)} \cdot \frac{w}{\sqrt{\frac{1}{2}(1-y)}} = w$$

Furthermore, $(z_1, z_2) \in \mathbb{S}^3$ as:

$$\begin{aligned} |z_1|^2 + |z_2|^2 =& \frac{1}{2}(1-y) + \frac{1}{4} \frac{|w|^2}{\frac{1}{2}(1-y)} \\ =& \frac{1}{2}(1-y) + \frac{1}{2} \frac{(1-y)(1+y)}{(1-y)} \\ =& \frac{1}{2}(1-y) + \frac{1}{2}(1+y) \\ =& 1 \end{aligned}$$

This then implies that F is surjective, and thus a smooth bijection.

We now wish to calculate the differential $D_p F$ and show it is an isomorphism of tangent spaces. Let $\gamma : I \to \mathbb{S}^3/\mathbb{S}^1$ be a smooth curve, then for all $t \in I$ we have that:

$$\gamma(t) = [z_1(t), z_2(t)]$$

for curves z_1, z_2 in $\mathbb{C}$ satisfying:

$$|z_1(t)|^2 + |z_2(t)|^2 = 1$$

Then:

$$\dot{\gamma}(t) = D_{z_1(t), z_2(t)}\pi(\dot{z}_1(t), \dot{z}_2(t))$$

which we denote by $[\dot{z}_1(t), \dot{z}_2(t)]$. Note that if $\gamma(t)$ is constant, then:

$$\gamma(t) = (z_1, z_2) \cdot e^{if(t)}$$

for some real valued function $f : I \to \mathbb{R}$. This implies that the kernel of $D_{z_1(t), z_2(t)}\pi$ is given by:

$$\ker D_{z_1(t), z_2(t)}\pi = \{(X_1, X_2) \in T_{z_1, z_2}\mathbb{S}^3 : \exists y \in \mathbb{R}, (X_1, X_2) = (iyz_1, iyz_2)\}$$

where:

$$T_{z_1, z_2}\mathbb{S}^3 = \{(X_1, X_2) \in \mathbb{C}^2 : \mathrm{Re}(\bar{z}_1 X_1 + \bar{z}_2 X_2) = 0\}$$

It is important to note that these are both real vector spaces/subspaces. Let $\dot{z}_1(0) = X_1$ and $\dot{z}_2(0) = X_2$ with $z_1(0) = z_1$, and $z_2(0) = z_2$, then we see that:

$$\left.\frac{d}{dt}\right|_{t=0} 2z_1(t)\bar{z}_2(t) = 2(X_1\bar{z}_2 + z_1\bar{X}_2) \tag{1.2.12}$$

[15]This choice fails when $y = 1$, but that is ok, as if $y = 1$, then $w = 0$, and $(1, 0) \in \mathbb{S}^3$ maps to $(0, 1)$.

and that:

$$\left.\frac{d}{dt}\right|_{t=0} 2|z_1(t)|^2 - 1 = 2(X_1\bar{z}_1 + \bar{X}_1 z_1) \tag{1.2.13}$$

So the differential is given by:

$$D_{[z_1,z_2]}F([X_1,X_2]) = 2(X_1\bar{z}_2 + z_1\bar{X}_2, X_1\bar{z}_1 + \bar{X}_1 z_1)$$

It suffices to show that the map is injective as $T_p\mathbb{S}^3/\mathbb{S}^2$ and $T_{F(p)}\mathbb{S}^2$ have the same dimension. In this case, that is equivalent to checking that $D_{[z_1,z_2]}F$ is zero only when $(X_1,X_2) \in \ker D_{(z_1,z_2)}\pi$. It is easy to that if $(X_1,X_2) \in \ker D_{(z_1,z_2)}\pi$ that $D_{[z_1,z_2]}F$ is zero. Let $(X_1,X_2) \in T_{z_1,z_2}\mathbb{S}^3$ for both $z_1, z_2 \neq 0$, and suppose that $D_{[z_1,z_2]}F[X_1,X_2] = 0$. Then setting equation (1.2.13) to zero gives:

$$\begin{aligned} X_1\bar{z}_1 &= -\bar{X}_1 z_1 \\ &= -(X_1\bar{z}_1)^* \end{aligned}$$

so $X_1\bar{z}_1$ is purely imaginary, and thus can be written as:

$$X_1\bar{z}_1 = ia$$

for some $a \in \mathbb{R}$. Now let:

$$y = \frac{a}{|z_1|^2}$$

for some $y \in \mathbb{R}$, thus X_1 can be written as:

$$X_1 = iyz_1 \tag{1.2.14}$$

since:

$$X_1\bar{z}_1 = iyz_1\bar{z}_1 = i\frac{a}{|z_1|^2}|z_1|^2 = ia$$

Using the fact that $(X_1,X_2) \in \mathbb{S}^3$ we have that:

$$\mathrm{Re}(\bar{z}_1X_1 + \bar{z}_2X_2) = 0 \implies \bar{z}_1X_1 + z_1\bar{X}_1 + \bar{z}_2X_2 + z_2\bar{X}_2 = 0$$

By (1.2.14) we have that this reduces to:

$$\bar{z}_2X_2 + z_2\bar{X}_2 = 0$$

Then the same argument used for X_1 demonstrates that:

$$X_2 = ivz_2$$

for some $v \in \mathbb{R}$. Finally setting (1.2.12) to zero gives:

$$\begin{aligned} 0 =& X_1\bar{z}_2 + z_1\bar{X}_2 \\ =& iyz_1\bar{z}_2 - ivz_1\bar{z}_2 \\ =& iz_1\bar{z}_2(v-y) \end{aligned}$$

implying that $v = y$, hence:

$$(X_1,X_2) = (iyz_1, iyz_2)$$

for some $y \in \mathbb{R}$, so $(X_1,X_2) \in \ker D_{z_1,z_2}\pi$, and $D_{z_1,z_2}F$ is injective for all $z_1, z_2 \neq 0$. Furthermore, if $z_1 = 0$, then the differential is given by:

$$D_{[0,z_2]}F([X_1,X_2]) = 2(X_1\bar{z}_2, 0)$$

which is only zero if X_1 is zero. We then obtain that:

$$\mathrm{Re}(X_2\bar{z}_2) = 0 \implies X_2 = iyz_2$$

for some $y \in \mathbb{R}$, so again $(X_1,X_2) \in \ker D_{z_1,z_2}\pi$, hence $D_{[0,z_1]}F$ is injective. Finally, if $z_2 = 0$, then we see that:

$$D_{[z_1,0]}F([X_1,X_2]) = 2(z_1\bar{X}_2, X_1\bar{z}_1 + \bar{X}_1 z_1)$$

which is only zero if X_2 is zero, and, as shown earlier, $X_1 = iyz_1$ for some $y \in \mathbb{R}$ hence $(X_1,X_2) \in \ker D_{z_1,z_2}\pi$ and $D_{[0,z_1]}F$ is injective. This then implies that D_pF is injective for all $p \in \mathbb{S}^3/\mathbb{S}^1$, thus D_pF is an isomorphism for all p, so F is a diffeomorphism. Therefore, $\mathbb{S}^3/\mathbb{S}^1 \cong \mathbb{S}^2$ as desired.

1.2.6 Representations

In **1.2.4**, we initially motivated group actions on manifolds by briefly mentioning representations; we are now ready to study this concept in detail. We will be particularly interested in representations of G on its Lie algebra $\mathfrak{g}$, and the existence of G-invariant scalar products on V. These ingredients will prove important for our work on Yang-Mills, as they are needed to construct the Yang-Mills action.

Definition 1.2.21. Let V be an n-dimensional $\mathbb{K}$-linear vector space for some field $\mathbb{K}$, and G a Lie group. A **representation of G on V** is a Lie group homomorphism:

$$\rho : G \longrightarrow GL(V)$$

where $GL(V)$ is isomorphic to $GL_n(\mathbb{K})$, where $\mathbb{K}$ is the field

If the representation is clear, we sometimes write:

$$\rho(g)(v) = \rho(g) \cdot v = g \cdot v = gv$$

for $g \in G$ and $v \in V$. Furthermore, by the properties of the a Lie group homomorphism, the following identities are automatic:

$$\begin{aligned} \rho(g \cdot h)(v) = \rho(g) \circ \rho(h)(v) = \rho(g) \circ \rho(h)v \\ \rho(g^{-1})(v) = \rho(g)^{-1}(v) = \rho(g)^{-1}v \end{aligned}$$

for all $g, h \in G$, and all $v \in V$.

Definition 1.2.22. A representation is called **faithful** if ρ is injective.

We have a similar definition for Lie algebras:

Definition 1.2.23. Let V be an n-dimensional $\mathbb{K}-$linear vector space for some field $\mathbb{K}$, and $\mathfrak{g}$ a Lie algebra. A **representation of $\mathfrak{g}$ on V** is a Lie algebra homomorphism:

$$\phi : \mathfrak{g} \longrightarrow \mathfrak{gl}(V) = \mathrm{End}(V)$$

In particular, a Lie algebra representation is a linear map such that:

$$\phi\left([X, Y]\right) = [\phi(X), \phi(Y)]$$

for all $X, Y \in \mathfrak{g}$.

Proposition 1.2.18. *Let $\rho : G \to GL(V)$ a representation of a Lie group G on some vector space V. Then the differential $\rho_* : \mathfrak{g} \to End(V)$ is a representation of the Lie algebra $\mathfrak{g}$*

Proof. This follows trivially from **Proposition 1.2.5**. □

Definition 1.2.24. Let $\rho : G \to GL(V)$ be a representation of G on a vector space V equipped with a (pseudo)-Euclidean scalar product of signature (t, s), $\langle \cdot, \cdot \rangle$. We say that ρ is **(pseudo)-orthogonal** if for all $v, w \in V$, and all $g \in G$:

$$\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle$$

Equivalently, the image of ρ lies in $O(t, s)$.

Theorem 1.2.5. *Let ρ be a representation of G on $V \cong \mathbb{R}^n$ equipped with a Euclidean scalar product, $\langle , \rangle$. Then, if G is compact, there exists another Euclidean scalar product on V such that ρ is an orthogonal representation with respect to the new scalar product.*

Proof. Let G have dimension n, and let $\{X_i, \ldots, X_n\}$ be a basis for T_eG, then and let $\{\tilde{\omega}^i, \ldots, \tilde{\omega}^n\}$ denote its dual basis. Then we obtain a set of right invariant one forms on G defined for all $g \in G$ via:

$$\omega^i_g(v) = R^*_{g^{-1}}\tilde{\omega}(v) = \tilde{\omega}(R_{g^{-1}*}v)$$

where $v \in T_gG$. It is clear that $\{\omega^1, \dots, \omega^n\}$ is a global frame of right invariant one forms, hence, the top form:

$$\sigma = \omega^1 \wedge \cdots \wedge \omega^n$$

is globally defined, and nowhere vanishing; in other words, σ is an orientation form on G. We assume that orientation on G agrees with σ[16], so that:

$$\int_G \sigma > 0$$

We define a smooth $\tau_{v,w} : G \to \mathbb{R}$, by:

$$\tau_{v,w}(g) = \langle \rho(g)v, \rho(g)w \rangle$$

and use $\tau_{v,w}$ to define a new scalar product on V by:

$$\eta(v, w) = \int_G \tau_{v,w}\sigma$$

for all $v, w \in V$. Note that η is finite, since G is compact

We first show that this indeed a Euclidean scalar product. Clearly, η is symmetric and bilinear because $\langle \cdot, \cdot \rangle$ is symmetric and bilinear, and integration is linear. Furthermore, note that $\langle v, v \rangle > 0$ for all non zero $v \in V$, so for any non zero $v \in V$, $\tau_{v,v}\sigma$ is a positively oriented orientation form. Thus, by **Theorem 1.1.5**:

$$\eta(v, v) = \int_G \tau_{v,v}\sigma > 0$$

implying that η is positive definite. Therefore, η is a Euclidean scalar product.

We now check that ρ is an orthogonal representation with respect to η. Let $g \in G$ be fixed, then we see that:

$$\begin{aligned} R^*_{g^{-1}}\tau_{\rho(g)v,\rho(g)w}(h) =& \tau_{\rho(g)v,\rho(g)w}(hg^{-1}) \\ =& \langle \rho(hg^{-1})\rho(g)v, \rho(hg^{-1})\rho(g)w \rangle \\ =& \langle \rho(h)v, \rho(h)w \rangle \\ =& \tau_{v,w}(h) \end{aligned}$$

where we have used the fact that ρ is a homomorphism. Since σ is right invariant, we then obtain that:

$$R^*_{g^{-1}}(\tau_{\rho(g)v,\rho(g)w}\sigma) = \tau_{v,w}\sigma$$

Then, since $R^*_{g^{-1}}$ is an orientation preserving isomorphism, by **Theorem 1.1.5** we see that for all $g \in G$, and all $v, w \in V$:

$$\begin{aligned} \eta(\rho(g)v, \rho(g)w) =& \int_G \tau_{\rho(g)v,\rho(g)w}\sigma \\ =& \int_G R^*_{g^{-1}}(\tau_{\rho(g)v,\rho(g)w}\sigma) \\ =& \int_G \tau_{v,w}\sigma \\ =& \eta(v, w) \end{aligned}$$

so ρ is orthogonal with respect to η. □

It is important to note that the same process will not in general work if V is equipped with a pseudo Euclidean scalar product. Indeed, let $V = \mathbb{R}^2$, and equip V with the pseudo Euclidean scalar product:

$$\eta(x, y) = x^1y^1 - x^2y^2$$

[16] If it does not, just swap ω^1 with ω^2 to pick up a minus sign.

Then, let $G = \mathbb{Z}/2\mathbb{Z}$ and let $\rho : \mathbb{Z}/2\mathbb{Z} \to GL(V)$ be the representation given by:

$$\rho(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \rho(1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The new scalar product would be given by:

$$\begin{aligned} \omega(x, y) =& \eta(\rho(0)x, \rho(0)y) + \eta(\rho(1)x, \rho(1)y) \\ =& x^1y^1 - x^2y^2 + x^2y^2 - x^1y^1 \\ =& 0 \end{aligned}$$

hence ω is identically zero on $\mathbb{R}^2$, and thus clearly degenerate, which is less than ideal.

Proposition 1.2.19. *Let V be an $\mathbb{R}$-linear vector space, equipped with a Euclidean scalar product $\langle \cdot, \cdot \rangle$, and $\rho : G \to GL(V)$ be a (pseudo) orthogonal representation. Then $\rho_* : \mathfrak{g} \to End(V)$ satisfies:*

$$\langle \rho_* X v, w \rangle + \langle v, \rho_* X w \rangle = 0$$

for all $X \in \mathfrak{g}$, and all $v, w \in V$.

Proof. From **Proposition 1.2.9**, we have that for all $X \in \mathfrak{g}$:

$$\rho(\exp(tX)) = \exp(t\rho_* X)$$

hence, for all $v, w \in V$:

$$\begin{aligned} \langle v, w \rangle =& \langle \rho(\exp(tX))v, \rho(\exp(tX))w \rangle \\ =& \langle \exp(t\rho_* X)v, \exp(t\rho_* X)w \rangle \end{aligned}$$

Differentiating at $t = 0$ we obtain that:

$$\langle \rho_* X v, w \rangle + \langle v, \rho_* X w \rangle = 0$$

implying the claim. □

We now turn to developing a very special type of representation, that is a representation of G on it's own Lie algebra. Recall the conjugation map:

$$\begin{aligned} c_g : G &\longrightarrow G \\ h &\longmapsto ghg^{-1} \end{aligned}$$

we can also right:

$$c_g = L_g \circ R_{g^{-1}}$$

This map is clearly a diffeomorphism, as it has a smooth inverse given by $c_{g^{-1}}$. Furthermore, the map is a homomorphism, as for all $h_i \in G$, we have that:

$$c_g(h_1) \cdot c_g(h_2) = gh_1g^{-1}gh_2g^{-1} = gh_1h_2g^{-1} = c_g(h_1h_2)$$

so c_g is a Lie group isomorphism, and thus $c_{g*} : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra isomorphism, motivating our next theorem.

Theorem 1.2.6. *The map:*

$$\begin{aligned} Ad : G &\longrightarrow GL(\mathfrak{g}) = Aut(\mathfrak{g}) \\ g &\longrightarrow Ad_g = c_{g*} \end{aligned}$$

*is a representation of G on $\mathfrak{g}$, called the **adjoint representation**.*

Proof. Let $g, h \in G$, then:

$$\begin{aligned} c_g \circ c_h =& \left(L_g \circ R_{g^{-1}}\right) \circ \left(L_h \circ R_{h^{-1}}\right) \\ =& L_g \circ (R_{g^{-1}} \circ R_{h^{-1}}) \circ L_h \\ =& L_g \circ R_{(gh)^{-1}} L_h \\ =& L_{gh} \circ R_{(hg)^{-1}} \\ =& c_{gh} \end{aligned}$$

hence:

$$\begin{aligned} \mathrm{Ad}_g \circ \mathrm{Ad}_h =& c_{g*} \circ c_{h*} \\ =& (c_g \circ c_h)_* \\ =& \mathrm{Ad}_{gh} \end{aligned}$$

so Ad is a homomorphism. We now need to show that Ad is a smooth map. We need only show that the map:

$$\mathrm{Ad}_{(\bullet)}(v) : G \longrightarrow \mathfrak{g}$$

is smooth for all $v \in \mathfrak{g}$, because if we choose a basis for $\mathfrak{g}$ then Ad is a smooth matrix representation into $GL_{\dim \mathfrak{g}}(\mathbb{R})$. We see the composition of smooth maps:

$$\begin{aligned} \mathrm{Ad}_{(\bullet)} v :& G \longrightarrow TG \times TG \longrightarrow T(G \times G) \longrightarrow TG \\ :& g \longmapsto ((g,0),(e,v)) \longmapsto ((g,e),(0,v)) \longmapsto D_{(g,e)} c(0,v) \end{aligned}$$

where:

$$\begin{aligned} c : G \times G &\longrightarrow G \\ (g,h) &\longmapsto ghg^{-1} \end{aligned}$$

is equal to $\mathrm{Ad}_{(\bullet)} v$ since for all $g \in G$:

$$\begin{aligned} D_{(g,e)} c(0,v) =& \frac{d}{dt}\Big|_{t=0} g \exp(tv) g^{-1} \\ =& gvg^{-1} \\ =& \mathrm{Ad}_g(v) \end{aligned}$$

Therefore, $\mathrm{Ad}_{(\bullet)}(v)$ is smooth for all $v \in V$, so the map Ad is smooth as well. Since Ad is smooth, and a homomorphism, it follows that Ad is a Lie group homomorphism $G \to GL(\mathfrak{g})$, and thus a representation of G on $\mathfrak{g}$ as desired. ☐

We can go a step further, and obtain an induced representation of the Lie algebra on itself as the next theorem shows.

Theorem 1.2.7. *Let G be a Lie group. Then the induced representation of Ad is the map:*

$$\begin{aligned} ad :& \mathfrak{g} \longrightarrow End(\mathfrak{g}) \\ & X \longrightarrow ad_X = Ad_*(X) \end{aligned}$$

and satisfies:

$$ad_X Y = [X, Y]$$

Proof. Let $X, Y \in \mathfrak{g}$, then:

$$\begin{aligned} \mathrm{ad}_X Y =& \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(tX)} Y_e \\ =& \frac{d}{dt}\Big|_{t=0} c_{\exp(tx)*} Y_e \\ =& \frac{d}{dt}\Big|_{t=0} L_{\exp(tX)*} R_{\exp(-tX)*} Y_e \\ =& \frac{d}{dt}\Big|_{t=0} R_{\exp(-tX)*} Y_{\exp(tX)} \end{aligned}$$

However, we see that:

$$\begin{aligned} R_{\exp(-tX)*}Y_{\exp(tX)} =& D_{\exp(tX)}R_{\exp(-tx)}(Y_{\exp(tx)}) \\ =& D_{\theta_t(e)}\theta_{-t}(Y(\theta_t(e))) \end{aligned} \tag{1.2.15}$$

since the flow of X, θ_t satisfies:

$$\theta(t,g) = \theta_t(g) = R_{\exp(tX)}g$$

With (1.2.15) in mind, we obtain:

$$\begin{aligned} \mathrm{ad}_X Y =& \frac{d}{dt}\Big|_{t=0} D_{\theta_t(e)}\theta_{-t}(Y(\theta_t(e))) \\ =& (\mathscr{L}_X Y)_e \\ =& [X,Y]_e \end{aligned}$$

as desired. $\square$

This theorem gives us the following corollary:

Corollary 1.2.9. *If G is abelian, then $\mathfrak{g}$ is abelian, i.e. for all $X, Y \in \mathfrak{g}$:*

$$[X,Y] = 0$$

Proof. If G is abelian then for all $g, h \in G$ we have that:

$$c_g(h) = h$$

hence Ad is the identity map on $\mathfrak{g}$. Furthermore, we have that:

$$\begin{aligned} [X,Y]_e =& \mathrm{ad}_X Y \\ =& \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(tX)} Y_e \\ =& \frac{d}{dt}\Big|_{t=0} Y_e \\ =& 0 \end{aligned}$$

which implies the claim. $\square$

We will develop more representation theory as needed throughout this paper, but for now what we have is sufficient. We end our with the following corollary which will be vital for our development of the Yang-Mills action.

Corollary 1.2.10. *Let G be a compact Lie group, then there exist a Ad invariant Euclidean inner product on $\mathfrak{g}$.*

Proof. This follows from **Theorem 1.2.5**. $\square$

Gauge Theory and Spinors

2.1 Principal Bundles and Connections

We are now ready to put all the pieces together and set the stage for our work on gauge theory. We begin by introducing general fibre bundles, a type of manifold that can be thought of as locally 'looking like' a product manifold, an idea which should feel familiar. Indeed, we have seen such manifolds before from our work on tangent bundles, and various other tensor bundles, but in this section we generalize this idea. From there, we will dive into principal bundles, a specific type of fibre bundle with a Lie group action which is transitive and free on the fibres; these types of bundles will help us make it clear what one means by a gauge transformation. We will then study how one can obtain vector bundles from principal bundles, which will be necessary for our work with Yang-Mills and, in particular, for our formalism of classical electromagnetism.

Connections on principal bundles, and their curvatures, will then take centre stage as the objects of primary interest. These can be loosely thought of as generalizations of the Levi-Civita connection, and the Riemann curvature tensor, and we will see in a later chapter how to succinctly connect these ideas together. In particular, our connections, and curvatures can be viewed as Lie algebra valued one and two forms on the principal bundle respectively. However, more will be true, as we will be able to locally identify both of these objects as Lie algebra valued differential forms on the base manifold, allowing us to view them as fields on our spacetime. We will end the chapter by discussing covariant derivatives on associated vector bundles, which will provide us with the necessary tools to *vary* the Yang-Mills action.

We continue to follow closely Hamilton's *Mathematical Gauge Theory*, and, as always, a more complete discussion can be found there.

2.1.1 Fibre Bundles

We begin with the following definition:

Definition 2.1.1. Let $\pi : E \to M$ be a smooth surjection between manifolds.

a) For all $x \in M$, the subset $E_x \subset E$ defined by:

$$E_x = \pi^{-1}(x) = \{p \in E : \pi(p) = x\}$$

is called the **fibre** of π over x

b) For any subset $U \subset M$, we set:

$$E_U = \pi^{-1}(U) = \{p \in E : \pi(p) \in U\}$$

E_U can be thought of as the part of E 'sitting over' U, and is clearly the union of all the fibres E_x for $x \in U$.

c) A smooth map $s : M \to E$ satisfying:

$$\pi \circ s = \mathrm{Id}_M$$

is a **global section** of π. Furthermore, a smooth map $s : U \subset M \to E$ satisfying:

$$\pi \circ s = \mathrm{Id}_U$$

is a **local section** of π.

Recall our work with the tangent bundle for some n-dimensional manifold M. In this case, TM comes equipped with a smooth projection map $\pi : TM \to M$, and each fibre of π is the vector space $T_xM \cong \mathbb{R}^n$, and the smooth sections of π are vector fields. For general surjective maps however, it need not be case that each fibre is diffeomorphic to one another, i.e in general $E_x \neq E_y$ for $x \neq y$ With this in mind, we define fibre bundles as follows:

Definition 2.1.2. Let E, M, F be smooth manifolds, and $\pi : E \to M$ a smooth surjective map. The quadruplet $(E, \pi, M; F)$ is then a **fibre bundle** if the following holds: For each $x \in M$, there exists an open neighborhood $U \subset M$ around x such that π restricted to E_U can be trivialized, i.e. there exists a diffeomorphism:

$$\phi_U : E_U \to U \times F$$

such that:

$$\mathrm{pr}_U \circ \phi_U = \pi$$

where pr_U is the projection map $U \times F \to U$. We call:

- E the **total space**
- M the **base manifold**
- F the **general fibre**
- π the **projection**
- (U, ϕ_U) a **local trivialization** or **bundle charts**

We see that the preceding definition guarantees that for each $x, y \in M$, we have that $E_x \cong E_y$. Furthermore, the existence of bundle charts, or local trivializations makes clear what one means by the statement: fibre bundles locally 'look like' product manifolds. Clearly, $E = M \times F$, with $\pi = \mathrm{pr}_M$ is a fibre bundle, albeit the easiest example of one.

Definition 2.1.3. Let $(E, \pi, M; F)$, and $(E', \pi', M; F')$ be fibre bundles, then a **bundle map** or **bundle homomorphism** is a smooth map $\Phi : E \to E'$ satisfying:

$$\pi' \circ \Phi = \pi$$

A **bundle isomorphism** is a bundle map which is also a diffeomorphism, and if such a maps exists we write $E \cong E'$.

In other words, bundle maps leave the base manifold fix, but changes the fibres, and hence the total space. If $E \cong M \times F$ we say that E is a **trivial** bundle. We see for each local trivialization (U, ϕ_U), that ϕ_U is a bundle isomorphism $E_U \to U \times F$, so each E_U is trivial bundle over the base space U. Furthermore, with some algebraic topology, one can show that any fiber bundle over some contractible manifold M is trivial; in particular, any fibre bundle over $\mathbb{R}^n$ is trivial.

Proposition 2.1.1. *Let $(E, \pi, M; F)$ be a fibre bundle. Then the fibres, E_x, are embedded submanifolds for all $x \in M$.*

Proof. We know that each fibre E_x is a subset of some $E_U \cong U \times F$, where $U \subset M$ is an open neighborhood of x. Therefore, we have that for all $v \in T_p E_U$:

$$D_p\pi(v) = D_p(\mathrm{pr}_U \circ \phi_U)(v)$$

Let $q \in \pi^{-1}(x)$, and $w \in T_x M$ be arbitrary, we want to find a $v \in T_q E$ such that $D_q\pi(v) = w$. We see that $\phi_U(q) = (x, f)$ for some $f \in F$, hence:

$$D_q\phi_U(v) = (v_m, v_f) \in T_x U \oplus T_f F$$

for some $v_m \in T_x U$, and some $v_f \in T_f F$. Therefore:

$$\begin{aligned} D_q\pi(v) =& D_{x,f}\mathrm{pr}_U \circ D_q\phi_U(v) \\ =& v_m \end{aligned}$$

Since ϕ_U is a diffeomorphism, there exists some $v \in T_q E_U$ such that $D_q\phi_U(v) = (w, v_f)$, hence for this v:

$$D_q\pi(v) = w$$

So the map $D_q\pi$ is a surjection onto $T_x M$ for all $q \in \pi^{-1}(x)$, implying that x is a regular value of π. Thus, by **Theorem 1.1.1** $\pi^{-1}(x)$ is an embedded submanifold of E. □

In the process of proving the proceeding proposition, we have also shown that π is a submersion, as each point in each fibre is a regular point of π, thus $D_p\pi$ is surjective for all $p \in E$. This allows us to prove the following:

Proposition 2.1.2. *Let $(E, \pi, M; F)$ and $(E', \pi', M; F')$, a $\Phi : E \to E'$ a bundle homomorphism between them. Φ is a bundle isomorphism if and only if the restriction of Φ to the fibre E_x is a diffeomorphism for all $x \in M$.*

Proof. If Φ is a diffeomorphism then the restriction of Φ to E_x is a smooth bijective map, which we denote by Φ_x. Since $D_p\Phi$ is an isomorphism for all $p \in E$, it follows that $D_p\Phi_x$ is an isomorphism for all $p \in E_x$, then Φ_x is a diffeomorphism for all $x \in M$.

Now suppose that the restriction of Φ, $\Phi_x : E_x \to E'_x$ is a diffeomorphism for all $x \in M$. Suppose $p, p' \in E$ such that:

$$\Phi(p) = \Phi(p')$$

then p and p' must lie in the same fibre E_x for some $x \in M$. Therefore:

$$\Phi_x(p) = \Phi_x(p')$$

so $p = p'$, and Φ is injective. Furthermore, for $q \in E'_x$, such that $\pi'(q) = x$, we have that there exists a $p \in E_x$ such that $\Phi_x(p) = q$. It then follows that:

$$\Phi(p) = q$$

so Φ is surjective, and thus a smooth bijection.

Since Φ_x is a diffeomorphism, it follows that $\dim F = \dim F'$, and since E is locally diffeomorphic to $(U \subset M) \times F$ we have that:

$$\dim E = \dim M + \dim F = \dim M + \dim F' = \dim E'$$

Therefore, it suffices to check that the differential is injective at each point by rank nullity. Let $p \in E$ such that $\pi(p) = x$, then since π is a submersion, after choosing a bundle chart, T_pE splits as:

$$T_pE \cong T_xM \oplus T_pE_x$$

as $T_pE_x = \ker D_p\pi$, and $T_xM = \text{im } D_p\pi$. Similarly we have that:

$$T_pE' \cong T_xM \oplus T_pE'_x$$

Therefore, any $v \in T_pE$ can be written as $v_m + v_e$ for some $v_m \in T_xM$ and $v_e \in T_pE_x$. We see that:

$$D_p\Phi(v_e) = D_p\Phi_x(v_e) \in T_pE'_x$$

which is only zero when v_e is zero. Furthermore, we see that for nonzero v_m:

$$\begin{aligned} D_{\Phi(p)}\pi' \circ D_p\Phi(v_m) =& D_p(\pi' \circ \Phi)(v_m) \\ =& D_p\pi(v_m) \in T_xM \end{aligned}$$

which can't be zero by assumption, implying that $D_p\Phi(v_m) \notin \ker D_{\Phi(p)}\pi'$. However, if $D_p\Phi(v_m) = 0$ then $D_p\Phi(v_m) \in \ker D_{\Phi(p)}\pi'$, so $D_p\Phi(v_m) \neq 0$. Hence if:

$$D_p\Phi(v) = D_p\Phi(v_m) + D_p\Phi_x(v_e) = 0$$

we need both $D_p\Phi_x(v_e) = 0$ and $D_p\Phi(v_m) = 0$, which as just shown only holds if $v_m = 0$ and $v_e = 0$, i.e. if $v = 0$. Therefore, the kernel of $D_p\Phi$ is trivial for all $p \in E$, implying that Φ is a diffeomorphism and thus a bundle isomorphism as desired. $\square$

Importantly, the above proposition relies on the fact that Φ is a a smooth map as a priori. In general, the condition that Φ restricts to a diffeomorphism on each fibre is not enough to prove this claim without this underlying assumption.

Much like manifolds, fibre bundles come equipped with an atlas:

Definition 2.1.4. Let $(E, \pi, M; F)$ be a fibre bundle, and $\{U_i\}_{i \in I}$ an open covering for an M. A **bundle atlas** is then the aforementioned open cover of M, with a set of bundle charts:

$$\phi_i : E_{U_i} \longrightarrow U_i \times F$$

We denote the atlas by $\{(U_i, \phi_i)\}_{i \in I}$

Furthermore, for a bundle atlas we have transition functions from one local trivialization to the next:

Definition 2.1.5. Let $(U_i, \phi_i)_{i\in I}$ be a bundle atlas for the fibre bundle $(E, \pi, M; F)$ be a fibre bundle. If $U_i \cap U_j \neq \emptyset$, we define the transition functions as:

$$\phi_j \circ \phi_i^{-1}\Big|_{(U_i\cap U_j)\times F} : (U_i \cap U_j) \times F \longrightarrow (U_i \cap U_j)$$

For $U_i, U_j \subset M$ such that $U_i \cap U_j \neq \emptyset$, we denote the restriction of ϕ_i to the fibre $E_x \subset E_{U_i}$ by ϕ_{ix}. Then, for each $x \in U_i \cap U_j$, we see that:

$$\phi_{jx} \circ \phi_{ix}^{-1} : F \longrightarrow F$$

is a diffeomorphism. Therefore, we obtain the following maps:

$$\begin{aligned} \phi_{ij} : U_i \cap U_j &\longrightarrow \mathrm{Diff}(F) \\ x &\longmapsto \phi_{jx} \circ \phi_{ix}^{-1} \end{aligned}$$

which we also call transition functions.

Proposition 2.1.3. *The transition functions $\{\phi_{ij}\}_{ij\in I}$ satisfy the following equations:*

$$\phi_{ii}(x) = Id_F \text{ for } x \in U_i \tag{2.1.1}$$
$$\phi_{ji}(x) \circ \phi_{ij}(x) = Id_F \text{ for } x \in U_i \cap U_j \tag{2.1.2}$$
$$\phi_{ki}(x) \circ \phi_{jk}(x) \circ \phi_{ij}(x) = Id_F \text{ for } x \in U_i \cap U_j \cap U_k \tag{2.1.3}$$

(2.1.3) *is called the* ***cocycle condition***

Proof. (2.1.1) is trivial. For (2.1.2) we see that for $x \in U_i \cap U_j$:

$$\phi_{ij}(x) = \phi_{jx} \circ \phi_{ix}^{-1}$$

while:

$$\phi_{ji}(x) = \phi_{ix} \circ \phi_{jx}^{-1}$$

Hence:

$$\begin{aligned} \phi_{ji}(x) \circ \phi_{ij}(x) =& (\phi_{ix} \circ \phi_{jx}^{-1}) \circ (\phi_{jx} \circ \phi_{ix}^{-1}) \\ =& \phi_{ix} \circ \phi_{ix}^{-1} \\ =& \mathrm{Id}_F \end{aligned}$$

as desired. For (2.1.3), we have that for $x \in U_i \cap U_j \cap U_k$:

$$\begin{aligned} \phi_{ki}(x) \circ \phi_{jk}(x) \circ \phi_{ij}(x) =& (\phi_{ix} \circ \phi_{kx}^{-1}) \circ (\phi_{kx} \circ \phi_{jx}^{-1}) \circ (\phi_{jx} \circ \phi_{ix}^{-1}) \\ =& \phi_{ix} \circ \phi_{ix}^{-1} \\ =& \mathrm{Id}_F \end{aligned}$$

so the cocycle conditions is satisfied. □

Before ending our discussion on general fibre bundles, recall our work with the tangent bundle: we began with a set, TM and a surjective map π, and then constructed a topology and smooth structure on TM such that it was a smooth manifold. In this process, we indirectly showed that TM is a fibre bundle; indeed, the coordinate charts for TM can be manipulated to yield bundle charts for TM. In a similar fashion, we would like to know when we can construct a fibre bundle out a set E, smooth manifolds M and F, and a surjective map $\pi : E \to M$. We need the following definition:

Definition 2.1.6. Let M and F be smooth manifolds, E a set, and $\pi : E \to M$ a surjective map.

a) Suppose $U \subset M$ is open and:

$$\phi_U : E_U \to U \times F$$

is a bijection with:

$$\mathrm{pr}_U \circ \phi_U = \pi|_{E_U}$$

then we call (U, ϕ) a **formal bundle chart** for E.

b) A collection of formal bundle charts $\{(U_i, \phi_i)\}_{i\in I}$ which is also an open cover for M is called a **formal bundle atlas**.

c) The charts in formal bundle atlas are **smooth compatible** if all transition functions:

$$\phi_j \circ \phi_i^{-1} : (U_i \cap U_j) \times F \to (U_i \cap U_j) \times F$$

are diffeomorphisms.

With this definition we end with the following theorem:

Theorem 2.1.1. *Let M and F be smooth manifolds, E a set, and $\pi : E \to M$ a surjective map. Suppose that $\{(U_i, \phi_i)\}_{i\in I}$ is a smoothly compatible formal bundle atlas. Then there exists a unique topology and smooth structure on E such that $(E, \pi, M; F)$ is a fibre bundle.*

Proof. The proof of this theorem is similar to **Proposition 1.1.3**, albeit with a few changes. We first define a basis for E by noting that M has a countable basis for it's topology $\{V_i\}_{i\in\mathbb{N}}$, and that each V_i in this basis is contained in a bundle chart by definition. Therefore, since each $V_i \subset U_j$ for some j, we define a new formal bundle atlas by taking $\{(V_i, \psi_i)\}_{i\in\mathbb{N}}$, where ψ_i is the restriction of ϕ_j to E_{V_i}. Now let $\{W_i\}_{i\in\mathbb{N}}$ be a countable basis for F, we construct a basis for the topology of E by:

$$\{\psi_i^{-1}(V_i \times W_j)\}_{i,j\in\mathbb{N}}$$

This basis is clearly countable, so E is second countable. Furthermore, let $p, q \in E$ such that $p \neq q$, if $\pi(q) = \pi(p) = x$, then $p, q \in E_x$, implying that $p, q \in E_{V_i}$ for some V_i. The topology on F is Hausdorff, so it follows that there exists disjoint open set $W_l, W_j \in \{W_i\}_{i\in\mathbb{N}}$, such that $p \in \psi^{-1}(V_i \times W_l)$, and $q \in \psi^{-1}(V_i \times W_j)$. If $\pi(p) \neq \pi(q)$, then, since the topology on M is Hausdorff there exist disjoint open sets V_i, V_j, such that $\pi(p) \in V_i$ and $\pi(q) \in V_j$. Thus for some non empty $W_k \in \{W_i\}_{i\in\mathbb{N}}$, $p \in \psi^{-1}(V_i \times W_k)$ and $q \in \psi^{-1}(V_j \times W_k)$, both of which are disjoint, so with this topology E is Hausdorff.

Let $O \subset U_i \times F$ be open for some $U_i \in \{U_i\}_{i\in I}$. Then there exists some open $V \subset U_i$ and $W \subset F$ such that:

$$O = V \times W$$

V must be the union of some subfamily of $\{V_i\}_{i\in\mathbb{N}}$ such that each $V_i \subset U_i$, and W must be the union of some subfamily of $\{W_i\}_{i\in\mathbb{N}}$, with this in mind we have:

$$O = \bigcup_{k\in K, j\in J} V_k \times W_j$$

where K and J are the indexing sets of the aforementioned subfamilies. Therefore, since each ψ_k is just ϕ_i restricted to V_k :

$$\phi_i^{-1}(O) = \bigcup_{k\in K, j\in J} \psi_k^{-1}(V_k \times W_j)$$

so ϕ_i^{-1} is a continuous bijection. We see that this map is also open as if $O \subset E_{U_i}$ is open then:

$$O = \bigcup_{k\in K, j\in J} \psi_k^{-1}(V_k \times W_j)$$

for some subfamilies of $\{V_i\}_{i\in\mathbb{N}}$, $\{W_i\}_{i\in\mathbb{N}}$ indexed by K and J respectively. Therefore, again since each ψ_k is just ϕ_i restricted to V_k :

$$\begin{aligned}\phi(O) &= \bigcup_{k\in K, j\in J} \phi_i(\psi_k^{-1}(V_k \times W_j)) \\ &= \bigcup_{k\in K, j\in J} V_k \times W_j\end{aligned}$$

which is open in $U_i \times F$. Thus, since ϕ_i is an open continuous bijection, it is a homeomorphism. Therefore, our basis defines a topology on E where our formal bundle atlas is a collection of homeomorphisms.

Let $p \in E$ such that $\pi(p) \in U_i$ for some $U_i \in \{U_i\}_{i\in I}$. Then there exists an open neighborhood of S such that $S \subset E_{U_i}$ and $\phi_i(S)$ is then open in $U_i \times F$. We can make S small enough such that $\phi(S)$ is contained in a coordinate chart for the product smooth manifold $U_i \times F$, hence E is locally Euclidean.

We have shown that with the topology defined above, E is Hausdorff, second countable, and locally Euclidean, and thus a topological manifold. It remains to be shown that E is a smooth manifold. For each $U_i \times F$, let $\{(O_{ij}, \psi_{ij})\}_{j\in J}$ be the product smooth atlas for the smooth manifold $U_i \times F$. Then, the collection:

$$\{E_{U_i}, \psi_{ij} \circ \phi_i\}_{i\in I, j\in J}$$

covers E, and is smoothly compatible since the transition functions:

$$\psi_{ij} \circ \phi_i \circ \phi_k^{-1} \circ \psi_{kl}^{-1} : \psi_{kl}(O_{kl} \cap O_{ij}) \longrightarrow \psi_{ij}(O_{kl} \cap O_{ij})$$

are smooth, as $\phi_i \circ \phi_k^{-1}$ is a diffeomorphism, and each ψ_{ij} is a local diffeomorphism by assumption. This clearly defines a smooth atlas for E such that each $\phi_i : E_{U_i} \to U_i \times F$ is a diffeomorphism, hence π is smooth as well, as locally π is the composition of $\mathrm{pr}_{U_i} \circ \phi_i$. Therefore, the quadruplet $(E, \pi, M; F)$ is a fibre bundle by **Definition 2.1.2** as desired. □

2.1.2 Principal Bundles

We are now in a position to define principal bundles. As mentioned earlier, these are essentially fibre bundles with a right group action of some Lie group G which acts transitively and freely on the fibres, though we will place extra restrictions on the bundle atlas. If G acts simply and transitively on the fibres, then this implies that the fibres are diffeomorphic to G. Importantly, the fibres will not be isomorphic to G in the sense of Lie groups.

Definition 2.1.7. Let G be a Lie group, M a a smooth manifold, and $(P, \pi, M; G)$ a fibre bundle with a smooth right action of G. P is called a **principal bundle** if:

a) The action of G preserves the fibres of P, and G is transitive and free on them. In other words, for all $x \in M$, the action of G restricts to:

$$P_x \times G \to P_x$$

and the orbit map:

$$\begin{aligned} G &\longrightarrow P_x \\ g &\longmapsto p \cdot g \end{aligned}$$

is a bijection.

b) The exists a bundle atlas $\{(U_i, \phi_i)\}_{i\in I}$ for P satisfying:

$$\phi_i(p \cdot g) = \phi_i(p) \cdot g \qquad \forall p \in P_{U_i}, g \in G$$

where G acts on the right hand side for $\phi_i(p) = (x, h) \in U_i \times G$ by:

$$(x, h) \cdot g = (x, hg)$$

such an atlas is called a **principal bundle atlas**.

We call G the **structure group** of P.

As the next proposition shows, the transition functions for a principal bundle atlas take values in $G \subset \mathrm{Diff}(G)$, and act on G by left multiplication.

Proposition 2.1.4. *Let P be a principal bundle over M with structure group G, and $\{(U_i, \phi_i)\}_{i\in I}$ a principal bundle atlas. The transition functions take values in $G \subset$ Diff(G):*

$$\begin{aligned}\phi_{ij} : U_i \cap U_j &\longrightarrow G \subset \mathit{Diff}(G)\\ x &\longmapsto \phi_{jx} \circ \phi_{ix}^{-1}\end{aligned}$$

where $g \in G$ acts as a diffeomorphism on G through left multiplication:

$$g(h) = g \cdot h$$

Proof. We first want to show that $\phi_{ix}^{-1} : G \to P_x$ is G-equivariant. By assumption we have that $\phi_{ix} : P_x \to G$ is a diffeomorphism hence for some $p \in P_x$, and some $g \in G$ we have that:

$$\phi_{ix}(p) = g$$

Furthermore, since ϕ_{ix} is G-equivariant, we have that for some $h \in G$:

$$\phi_{ix}(p \cdot h) = \phi_{ix}(p) \cdot h = g \cdot h$$

This implies that:

$$\phi_{ix}^{-1}(g) = p \qquad \text{and} \qquad \phi_{ix}^{-1}(g \cdot h) = p \cdot h$$

hence:

$$\phi_{ix}^{-1}(g \cdot h) = \phi_{ix}^{-1}(g) \cdot h$$

thus ϕ_{ix}^{-1} is G-equivariant as well. We now see that the transition function:

$$\begin{aligned}\phi_{ij} : U_i \cap U_j &\longrightarrow G \subset \mathrm{Diff}(G)\\ x &\longmapsto \phi_{jx} \circ \phi_{ix}^{-1}\end{aligned}$$

satisfy the following condition for all $x \in U_i \cap U_j$, and $g, h \in G$:

$$\begin{aligned}\phi_{ij}(x)(g \cdot h) =& \phi_{jx} \circ \phi_{ix}^{-1}(g \cdot h)\\ =& \phi_{jx}(\phi_{ix}^{-1}(g) \cdot h)\\ =& (\phi_{jx} \circ \phi_{ix}^{-1}(g)) \cdot h\\ =& \phi_{ij}(x)(g) \cdot h\end{aligned}$$

Therefore, for some $g \in G$, and some $x \in U_i \cap U_j$:

$$\phi_{ij}(x)(e) = g$$

hence for any $h \in G$:

$$\begin{aligned}\phi_{ij}(x)(h) =& \phi_{ij}(x)(e) \cdot h\\ =& g \cdot h\\ =& g(h)\end{aligned}$$

Thus the transition functions take values in $G \subset \mathrm{Diff}(G)$, and act on G via left multiplication. □

Proposition 2.1.5. *Let $\pi : P \to M$ be a smooth submersion between the smooth manifolds P and M. Furthermore, let G be a Lie group which acts on P from the right, and preserves the fibres of π, and is free and transitive on them. Then, P is a principal bundle.*

Proof. We need only show that we can construct a principal bundle atlas, as that will automatically imply that $(P, \pi, M; G)$ is a fibre bundle, and thus a principal bundle. As π is a smooth submersion, by **Lemma 1.2.8** there exists a covering $\{U_i\}_{i\in I}$ of M such that each U_i admits a local section $s : U_i \to P_{U_i}$ of π. Let $U \in \{U_i\}_{i\in I}$, and $s : U \to P_U$ be a local section. We now need to show that the map:

$$\begin{aligned} \psi : U \times G &\longrightarrow P_U \\ (x, g) &\longmapsto s(x) \cdot g \end{aligned}$$

is a G-equivariant diffeomorphism. It is smooth as the composition of smooth maps, and we see that:

$$\psi(x, gh) = s(x) \cdot (gh) = \psi(x, g) \cdot h$$

hence ψ is also G-equivariant.

For $(x, g), (y, h) \in U \times G$, we have that:

$$\psi(x, g) = \psi(y, h) \Longrightarrow s(x) \cdot g = s(y) \cdot h$$

The action of G preserves fibres, so $\psi(x, g)$ and $\psi(y, g)$ must be in the same fibre. However, by the same logic, this implies that $s(x)$ and $s(y)$ must lie in the fibre, i.e.:

$$\pi \circ s(x) = \pi \circ s(y)$$

but s is a local section, so $x = y$, and $s(x) = s(y)$. Furthermore, since G acts freely on the fibres we have that $g = h$ as well, hence ψ is injective. For $q \in P_U$, we have that $\pi(q) = x$ for some $x \in U$, and that $s \circ \pi(q) = p$ for some $p \in P_x$. Then, since the action of G is transitive on the fibres, we have that there exists a $g \in G$ such that:

$$\psi(x, g) = s(x) \cdot g = p \cdot g = q$$

hence ψ is surjective, thus ψ is a bijection.

Since π is a smooth submersion, we know that each P_x is an embedded submanifold of P with $\dim P_x = \dim G$. Furthermore, by **Theorem 1.1.1** we have that:

$$\dim G = \dim P - \dim M \Rightarrow \dim P = \dim M + \dim G$$

Since $U \times G$ has dimension $\dim M + \dim G$, and P_U is open in P, if we can show that the differential:

$$D_{(x,g)}\psi : T_xU \times T_gG \longrightarrow T_{(s(x)\cdot g)}P_U$$

has trivial kernel for all $x, g \in U \times G$, then by rank-nullity the differential will be an isomorphism, so ψ will be a diffeomorphism. We see that for $(X, Y) \in T_xU \times T_gG$, that the differential of ψ is given by **Proposition 1.2.17**:

$$D_{(x,g)}\psi(X, Y) = D_x(R_g \circ s)(X) + \widetilde{\mu_G(Y)}_{s(x)\cdot g}$$

We first see that:

$$R_g \circ s = s(x) \cdot g$$

is another another section as G preserves the fibres of π, hence:

$$\pi \circ (R_g \circ s) = \mathrm{Id}$$

So by the chain rule:

$$D_{s(x)}\pi \circ D_x(R_g \circ s) = \mathrm{Id}_{T_xU}$$

Suppose now that $D_x(R_g \circ s)(X) = 0$ for some $X \neq 0$, then we see that:

$$D_{s(x)\cdot g}\pi \circ D_x(R_g \circ s)(X) = 0$$

a contradiction, so $D_x(R_g \circ s)$ is injective, and the image of $D_x(R_g \circ s)$ intersected with $\ker D_{s(x)\cdot g}\pi$ is the zero vector. Since the action of G is free and transitive on the fibres, we have that the map:

$$\begin{aligned} T_gG &\longrightarrow T_{s(x)\cdot g}P_x \\ Y &\longmapsto \widetilde{\mu_G(Y)}_{s(x)\cdot g} \end{aligned}$$

is an isomorphism. Furthermore, let $\gamma(t)$ be a curve in P_x such that $\gamma(0) = s(x)\cdot g$, and $\dot{\gamma}(0) = Z \in T_{s(x)\cdot g}P_x$, then:

$$\begin{aligned} D_{s(x)\cdot g}\pi(Z) =& \frac{d}{dt}\Big|_{t=0}\pi(\gamma(t)) \\ =& \frac{d}{dt}\Big|_{t=0} x \\ =& 0 \end{aligned}$$

so $T_{s(x)\cdot g}P_x \subset \ker D_{s(x)\cdot g}\pi$. This then implies that the only way for $D_{(x,g)}\psi(X,Y)$ to be zero is if:

$$D_x(R_g \circ s)(X) = 0 \qquad \text{and} \qquad \widetilde{\mu_G(Y)}_{s(x)\cdot g} = 0$$

However, as we have just shown, the only way these can both zero is if $X = Y = 0$, implying that $D_{(x,g)}\psi$ is injective as desired.

Therefore, ψ is a G equivariant diffeomorphism, and for each $\{U_i\}_{i\in I}$:

$$\begin{aligned} \psi_i : U_i \times G &\longrightarrow P_{U_i} \\ (x,g) &\longmapsto s_i(x)\cdot g \end{aligned}$$

is a G equivariant diffeomorphism. The collection $\{(U_i, \psi_i^{-1})\}_{i\in I}$ is then a principal bundle atlas, as desired. $\square$

Our work in the preceding proposition gives the following the corollary:

Corollary 2.1.1. *Let M be a smooth manifold, and Φ be a principal right action of G on M. Then M is a principal bundle over M/G with structure group G*

Proof. By **Theorem 1.2.4**, $\pi : M \to M/G$ is a smooth submersion. In particular, the fibres of π are preserved by G, and G acts freely and transitively on them. By **Lemma 1.2.9**, we have that π admits local sections, so we can construct a principal bundle atlas in the same way as **Proposition 2.1.5**, and the claim follows from the definition of fibre bundles and principal bundles. $\square$

Furthermore, **Proposition 2.1.5** implies that any fibre bundle with a right group action which preserves the fibres, and is simply transitive on them is a principal G bundle, as the projection from the total space to the base space is always a surjective submersion.

Example 2.1.1. By **Corollary 2.1.1**, we see that the Hopf fibration from **Example 1.2.15** is a principal bundle over $\mathbb{S}^2$, where the total space is $\mathbb{S}^3$ and the structure group is $\mathbb{S}^1$.

We would like to obtain a converse to **Corollary 2.1.1**, i.e. that every principal bundle can be thought of as a quotient manifold. We need the following proposition:

Proposition 2.1.6. *Let P be a principal bundle over M, with structure group G. Then, the right action of G on P is principal.*

Proof. By assumption, the action of G is free, so we need only show that the map:

$$\begin{aligned} \Psi : P \times G &\longrightarrow P \times P \\ (p,g) &\longmapsto (p, p\cdot g) \end{aligned}$$

is closed. We will employ the sequence definition of closed sets to prove this, i.e. that a set A is closed if it contains all of its limits points. Let $A \subset P \times G$ be a closed set, and $(p_i, q_i)_{i\in\mathbb{N}} \in \Psi(A)$ a sequence converging to $(p,q) \in P \times P$. Since the action of G preserves the fibres, and is free and transitive on them, we have that there exists a sequence $g_i \in G$ such that $q_i = p_i \cdot g_i$. If we can show that g_i converges to $g \in G$, then the sequence (p_i, g_i) converges in A as A is closed, and thus $(p, q) \in \Psi(A)$ so $\Psi(A)$ will be closed as well.

Let $\pi(p) = x$, and U an open neighborhood of x with bundle chart:

$$\phi : P_U \longrightarrow U \times G$$

Since (p_i, q_i) is a convergent sequence, there exists an N such that for all $i \geq N$, $p_i, q_i \in P_U$. For some sequence $x_i \in U$, and $h_i \in G$ converging to x and h respectively, we can then write:

$$\begin{aligned}\phi(p_i) &= (x_i, h_i)\\ \phi(q_i) &= (x_i, h_i g_i)\\ \phi(p) &= (x, h)\end{aligned}$$

Since $q_i \to q$ and $x_i \to x$, we find that:

$$\phi(q) = (x, h')$$

for some $h' \in G$. We see that the sequence g_i is then given by:

$$g_i = h_i^{-1}(h_i g_i)$$

Therefore, since h_i converges to h, and $h_i g_i$ converges h', we have that g_i converges to $h^{-1}h'$. As mentioned earlier, A is closed so since the sequence (p_i, g_i) converges to $(p, h^{-1}h')$ we have that $(p, h^{-1}h') \in A$. We see that:

$$\Psi(p, h^{-1}h') = (p, p \cdot (h^{-1}h)) \tag{2.1.4}$$

and note that:

$$\phi(p \cdot (h^{-1}h)) = (x, h') = \phi(q)$$

Then since ϕ is injective, (2.1.4) reduces to:

$$\Psi(p, h^{-1}h') = (p, q)$$

Therefore $(p, q) \in \Psi(A)$, so $\Psi(A)$ is closed and Ψ is a closed map, making the action of G a principal right action, as desired. □

Corollary 2.1.2. *Let P be a principal bundle over M with structure group G. Then P/G has the unique structure of a smooth manifold such that $\pi_Q : P \to P/G$ is a smooth submersion. In particular:*

$$P/G \cong M$$

as smooth manifolds.

Proof. By **Proposition 2.1.6** the right action of G on P is principal, hence by **Theorem 1.2.4** P/G has the unique structure of a smooth manifold such that $\pi_Q : P \to P/G$ is a smooth submersion.

We now need to show that P/G and M are diffeomorphic. Let $x \in M$, and $p \in \pi^{-1}(x)$, then since G is free and transitive on the fibres we have that for any $q \in \pi^{-1}(x)$, there exists a unique $g \in G$ such that $p \cdot g = q$. We then see that the orbit of p, $\mathcal{O}_p$, is equal to the fibre $\pi^{-1}(x)$, as for any element $q \in \pi^{-1}(x)$, we have that $p \cdot g = q$ for some $g \in G$, and for any element $q \in \mathcal{O}_p$ we have that:

$$\pi(q) = \pi(p \cdot g) = x$$

for some $g \in G$. Define a new equivalence relation on P:

$$p \sim q \Longleftrightarrow \pi(p) = \pi(q)$$

We see that this is exactly the equivalence relation defining P/G, as if $\pi(q) = x$, then $q, p \in \pi^{-1}(x)$ so p and q belong to the same orbit. Furthermore, by construction, π is the quotient map $\pi : P \to P/\sim$, hence $P/\sim = M$. Finally, since the defining equivalence relation for P/G is the same as $\sim$, we have that by **Theorem 1.2.3** $P/G \cong M$, as desired. □

Example 2.1.2. Let M be a smooth n dimensional manifold. For each $p \in M$ we define:

$$\mathrm{Fr}_{GL}(M)_p = \{(v_1, \dots, v_n) \in T_pM^n : (v_1, \dots, v_n) \text{ is a basis for } T_pM\}$$

In other words, at each point p, $\mathrm{Fr}_{GL}(M)_p$ is the set of all frames for T_pM. Just as we have done with the tangent spaces, we define the **Frame Bundle** of M as the disjoint union:

$$\mathrm{Fr}_{GL}(M) = \coprod_{p \in M} \mathrm{Fr}_{GL}(M)_p$$

This set comes equipped with a natural projection:

$$\begin{aligned} \pi : \mathrm{Fr}_{GL}(M) &\longrightarrow M \\ (v_1, \dots, v_n)_p &\longmapsto p \end{aligned}$$

Furthermore, there is a natural action of $GL_n(\mathbb{R})$ on $\mathrm{Fr}_{GL}(M)$ given by:

$$\begin{aligned} \Psi : \mathrm{Fr}_{GL}(M) \times GL_n(\mathbb{R}) &\longrightarrow \mathrm{Fr}_{GL}(M) \\ ((v_1, \dots, v_n)_p, g) &\longmapsto \left(\sum_i^n v_i g_{i1}, \dots, \sum_i^n v_i g_{in} \right) \end{aligned}$$

which is essentially just multiplying the matrix $(v_1, \dots, v_n)$ by g on the right. It is clear that this action preserves the fibres of π, and is free and transitive on them. Our goal is to use **Theorem 2.1.1** to show that $\mathrm{Fr}_{Gl}(M)$ is a fibre bundle, and then deduce from **Proposition 2.1.5** that $\mathrm{Fr}_{GL}(M)$ is a principal bundle over M with structure group G.

Let (U_i, ψ_i) be a local chart for M with coordinates $(x^1, \cdots, x^n)$, then a local section of π can be given by:

$$\begin{aligned} s_i : U_i &\longrightarrow \mathrm{Fr}_{GL}(M) \\ p &\longmapsto (\partial_{x^i}, \dots, \partial_{x^n})_p \end{aligned}$$

where ∂_{x^j} is short hand for:

$$D_{\psi_i(p)} \psi_i^{-1} \left(\frac{\partial}{\partial x^j} \Big|_{\psi_i(p)} \right)$$

so $s_i(p)$ can be written as:

$$\begin{aligned} s_i(p) =& \left(D_{\psi_i(p)} \psi_i^{-1} \left(\frac{\partial}{\partial x^1} \Big|_{\psi_i(p)} \right), \dots, D_{\psi_i(p)} \psi_i^{-1} \left(\frac{\partial}{\partial x^n} \Big|_{\psi_i(p)} \right) \right) \\ =& D_{\psi_i(p)} \psi_i^{-1} \cdot \left(\frac{\partial}{\partial x^1} \Big|_{\psi_i(p)}, \dots, \frac{\partial}{\partial x^n} \Big|_{\psi_i(p)} \right) \end{aligned}$$

In the standard $\mathbb{R}^n$ basis where:

$$\frac{\partial}{\partial x^i} \Big|_{\psi_i(p)} = e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

the right term is simply the identity matrix, hence:

$$s_i(p) = D_{\psi_i(p)} \psi_i^{-1}$$

We define the inverse of the bundle chart in the same way we did for **Proposition 2.1.5**:

$$\begin{aligned} \phi_i^{-1} : U_i \times GL_n(\mathbb{R}) &\longrightarrow \mathrm{Fr}_{GL}(M)_{U_i} \\ (p, g) &\longmapsto s_i(p) \cdot g = (D_{\psi_i(p)} \psi_i^{-1}) \cdot g \end{aligned}$$

which is a set bijection by the same argument in **Proposition 2.1.5**. It's inverse is given by:

$$\begin{aligned}\phi_i : \mathrm{Fr}_{GL}(M)_{U_i} &\longrightarrow U_i \times GL_n(\mathbb{R})\\ f &\longmapsto (\pi(f), (D_{\pi(f)}\psi_i) \cdot f)\end{aligned}$$

We can check this an inverse by noting that for any $(p, g) \in U_i \times G$:

$$\begin{aligned}\phi_i \circ \phi_i^{-1}(p, g) =&\phi_i \left((D_{\psi_i(p)}\psi_i^{-1}) \cdot g\right)\\ =&(p, D_p\psi_i \cdot \left((D_{\psi_i(p)}\psi_i^{-1}) \cdot g\right))\\ =&(p, D_{\psi_i(p)}(\psi_i \circ \psi_i^{-1}) \cdot g)\\ =&(p, g)\end{aligned}$$

and that for any $f \in \mathrm{Fr}_{GL}(M)$:

$$\begin{aligned}\phi_i^{-1} \circ \phi_i(f) =&\phi_i^{-1}((\pi(f), (D_{\pi(f)}\psi_i) \cdot f))\\ =&D_{\psi_i(\pi(f))}\psi_i^{-1} \cdot ((D_{\pi(f)}\psi_i) \cdot f)\\ =&(D_{\psi_i(\pi(f))}\psi_i^{-1} \circ D_{\pi(f)}\psi_i)) \cdot f\\ =&f\end{aligned}$$

The transition functions for $U_i \cap U_j \neq \emptyset$ are then given by:

$$\begin{aligned}\phi_j \circ \phi_i^{-1} : (U_i \cap U_j) \times GL_n(\mathbb{R}) &\longrightarrow (U_i \cap U_j) \times GL_n(\mathbb{R})\\ (p, g) &\longmapsto (p, D_{\psi_i(p)}(\psi_j \circ \psi_i^{-1}) \cdot g)\end{aligned}$$

These transition functions are smooth because each $(\psi_j \circ \psi_i)$ is smooth, and clearly have a smooth inverse, hence they are diffeomorphisms. The formal bundle atlas:

$$\{U_i, \phi_i\}$$

is then smoothly compatible, hence $(\mathrm{Fr}_{GL}(M), \pi, M; GL_n(\mathbb{R}))$ is a smooth fibre bundle by **Theorem 2.1.1**. In particular, the action of $GL_n(\mathbb{R})$ is smooth since it is smooth in a smooth bundle chart, so since the action also preserves the fibres, and is free and transitive on them, by **Proposition 2.1.5**, $\mathrm{Fr}_{Gl}(M)$ is a principal bundle over M with structure group $GL_n(\mathbb{R})$.

If (M, g) is a (pseudo)-Riemannian manifold, we could also construct a principal $O(n)$ bundle over M, where the fibres consist of orthornormal frames of T_pM. If in addition M is orientable we can further reduce to the structure group to $SO(n)$, where the fibres consist of the oriented orthonormal frames for T_pM. We denote these bundles by $O(M)$, and $SO(M)$ respectively. This notion of 'reducing' the structure group motivates the following definition.

Definition 2.1.8. Suppose $\pi : P \to M$ and $\pi' : P' \to M$ are principal bundles over M with structure groups G and G' respectively, and $f : G \to G'$ is a Lie group homomorphism. A **principal bundle homomorphism** between P and P' is an f-equivariant map smooth bundle map $F : P \to P'$, i.e.

$$\pi' \circ F = \pi$$

and

$$F(p \cdot g) = F(p) \cdot f(g)$$

The principal bundle P together with the bundle homomorphism F is called a f**-reduction** of P'. In particular, if $f : G \to G'$ is an embedding, then F is G **reduction** of P', and the image of F is a **principal subbundle** of P'.

With this definition, we see that if (M, g) is an orientable (pseudo)-Riemannian manifold, then the inclusion map:

$$i : SO(n) \to GL_n(\mathbb{R})$$

is an embedding, and the bundle inclusion map $SO(M) \to \mathrm{Fr}_{GL}(M)$ determines an $SO(n)$ reduction of $\mathrm{Fr}_{GL}(M)$, as expected.

We end our discussion on principal bundles by finally introducing the notion of a (local) 'gauge transformation'. We need the following definition:

Definition 2.1.9. Let $P \to M$ be a principal bundle with structure group G. A **global gauge** is a global section $s : M \to P$. A **local gauge** is a local section $s : U \to P_U$ defined on an open set $U \subset M$.

By our work in **Proposition 2.1.5** we have the following corollary:

Corollary 2.1.3. *Let $P \to M$ be a principal bundle with structure group G, and $s : U \to P_U$ a local gauge. Then the map:*

$$\begin{aligned} \phi^{-1} : U \times G &\longrightarrow P_U \\ (x, g) &\longmapsto s(x) \cdot g \end{aligned}$$

is a G-equivariant diffeomorphism, and it's inverse is a local trivialization of P. In particular, if $s : M \to P$ is a global gauge then P is a trivial bundle with trivialization given by ϕ:

$$\phi : P \longrightarrow M \times G$$

With the corollary above, we see that gauge's correspond to (local) trivializations of the principal bundle. The core idea behind gauge theories is that the physics is independent of the choice of gauge, or trivialization. This idea mimics those found in geometry, and special/general relativity, i.e. that geometry and thus the physics of gravity should be independent of the coordinates, or the rest frame one chooses in flat spacetime. With this in mind, we see that a gauge transformation is nothing more than a change of local trivialization; in future sections we will further develope this idea, and see how this mathematical gauge transformation aligns with the gauge transformations found in classical electromagnetism.

2.1.3 Associated Vector Bundles

We have already come across a multitude vector bundles in our earlier discussions on smooth manifolds, though we have not yet provided a succinct definition of what a vector bundle actually is; we do so now:

Definition 2.1.10. A fibre bundle $(E, \pi, M; V)$ is called a **real or complex vector bundle** of rank m if:

a) The general fibre V, and every fibre E_x, for $x \in M$ are m-dimensional vector spaces over $\mathbb{R}$ or $\mathbb{C}$.

b) There exists a bundle atlas $\{(U_i, \phi_i)\}_{i \in I}$ for E such that the restriction of ϕ_i to the fibre E_x, denoted ϕ_{ix}:

$$\phi_{ix} : E_x \longrightarrow V$$

is a vector space isomorphism for all $x \in M$. Such an atlas is called a **vector bundle atlas**, and the bundle charts are similarly called **vector bundle charts**.

With this definition above, it should be clear that the tangent bundle is indeed a vector bundle, though the vector bundle structure of T^*M, and $\Lambda^k(T^*M)$, is perhaps less clear. However, by **Theorem 2.1.1** we can obtain new vector bundles from purely linear algebraic constructions on the fibres. Indeed, if E and F are vector bundles, then we can construct vector bundles:

$$E \oplus F \qquad E \otimes F \qquad E^* \qquad \Lambda^k(E)$$

whose fibres are:

$$(E \oplus F)_x = E_x \oplus F_x \qquad (E \otimes F)_x = E_x \otimes F_x \qquad (E^*)_x = E_x^* \qquad (\Lambda^k(E))_x = \Lambda^k(E_x)$$

Then, since E and F come equipped with a vector bundle atlas, we can combine said vector bundle atlas's, and compose them with them various maps in linear algebra, to construct a formal bundle atlas satisfying the assumptions of **Theorem 2.1.1**. In particular, if E is a real rank m vector bundle with bundle charts $\{(U_i, \phi_i)\}$, we can define the bundle charts for E^* by:

$$\begin{aligned} \phi_i^* : E_{U_i}^* &\longrightarrow U_i \times V^* \\ \omega_x &\longmapsto (x, \left((\phi_{ix})^{-1}\right)^T \omega_x) \end{aligned}$$

as $\phi_{ix} : E_x \to V$, we have that $\phi_{ix}^{-1} : V \to E_x$, so $\left(\phi_{ix}^{-1}\right)^T : E_x^* \to V^*$. The rest of the constructions then follow similarly.

Example 2.1.3. Let M be a smooth manifold, and E a vector bundle. Recall the vector bundle $\Lambda^k(T^*M)$, and that:

$$\Gamma\left(\Lambda^k(T^*M)\right) = \Omega^k(M)$$

i.e. smooth sections of this bundle are differential k forms. A vector bundle of interest is then $\Lambda^k(M) \otimes E$. For $x \in M$, if $\omega_x \in (\Lambda^k(M) \otimes E)_x$, then ω_x is the alternating, multilinear map:

$$\omega_x : T_xM \times \cdots \times T_xM \longrightarrow E_x$$

With this in mind, the sections of this bundle are differential k forms with values in $\Gamma(E)$. For brevity we set:

$$\Gamma(\Lambda^k(T^*M) \otimes E) = \Omega^k(M, E)$$

Elements of $\Omega^k(M, E)$ are often called **k-forms twisted with E**.

Much like principal bundles, the transition functions for a vector bundle have the following special property:

Proposition 2.1.7. *Let $(E, \pi, M; V)$ be a rank n vector bundle, where V is a $\mathbb{K}$-linear vector space, with $\mathbb{K}$ being $\mathbb{C}$ or $\mathbb{R}$. The transition functions then take values in $GL_n(\mathbb{K})$:*

$$\begin{aligned} \phi_{ij} : U_i \cap U_j &\longrightarrow GL_n(\mathbb{K}) \subset \mathit{Diff}(V) \\ x &\longmapsto \phi_{jx} \circ \phi_{ix}^{-1} \end{aligned}$$

Proof. Both $\phi_{ix}^{-1} : V \to E_x$, and $\phi_{jx} : E_x \to V$ are isomorphisms, hence $\phi_{jx} \circ \phi_{ix}^{-1} : V \to V$ is an isomorphism. Therefore, since $V \cong \mathbb{K}^n$, and $\det\left(\phi_{jx} \circ \phi_{ix}^{-1}\right) \neq 0$, we have that $\phi_{jx} \circ \phi_{ix} \in GL_n(\mathbb{K})$, as desired. $\square$

Given a vector bundle E, we can also define bundle metrics by constructing the bundle $E^* \otimes E^*$ and then taking non vanishing, smooth, symmetric sections of that bundle.

Definition 2.1.11. Let $E \to M$ be a $\mathbb{C}$ or $\mathbb{R}$ linear vector bundle. A **Euclidean or Hermitian bundle metric** is a metric on each fibre E_x which varies smoothly, i.e. if E is a real vector bundle then it is a section:

$$\langle \cdot, \cdot \rangle \in \Gamma(E^* \otimes E^*)$$

or if E is a complex vector bundle:

$$\langle \cdot, \cdot \rangle \in \Gamma(\bar{E}^* \otimes E^*)$$

where $\bar{E}$ denotes the complex conjugate vector bundle. The section at each point must define a positive definite, symmetric (or hermitian) bilinear map:

$$E_x \times E_x \longrightarrow \mathbb{C} \text{ or } \mathbb{R}$$

Furthermore, every real or complex vector bundle admits such bundle metric.

Proposition 2.1.8. *Let $E \to M$ be a real or complex vector bundle. Then E admits a Euclidean or Hermitian bundle metric.*

Proof. This follows from the exact same partition of unity argument found in the proof of **Theorem 1.1.10**, that every smooth manifold M admits a Riemannian metric. $\square$

Like principal bundles, homomorphisms between vector bundles must satisfy an extra condition, namely that they respect the vector space structure of the fibres.

Definition 2.1.12. Let $\pi : E \to M$ and $\pi' : E' \to M$ be vector bundles. Then a map $F : E \to E'$ is a **vector bundle homomorphism** if it is a bundle homomorphism such that F restricted to the fibres of E is a linear map. Furthermore, F is called a **vector bundle isomorphism** if the restriction to each fibre is a linear isomorphism, and E is called **trivial** if $E \cong M \times V$.

With the definition above, and our discussion at the beginning of this section we wish to understand the tangent bundle of a product manifold $M \times N$. We first cite the following theorem, a proof of which can be found in Hamilton's *Mathematical Gauge Theory.*

Theorem 2.1.2. *Let $(E, \pi, M; F)$ be a fibre bundle, and $f : N \to M$ a smooth map between manifolds. Then the set:*

$$f^*E = \{(x, e) \in N \times E : f(x) = \pi(e)\}$$

has the structure of a fibre bundle over N with general fibre F.

We can now prove the following result:

Proposition 2.1.9. *Let M and N be smooth manifolds. Then, as vector bundles:*

$$T(M \times N) \cong (\pi_M^* TM) \oplus (\pi_N^* TN) \cong TM \times TN$$

where π_M and π_N denote the projection onto M and N respectively.

Proof. We see that:

$$\pi_M^* TM = \{((p, q), (p, v)) \in (M \times N) \times TM\}$$

while:

$$\pi_N^* TN = \{((p, q), (q, w)) \in (M \times N) \times TN\}$$

So the fibres of each then satisfy:

$$(\pi_M^* TM)_{(p,q)} = T_p M \qquad \text{and} \qquad (\pi^* TN)_{(p,q)} = T_q N$$

Hence:

$$((\pi_M^* TM) \oplus (\pi^* TN))_{(p,q)} = T_p M \oplus T_q N$$

We thus define a map by:

$$\begin{aligned} F : T(M \times N) &\longrightarrow (\pi_M^* TM) \oplus (\pi^* TN) \\ (p, q, v) &\longmapsto (p, q, D_{(p,q)}\pi_M(v), D_{(p,q)}\pi_N(v)) \end{aligned}$$

We see that this map is smooth as the global differential of a smooth map is smooth. Furthermore, for $\pi : T(M \times N) \to M \times N$ and $\pi' : (\pi_M^* TM) \oplus (\pi_N^* TN) \to M \times N$ we have that:

$$\pi' \circ F(p, q, v) = (p, q) = \pi(p, q, v)$$

hence F is a smooth bundle homomorphism. We need to check that $F_{(p,q)}$ is a linear isomorphism $T(M \times N)_{(p,q)} \to T_p M \oplus T_q M$. It is clear from the construction of F that $F_{(p,q)}$ is linear. Since the dimension of these two vector spaces is clearly the equal, we need only check that $F_{(p,q)}$ has trivial kernel. Suppose $v \neq 0 \in T_{(p,q)}(M \times N)$ such that:

$$F_{(p,q)}(v) = 0$$

then $v \in \ker D_{(p,q)}\pi_M \cap \ker D_{(p,q)}\pi_N$. Let $\gamma : I \to M \times N$ be a smooth curve such that $\dot{\gamma}(0) = v \in T_{(p,q)}(M \times N)$. Since γ is a smooth curve in $M \times N$, we see that:

$$\gamma(t) = (p(t), q(t))$$

for smooth curves $p : I \to M$, and $q : I \to n$ such that $p(0) = p$ and $q(0) = q$. If $v \in \ker D_{(p,q)}\pi_M$ we have that:

$$D_{(p,q)}\pi_M v = \frac{d}{dt}\Big|_{t=0} \pi_M(\gamma(t)) = \dot{p}(0) = 0$$

and if $v \in \ker D_{(p,q)}\pi_N$ then:

$$D_{(p,q)}\pi_N v = \frac{d}{dt}\Big|_{t=0} \pi_N(\gamma(t)) = \dot{q}(0) = 0$$

hence:

$$\dot{\gamma}(0) = v = (\dot{p}(0), \dot{q}(0)) = (0,0)$$

Therefore, $v \in \ker D_{(p,q)}\pi_M \cap \in \ker D_{(p,q)}\pi_N$ implies that v is zero, so $F_{(p,q)}$ is injective and thus an isomorphism as desired.

We now need to show that:

$$(\pi_M^* TM) \oplus (\pi_N^* TN) \cong TM \times TN$$

We first need to show that $TM \times TN$ is a vector bundle over $M \times N$. Note that $TM \times TN$ comes equipped with the natural projections $\pi_{TM} : TM \times TN \to TM$, and $\pi_{TN} : TM \times TN \to TN$, and further that TM and TN come equipped with the projections $\pi_1 : TM \to M$ and $\pi_2 : TN \to N$. We define the smooth projection $\pi'' : TM \times TN \to M \times N$ by:

$$\pi'' = (\pi_1 \circ \pi_{TM}, \pi_2 \circ \pi_{TN})$$

as for all $((p,v),(q,w)) \in TM \times TN$ we have:

$$\pi''((p,v),(q,w)) = (p,q)$$

Furthermore, Let $\{(U_i, \phi_i)\}$ and $\{(V_i, \psi_i)\}$ be a countable covering of M and N by coordinate carts, then we have coordinate charts for TM and TN constructed in **Proposition 1.1.3** given by:

$$\phi_{TM_i}(v^i \partial_i|_p) = (\phi(p), v^1, \dots, v^m) \qquad \text{and} \qquad \psi_{TN_i}(w^i \partial_i|_q) = (\psi(q), w^1, \dots, w^n)$$

where $m = \dim M$ and $n = \dim N$. From here we construct fibre bundle charts:

$$\begin{aligned} \tilde{\phi}_{TM_i} :& TM_{U_i} \longrightarrow U_i \times \mathbb{R}^m \\ (v^i \partial_i|_p) &\longmapsto (p, v^1, \dots, v^m) \end{aligned}$$

which is smooth as it is the composition of smooth maps:

$$\tilde{\phi}_{TM_i} = \phi^{-1} \times \mathrm{Id}_{\mathbb{R}^m} \circ \phi_{TM}$$

and clearly satisfies:

$$\pi_{TM}|_{TM_U} = \pi_U \circ \tilde{\phi}_{TM_i}$$

Furthermore, these are vector bundle charts as for all $p \in U_i$ each $\tilde{\phi}_{TM_i}$ restricts to a linear isomorphism $\tilde{\phi}_{TM_{ip}} : T_pM \to \mathbb{R}^n$ given by $D_p\phi_i$, with inverse $D_{\phi_i(p)}\phi_i^{-1}$. In a similar manner, we see that:

$$\begin{aligned} \tilde{\psi}_{TN_i} :& TN_{V_i} \longrightarrow V_i \times \mathbb{R}^n \\ (w^i \partial_i|_q) &\longmapsto (p, w^1, \dots, w^n) \end{aligned}$$

are vector bundle charts for TN. We now want to construct a formal bundle atlas for $TM \times TN$; recall that that $M \times N$ comes equipped with a product manifold structure, so $\{(U_i \times V_j, \phi_i \times \psi_j)\}$ covers $M \times N$. For each i and j, we construct the bundle charts:

$$\begin{aligned} \tilde{\phi}_{TM_i} \times \tilde{\psi}_{TN_j} : (TM \times TN)_{U_i \times V_j} &\longrightarrow (U_i \times V_j) \times (R^m \oplus \mathbb{R}^n) \\ ((p,v),(q,w)) &\longmapsto ((p,q),(D_p\phi_i(v), D_p\psi_j(w))) \end{aligned}$$

which are clearly smoothly compatible as the global differential $D(\phi_i \circ \phi_k^{-1})$ is smooth, so by **Theorem 2.1.1**, $TM \times TN$ is a fibre bundle with model fibre $F = \mathbb{R}^m \oplus \mathbb{R}^n$. Clearly the restriction of $\tilde{\phi}_{TM_i} \times \tilde{\psi}_{TN_j}$ to the point (p,q) is then an isomorphism:

$$(TM \times TN)_{(p,q)} \longrightarrow \mathbb{R}^m \oplus \mathbb{R}^n$$

so $TM \times TN$ is a vector bundle over $M \times N$.

Now that we know $TM \times TN$ is a vector bundle over $M \times N$, we construct the smooth map:

$$\begin{aligned} F : (\pi_M^* TM) \oplus (\pi_N^* TN) &\longrightarrow TM \times TN \\ (p,q,v,w) &\longmapsto ((p,v),(q,w)) \end{aligned}$$

which is a bundle homomorphism as for all $(p,q,v,w) \in (\pi_M^* TM) \oplus (\pi_N^* TN)$ we have:

$$\pi'' \circ F(p,q,v,w) = \pi''((p,v),(q,w)) = (p,q) = \pi'(p,q,v,w)$$

The restriction of F to the the fibre over $(p,q) \in M \times N$ is just the identity map, since for $(v,w) \in T_pM \oplus T_qN$:

$$F_{(p,q)}(v,w) = (v,w)$$

Thus, $F_{(p,q)}$ is a linear isomorphism for all $(p,q) \in M \times N$, so F is a bundle isomorphism, implying the claim. □

Though there is much to say about vector bundles in general, we are interested in a specific type of vector bundle, namely ones that are related in a way to a principal bundle in a special way. In physics, we will be able to identify the sections of these associated vector bundles as matter fields, and, as mentioned earlier, a very special associated vector bundle will allow us to write the Yang-Mill's Lagrangian down without reference to the underlying principal bundle.

To begin, we need the following lemma:

Lemma 2.1.1. *Let $P \to M$ be a principal G bundle, and ρ a representation of G on a $\mathbb{K}$-linear vector space V. Then the map:*

$$\begin{aligned} \Phi : (P \times V) \times G &\longrightarrow P \times V \\ (p,v,g) &\longmapsto (p \cdot g, \rho(g)^{-1} \cdot v) \end{aligned}$$

defines a principal right action of the Lie group G on the product manifold $P \times V$. In particular, the quotient space $E = (P \times V)/G$ has the unique structure of a smooth manifold such that $\pi : P \times V \to E$ is a submersion.

Proof. It is clear that the map above is smooth. It is also a right action as:

$$\begin{aligned} \Phi(\Phi((p,v,g),h)) =& \Phi(p \cdot g, \rho(g)^{-1} \cdot v, h) \\ =& (p \cdot (g \cdot h), \rho(g \cdot h)^{-1} v) \\ =& \Phi(p,v,gh) \end{aligned}$$

Let $\phi_{p,v}$ denote the orbit map through $(p,v) \in P \times V$. For $g,h \in G$, if:

$$\phi_{p,v}(g) = \phi_{p,v}(h)$$

we have that:

$$(p \cdot g, \rho(g)^{-1} v) = (p \cdot h, \rho(g)^{-1} h)$$

However, the action of G on P is free, so if $p \cdot g = p \cdot h$, we must have that $g = h$, so Φ is a free a right action. We now need only show that the action is principal, i.e. that the map:

$$\begin{aligned} \Psi : (P \times V) \times G &\longrightarrow (P \times V) \times (P \times V) \\ (p,v,g) &\longmapsto \big((p,v),(p \cdot g, \rho(g)^{-1} v)\big) \end{aligned}$$

is closed. We will proceed similarly to **Proposition 2.1.6**. Let $A \subset (P \times V) \times G$ be closed, and $((p_i,v_i),(q_i,w_i))_{i\in\mathbb{N}}$ be a sequence in $\Psi(A)$ converging to $((p,v),(q,w))$ in $(P \times V) \times (P \times V)$. Since the action of G on P preserves the fibres of P, and is free and transitive on them, there exists a unique sequence $g_i \in G$ such that $q_i = p_i \cdot g_i$. Since $((p_i,v_i),(q_i,w_i)) \in \Psi(A)$, it then follows that $w_i = \rho(g_i)^{-1} v_i$ for all $i \in \mathbb{N}$, so:

$$\Psi(p_i,v_i,g_i) = ((p_i,v_i),(q_i,w_i))$$

for all $i \in \mathbb{N}$. Let $\pi(p) = x$ and U an open neighborhood of x with bundle chart:

$$\phi : P_U \to U \times G$$

Then, there exists an $N \in \mathbb{N}$ such that for all $i \geq N$ we have that $(p_i, v_i), (q_i, w_i) \in P_U \times V$. Consider the map:

$$\begin{aligned} \psi : P_U \times V &\longrightarrow U \times G \times V \\ (p, v) &\longmapsto (\phi(p), \rho(\mathrm{pr}_G \circ \phi(p))v) \end{aligned}$$

which is clearly a bijection as ϕ is a diffeomorphism and $\rho(g)$ is an isomorphism for each $g \in G$. Furthermore, the group action on $U \times G \times V$ compatible with the group action on $P \times V$ is given by:

$$(p, h, v) \cdot g = (p, hg, v)$$

as if:

$$\psi(p, v) = (p, h, \rho(h)v)$$

then:

$$\begin{aligned} \psi(p \cdot g, \rho(g)^{-1}v) =& (x, hg, \rho(hg)\rho(g)^{-1}v) \\ =& (x, hg, \rho(h)v) \\ =& (x, h, \rho(h)v) \cdot g \\ =& \psi(p, v) \cdot g \end{aligned}$$

We also see that this map is constant in the first and third component for all $(r, u) \in \mathcal{O}_{p,v}$ as for some $g \in G$:

$$\begin{aligned} \psi(r, u) =& \psi(p \cdot g, \rho(g)^{-1}v) \\ =& (x, hg, \rho(hg)\rho(g)^{-1}v) \\ =& (x, hg, \rho(h)v) \end{aligned}$$

Then for some sequences x_i and h_i converging to x and h respectively, we have that:

$$\begin{aligned} \psi(p_i, v_i) =& (x_i, h_i, \rho(h_i)v_i) \\ \psi(q_i, w_i) =& (x_i, h_i g_i, \rho(h_i)v_i) \\ \psi(p, v) =& (x, h, \rho(h)v) \end{aligned}$$

Since $q_i \to q$, $w_i \to w$, $\rho(h_i)v_i \to \rho(h)v$, and $x_i \to x$ we also have that for some $h' \in G$:

$$\psi(q, w) = (x, h', \rho(h)v)$$

We write g_i as:

$$g_i = h_i^{-1}(h_i g_i)$$

then since h_i converges to h, and $h_i g_i$ converges h', we see that g_i converges to $g = h^{-1}h'$, so $(p, v, g) \in A$ as A is closed. We now see that:

$$\psi(p \cdot g, \rho(g)^{-1}v) = (x, h', \rho(h)v) = \psi(q, w)$$

hence since ψ is injective:

$$\Psi(p, v, g) = ((p, v), (q, w))$$

Therefore $((p, v), (q, w)) \in \Psi(A)$, so $\Psi(A)$ is closed, making Ψ a closed map, and Φ a principal right action as desired. The rest of the claim follows from **Theorem 1.2.4**. $\square$

Now that we know that $E = (P \times V)/G$ has the structure of a smooth manifold, we would like to go one step further and prove that E is actually a vector bundle over M with general fibre isomorphic to V.

Theorem 2.1.3. *Let P be a principal bundle over M with structure G, V a $\mathbb{K}$ linear vector space, and ρ a representation of G on V. Then the quotient space $E = (P \times V)/G$ has the structure of a $\mathbb{K}$-vector bundle over M with projection:*

$$\begin{aligned} \pi_E : E &\longrightarrow M \\ [p, v] &\longmapsto \pi_P(p) \end{aligned}$$

and fibres:

$$E_x = (P_x \times V)/G$$

isomorphic to V. The vector space structure of the fibre E_x over $x \in M$ is given by:

$$\lambda[p, v] + \eta[p, w] = [p, \lambda v + \eta w], \qquad \forall \lambda, \eta \in \mathbb{K},\ v, w \in V,\ p \in P$$

where $\pi_P(p) = x$.

Proof. We first show that the map $\pi_E : E \to M$ is well defined. Let $[p, v] \in E$, and $[q, w]$ be another representative of the equivalence class $[p, v]$. Then, for some $g \in G$ we see that $q = p \cdot g$ and $w = \rho(g)^{-1} \cdot v$. We then see that:

$$\pi_E([q, w]) = \pi_P(q) = \pi_P(p \cdot g) = \pi_P(p)$$

so the map is well defined. Furthermore, the fibre $E_x = (P_x \times V)/G$ is isomorphic to V under the map:

$$\begin{aligned} f : V &\longrightarrow E_x \\ v &\longmapsto [p, v] \end{aligned}$$

for a fixed $p \in P_x$. The map is clearly linear, and is injective as if $[p, v] = [p, w]$, then $v = w$, since (p, v) and (p, w) can't be in the same orbit. Finally, the map is surjective, since if $[p, v] \in E_x$, we see that $f(v) = [p, v]$, so f is an isomorphism.

We now need to show that E has the structure of vector bundle, so we need to show that for each $x \in M$, there exists an open neighborhood U of x such that E_U is diffeomorphic to $U \times V$. Let U be an open neighborhood of $x \in M$, and ϕ_U a local trivialization:

$$\phi : P_U \longrightarrow U \times G$$

We define a map:

$$\begin{aligned} \psi : E_U &\longrightarrow U \times V \\ [p, v] &\longmapsto (\pi_P(p), \rho(\mathrm{pr}_G \circ \phi(p))v) \end{aligned}$$

As shown earlier, $\rho(\mathrm{pr}_G \circ \phi(p))v$ is constant for all $(q, w) \in \mathcal{O}_{p,v}$, so the map is well defined. Furthermore, it is smooth as π_P is a submersion, and $\rho(\mathrm{pr}_G \circ \phi(p))v$ is the composition of smooth maps. We see that it has an inverse given by:

$$\begin{aligned} \psi^{-1} : U \times V &\longrightarrow E_U \\ (x, v) &\longmapsto [\phi^{-1}(x, e), v] \end{aligned}$$

We now check that this indeed an inverse. Let $[p, v] \in E_U$, $\pi(p) = x$, and $\phi(p) = (x, h)$ for some $x \in M$, $h \in G$, then:

$$\begin{aligned} \psi^{-1} \circ \psi([p, v]) =& \psi^{-1}(x, \rho(h)v) \\ =& [\phi^{-1}(x, e), \rho(h)v] \\ =& [\phi^{-1}(x, h) \cdot h^{-1}, \rho(h)v] \\ =& [p \cdot h^{-1}, \rho(h^{-1})^{-1}v] \\ =& [p, v] \end{aligned}$$

and:

$$\begin{aligned}\psi \circ \psi^{-1}(x,v) =&\psi([\phi^{-1}(x,e),v])\\ =&(x,\rho(e)v)\\ =&(x,v)\end{aligned}$$

hence $\psi^{-1} \circ \psi = \mathrm{Id}_{E_U}$, and $\psi \circ \psi^{-1} = \mathrm{Id}_{U \times V}$ as desired. Furthermore, the inverse map is smooth, as ϕ^{-1} is smooth, and the projection $P \times V \to E$ is a smooth submersion, so ψ is a diffeomorphism. Finally, the restriction of ψ to the fibre E_x is a linear isomorphism, as $\psi_x^{-1} : V \to E_x$ is just the isomorphism f, so $\psi_x = f^{-1}$ is an isomorphism as well. Therefore ψ is a vector bundle chart for U, so E is a vector bundle over M as desired. □

With the theorem above, we can now properly define associated vector bundles:

Definition 2.1.13. Let P be a principal bundle over M with structure group G, and ρ a representation on a $\mathbb{K}$ linear vector space V. The vector bundle:

$$E = P \times_\rho V = (P \times V)/G$$

is called the **vector bundle associated** to P and the representation ρ on V. The group G is also called the **structure group** of E.

Now that we an apt description of associated vector bundles, we would like to know how to construct local vector fields on said bundle, i.e. local sections of E. As mentioned earlier, these sections will be thought of as matter fields on our spacetime.

Proposition 2.1.10. *Let P be a principal bundle over M, and ρ a representation of the structure group G on a $\mathbb{K}$ linear vector space V. Let $s : U \to P_U$ be a local gauge, then there is a one to one correspondence between the smooth sections of $E = P \times_\rho V$, $\tau : U \to E_U$ and smooth maps $f : U \to V$.*

Proof. Let $f : U \to V$, then the map:

$$\begin{aligned}\tau : U &\longrightarrow E_U\\ x &\longmapsto [s(x), f(x)]\end{aligned}$$

is smooth as both s and f are smooth, and the map $P \times V \to E$ is a smooth submersion. Furthermore, it is a section of E as for $x \in M$:

$$\pi_E \circ \tau(x) = \pi_P \circ s(x) = x$$

Now conversely, let τ be a smooth section E, and suppose that for $v, w \in V$:

$$[s(x), v] = \tau(x) = [s(x), w]$$

then we must have that $v = w$, as $(s(x), v)$ and $(s(x), w)$ can't be in the same orbit. Therefore, there exists a unique $f(x) \in V$ for each $x \in U$ such that:

$$\tau(x) = [s(x), f(x)]$$

We now need to show that $f(x)$ is smooth. Define a bundle chart by the section s:

$$\begin{aligned}\phi^{-1} : U \times G &\longrightarrow P_U\\ (x,g) &\longmapsto s(x) \cdot g\end{aligned}$$

and a vector bundle chart by:

$$\begin{aligned}\psi : E_U &\longrightarrow U \times V\\ [p,v] &\longmapsto (\pi_P(p), \rho(\mathrm{pr}_G \circ \phi(p))v)\end{aligned}$$

Then we see that:

$$\phi^{-1}(x,e) = s(x) \Rightarrow \phi(s(x)) = (x,e)$$

so:

$$\begin{aligned}\psi \circ \tau(x) =&\psi([s(x), f(x)])\\ =&(x, \rho(\mathrm{pr}_G(x, e))f(x))\\ =&(x, f(x))\end{aligned}$$

Since $\psi \circ \tau$ is smooth, it must be smooth in both of it's components hence $f(x)$ is smooth. $\square$

Let $P \to M$ be a principal bundle with structure group G, and $E = P \times_\rho V$ a vector bundle associated to P. Then, for a principal bundle atlas $\{U_i, \phi_i\}$ determined by local gauges s_i, we can construct a vector bundle atlas for E.

Definition 2.1.14. The principal bundle atlas, mentioned above, determines an **adapted bundle** atlas for E with trivializations:

$$\begin{aligned}\psi_i : E_{U_i} &\longrightarrow U_i \times V\\ [p, v] &\longmapsto (\pi_P(p), \rho(\mathrm{pr}_G \circ \phi_i(p))v)\end{aligned}$$

whose inverses are given by:

$$\begin{aligned}\psi_i^{-1} : U_i \times V &\longrightarrow E_{U_i}\\ (x, v) &\longmapsto [s_i(x), v]\end{aligned}$$

Note that we constructed bundle charts for an associated vector bundle in **Theorem 2.1.2** in the same way, however those had an inverse given by:

$$\begin{aligned}\psi_i^{-1} : U_i \times V &\longrightarrow E_{U_i}\\ (x, v) &\longmapsto [\phi_i^{-1}(x, e), v]\end{aligned}$$

But, for a principal bundle chart defined by a section we had that:

$$\phi_i^{-1}(x, g) = s_i(x) \cdot g \Longrightarrow \phi_i^{-1}(x, e) = s_i(x)$$

So the only new part of this construction is the replacement of the local gauge s_i with $\phi_i^{-1}(x, e)$. Though this may seem inconsequential at moment, it often more practical to work with arbitrary local gauges than it is with arbitrary bundle charts.

As one should now expect, the transition functions for an adapted bundle atlas of an associated vector bundle have the following special property:

Proposition 2.1.11. *Let $P \to M$ be a principal bundle with structure group G, $E = P \times_\rho G$ a vector bundle associated to P, and $\{U_i, \phi_i\}$ a principal bundle atlas for P determines by local gauges s_i. The transition functions for P are given by:*

$$\begin{aligned}\phi_{ij} : U_i \cap U_j &\longrightarrow G\\ x &\longmapsto \phi_{jx} \circ \phi_{ix}^{-1}\end{aligned}$$

The transition functions for E are then:

$$\begin{aligned}\psi_{ij} : U_i \cap U_j &\longrightarrow GL(V)\\ x &\longmapsto \psi_{jx} \circ \psi_{ix}^{-1} = \rho(\phi_{ij}(x))\end{aligned}$$

Thus the transition function of E have image in the subgroup $\rho(G) \subset GL(V)$.

Proof. Let s_i and s_j be the local gauges which determine ϕ_i and ϕ_j. Then, since the action of G on P is free and transitive, we have that there exist uniquely determined $g(x)$ such that for all $x \in M$::

$$s_i(x) = s_j(x) \cdot g(x)$$

Furthermore, we have that for all $h \in H$:

$$\phi_{ix}^{-1}(h) = s_i(x) \cdot h$$

So:

$$\begin{aligned}\phi_{jx} \circ \phi_{ix}^{-1}(h) =&\phi_{jx}(s_i(x)) \cdot h \\ =&\phi_{jx}(s_j(x) \cdot g(x)) \cdot h \\ =&g(x) \cdot h\end{aligned}$$

as $\phi_{jx}(s_j(x)) = e$. Since $\phi_{ij}(x) = \phi_{jx} \circ \phi_{ix}^{-1}$, and $\phi_{ij}(x) \in G$ we see that:

$$\phi_{ij}(x) = g(x)$$

so:

$$s_i(x) = s_j(x) \cdot \phi_{ij}(x)$$

With this in mind, we see that for all $v \in V$:

$$\begin{aligned}\psi_{ix}^{-1}(v) =&[s_i(x), v] \\ =&[s_j(x) \cdot \phi_{ij}(x), v] \\ =&[s_j(x), \rho(\phi_{ij}(x)v)] \\ =&\psi_{jx}^{-1}(\rho(\phi_{ij}(x))v)\end{aligned}$$

Thus we obtain:

$$\begin{aligned}\psi_{ij}(x)(v) =&\psi_{jx} \circ \psi_{ix}^{-1}(v) \\ =&\psi_{jx} \circ \psi_{jx}^{-1}(\rho(\phi_{ij}(x))v) \\ =&\rho(\phi_{ij}(x))v\end{aligned}$$

Hence:

$$\psi_{ij}(x) = \rho(\phi_{ij}(x))$$

as desired. □

Example 2.1.4. Let M be a smooth n-dimensional manifold and $\mathrm{Fr}_{GL}(M)$ the frame bundle of M. We want to see that with standard representation, ρ, of $GL_n(\mathbb{R})$ on $\mathbb{R}^n$ given by multiplication on the left by column vectors that:

$$TM \cong \mathrm{Fr}_{GL}(M) \times_\rho \mathbb{R}^n$$

Consider the smooth map:

$$\begin{aligned}F : \mathrm{Fr}_{GL}(M) \times_\rho \mathbb{R}^n &\longrightarrow TM \\ [(v_1, \dots, v_n)_p, (x^1, \dots, x^n)] &\longmapsto v_i x^i = \rho((v_1, \dots, v_n)_p) \cdot \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}\end{aligned}$$

where we have that $(v_1, \dots, v_n)_p$ is a frame for T_pM, so $v_i x^i \in T_pM$. We need to check that this map is well defined. Let $[(w_1, \dots, w_n)_p, (y_1, \dots, y_n)] = [(v_1, \dots, v_n)_p, (x^1, \dots, x^n)]$, then for some $g \in G$:

$$\begin{aligned}F([(w_1, \dots, w_n)_p, (y_1, \dots, y_n)]) =&F([(v_1, \cdots, v_n)_p \cdot g, \rho(g)^{-1}(x^1, \dots, x^n)]) \\ =&\rho((v_1, \dots, v_n)_p \cdot g)\rho(g^{-1}) \cdot \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \\ =&v_i x^i\end{aligned}$$

so F is well defined. Furthermore, F is a bundle homomorphism as for all $p \in M$ we have that:

$$\pi_{TM} \circ F([(v_1, \dots v_n)_p, (x_1, \dots, x_n)]) =\pi_{TM}(v_i x^i) = p$$

while:

$$\pi_{\mathrm{Fr}}([(v_1,\dots,v_n)_p,(x_1,\dots,x_n)]) = p$$

It is clear that F respects the vector space structure of the fibres, so it then suffices to check that F restricts to an isomorphism of vector spaces on the fibres, however T_pM has the same dimension as $\mathrm{Fr}_{GL}(M)\times_\rho \mathbb{R}^n_p$, so we need only check that the restriction of F to $\mathrm{Fr}_{GL}(M)\times_\rho \mathbb{R}^n_p$ has trivial kernel. For some $[(v_1,\dots,v_n)_p,(x_1,\dots,x_n)]$ suppose we have that:

$$v_i x^i = 0$$

However, this either implies that $(v_1,\dots,v_n)_p$ is not a linear independent set of vectors or that every x^i is zero, and since $(v_1,\dots,v_n)$ is assumed to be a frame, we conclude that every x^i is zero, so the kernel is trivial. Therefore, F is a bundle isomorphism, so:

$$TM \cong \mathrm{Fr}_{GL}(M)\times_\rho \mathbb{R}^n$$

as desired.

Proposition 2.1.12. *Let $P\to M$ be a principal bundle with structure group G, and $E = P\times_\rho V$ a vector bundle associated to P for some n-dimensional $\mathbb{K}$-linear vector space V. If ρ is the trivial representation then $E\cong M\times V$.*

Proof. Since ρ is trivial, we have that for any $[p,v]\in E$:

$$[p,v] = [p\cdot g, v]$$

for all $g\in G$. We define a map F by:

$$\begin{aligned} F : E &\longrightarrow M\times V \\ [p,v] &\longmapsto (\pi(p),v) \end{aligned}$$

which is clearly smooth. It is well defined as for any $[p,v]\in E$ we have:

$$F([p\cdot g,v]) = (\pi(p\cdot g),v) = (\pi(p),v)$$

Furthermore, we see that if $\pi(p)=x$:

$$\pi_M\circ F([p,v]) = \pi_M(x,v) = x$$

hence F is a smooth bundle homomorphism. As before it now suffices to check that F_x is a linear isomorphism for all $x\in M$. Clearly F_x is linear, furthermore it's kernel is trivial as $F_x(v)=v$ for all $v\in E_x$, so F_x is a linear isomorphism for all $x\in M$. Thus:

$$E\cong M\times V$$

as desired. □

We will need a bundle metric on the vector bundle associated to some principal bundle in order to write down the Yang-Mills Lagrangian. For the Lagrangian to be 'gauge invariant' however, we will need a specific type of bundle metric which has this invariance baked into it. Fortunately for us, we can obtain such a bundle metric quite easily if we are first given a G-invariant scalar product on the model fibre V, where G is the structure group of the associated bundle. Gauge invariance will be a much clearer concept by the time we write down the Yang-Mill's Lagrangian, so for now we only show existence.

Proposition 2.1.13. *Suppose that $P\to M$ is a principal bundle with structure group G, and $E = P\times_\rho V$ a vector bundle associated to P. Let $\langle\cdot,\cdot\rangle_V$ be a G-invariant inner product on V. Then the bundle metric $\langle\cdot,\cdot\rangle_E$ on E given by:*

$$\langle [p,v],[p,w]\rangle_{E_x} = \langle v,w\rangle_V$$

for arbitrary $p\in P_x$ is well defined.

Proof. First note that if $\langle \cdot, \cdot \rangle_E$ is well defined, then for all $x \in M$, $\langle \cdot, \cdot \rangle_{E_x}$ is non-degenerate, and smooth as for a local gauge $s : U \to P_U$, and smooth map's $\phi, \psi : U \to V$:

$$\langle [s, \phi], [s, \psi] \rangle_E = \langle \phi, \psi \rangle_V$$

hence $\langle \cdot, \cdot \rangle_E$ is a bundle metric.

To see that the map is well defined, let $[q, u] = [p, v]$ and $[q, y] = [p, v]$, then for some $g \in G$ we have that $q = p \cdot g$, implying that $u = \rho(g)^{-1} v$, and $y = \rho(g)^{-1} v$. Therefore:

$$\begin{aligned} \langle [q, u], [q, y] \rangle_{E_x} =& \langle u, y \rangle_V \\ =& \langle \rho(g)^{-1} v, \rho(g)^{-1} w \rangle_V \\ =& \langle v, w \rangle_V \end{aligned}$$

So the value of the bundle metric is independent of the class representative we choose, implying the claim. □

We end our discussion on associated vector bundles with the following example:

Example 2.1.5. Let $P \to M$ be a principal bundle with structure group G. Denote the Lie algebra of G by $\mathfrak{g}$, and consider the adjoint representation of G on $\mathfrak{g}$:

$$\mathrm{Ad} : G \longrightarrow GL(\mathfrak{g})$$

The associated vector bundle:

$$\mathrm{Ad}(P) = P \times_{\mathrm{Ad}} \mathfrak{g}$$

is called the **adjoint bundle**. As we shall see later, the curvature form on P has a unique representative in $\Omega(M, \mathrm{Ad}(P))$. In particular, the adjoint bundle of the principal bundle $\mathbb{S}^3 \to \mathbb{S}^2$, with structure group $\mathbb{S}^1$, is $\mathbb{S}^3 \times_{\mathrm{Id}} \mathfrak{u}(1) \cong \mathbb{S}^3 \times i\mathbb{R}$.

2.1.4 Connections

We first need the following definition:

Definition 2.1.15. Let $(E, \pi, M; V)$ and $(E', \pi, M; W \subset V)$ be a vector bundles. Then E' is a **vector subbundle** of E if for all $x \in M$, E'_x is a vector subspace of E. In particular, if $E = TM$ a vector subbundle of TM is called a **distribution**.

We will define connections first as distributions of TP, and then show that we can analogously view them in a less abstract manner: as Lie algebra valued one forms on P. To begin, we first want to show that for every principal bundle there exists a canonical vertical bundle.

Definition 2.1.16. The **vertical tangent space** V_p at the point p is the tangent space of the fibre, $T_p P_x$, where $\pi(p) = x$.

Proposition 2.1.14. *The vertical tangent space satisfies the following properties:*

a) $V_p = \ker D_p \pi$

b) The map:

$$\begin{aligned} \phi_* : \mathfrak{g} &\longrightarrow V_p \\ X &\longmapsto \tilde{X}_p \end{aligned}$$

where $\tilde{X}$ is the fundamental vector associated to X determined by the G-action on P is a vector space isomorphism.

c) The set of all vertical tangent spaces V_p for $p \in P$ forms a smooth distribution on P, called the ***vertical tangent bundle****, and is denoted by V. The distribution is a trivial vector bundle via the map:*

$$\begin{aligned} F : P \times \mathfrak{g} &\longrightarrow V \\ (p, X) &\longmapsto \tilde{X}_p \end{aligned}$$

d) The vertical tangent bundle is ***right invariant****, i.e:*

$$R_{g*}(V_p) = V_{p\cdot g}$$

for all $g \in G$.

Proof. We see that $a)$ follows from **Corollary 2.1.2** and **Corollary 1.2.6** as if $\pi(p) = x$ we have that $\mathcal{O}_p = P_x$, hence $\ker D_p\pi = T_p\mathcal{O}_p = T_pP_x$.

For $b)$, we see that by the definition of fundamental vector fields:

$$\begin{aligned}
D_p\pi(\tilde{X}_p) =& \frac{d}{dt}\Big|_{t=0} \pi(p \cdot \exp(tX)) \\
=& \frac{d}{dt}\Big|_{t=0} \pi(p) \\
=& 0
\end{aligned}$$

so $\phi_* X \in V_p$. Since $\dim T_pP_x = \dim \mathfrak{g}$ it then suffices to show that ϕ_* is injective. Let $p \in P$, ϕ_p be the orbit map through p and $X \in \mathfrak{g}$; by our work in **Proposition 1.2.15**:

$$\phi_*(X) = \tilde{X}_p = D_e\phi_p(X)$$

Then **Corollary 2.1.2** and our work in **Corollary 1.2.6** imply that ϕ_p is an embedding, so $D_e\phi_p(X)$ is injective, hence ϕ_* is injective.

It is clear that $\pi_V : V \to P$ is a smooth vector subbundle of TP. Furthermore, $\pi_P : P \times \mathfrak{g} \to P$ is a trivial vector bundle. We see that $\pi_P(p, X) = p$, thus:

$$\begin{aligned}
\pi_V \circ F(p, X) =& \pi_V(\tilde{X}_p) \\
=& p \\
=& \pi_P(p, X)
\end{aligned}$$

so F is a bundle homomorphism. Furthermore, F is clearly smooth as it is the restriction of :

$$D\Phi : TP \times TG \longrightarrow TP$$

to $P \times T_eG$, where $D\Phi$ is the global differential of the right group action on P. It then suffices to show that the restriction of F to the fibre $\{p\} \times \mathfrak{g}$ is an isomorphism, but this follows from $b)$, hence V is a trivial distribution.

To prove $d)$ recall **Proposition 1.2.16**, then we have that a fundamental vector field $\tilde{X}$satisfies:

$$R_{g*}\tilde{X} = \tilde{Y}$$

for $Y = \mathrm{Ad}_{g^{-1}}(X)$, so for all $\tilde{X}_p \in V_p$:

$$R_{g*}(\tilde{X}_p) = \tilde{Y}_{p\cdot g}$$

Hence the isomorphism $R_{g*} : T_pP \to T_{p\cdot g}P$ sends vertical vectors to vertical vectors, thus $R_{g*}V_p = V_{p\cdot g}$. □

In contrast, as we are about to see, there exists no canonical choice of a horizontal bundle.

Definition 2.1.17. Let $P \to M$ be a principal bundle with structure group G. Then, a **horizontal subspace** is any vector subspace $H_p \subset T_pP$ such that:

$$T_pP = H_p \oplus V_p$$

A horizontal bundle is then just that, a subbundle H of TP such that:

$$TP = H \oplus V$$

Definition 2.1.18. Let H be a horizontal distribution of TP, then H is a **Ehresmann connection** on P if it is right invariant, i.e

$$R_{g*}(H_p) = H_{p\cdot g}$$

for all $p \in P$ and $g \in G$.

The right invariance of an Ehresmann connection implies that all H_p for $p \in P_x$ are determined by a singular point in the fibre. In other words, given H_{p_0} we can obtain every other horizontal space via the pushforward of right multiplication by G.

Proposition 2.1.15. *Let $\pi : P \to M$ be a principal bundle and H_p a horizontal tangent space of T_pP. Then $D_p\pi : H_p \to T_{\pi(p)}M$ is an isomorphism.*

Proof. We see that T_pP as dimension $\dim M + \dim G = n$, and that $\dim V_p = \dim G$ hence $\dim H_p = \dim M$. Since no $v \in H_p$ satisfies $D_p\pi(v) = 0$ other than the zero vector, by rank nullity $D_p\pi$ is an isomorphism. □

Example 2.1.6. Let $P = M \times G$, i.e. a trivial principal bundle over M with structure group G. We see that the vertical subspaces are given by:

$$V_{x,g} = T_{x,g}(\{x\} \times G) \cong T_gG$$

We can then choose:

$$H_{x,g} = T_{x,g}(M \times \{g\}) \cong T_xM$$

These horizontal tangent spaces define a distribution H of TP, and this distribution is clearly right invariant, so H is a connection on the trivial bundle such that:

$$H \cong \pi_M^*TM$$

where $\pi_M : M \times G \to M$, and that:

$$TP = H \oplus V \cong (\pi_M^*TM) \oplus (\pi_G^*G)$$

We shall see later that such a connection is flat, i.e. has vanishing curvature.

Example 2.1.7. Let P be a principal bundle, and g a G invariant (pseudo)-Riemannian metric on P, by which we mean:

$$R_g^*g = g$$

for all $g \in G$. Then there clearly exists a canonical choice for H, by taking the orthogonal compliment of V_p at each point $p \in P$.

The following proposition then shows that every principal bundle with a compact structure group has a connection.

Proposition 2.1.16. *Let $\pi : P \to M$ be a principal bundle with compact structure group G, then P has a G-invariant Riemannian metric.*

Proof. As P is a smooth manifold there exists a Riemannian metric s such that (P, s) is a Riemannian manifold. We define a new metric η on P by:

$$\eta_p(v, w) = \int_G s_{p\cdot g}(D_pR_g(v), D_pR_g(w))\sigma$$

where σ is the orientation form on G constructed in **Theorem 1.2.5**. The integral converges as G is compact. Since g is positive definite it follows from **Theorem 1.1.5** that:

$$\eta_p(v, v) > 0$$

for all nonzero $v \in T_pP$. It is also clear that η is symmetric, smooth, and bilinear, so h is a Riemannian metric for P. To see that that η is G invariant, take $h \in G$, and $v, w \in T_pP$, then:

$$\begin{aligned}(R_h^*\eta)_p(v, w) =& \eta_{p\cdot h}(D_pR_h(v), D_pR_h(w))\\ =& \int_G s_{p\cdot hg}(D_p(R_g \circ R_h)(v), D_p(R_g \circ R_h)(w))\sigma\\ =& \int_G s_{p\cdot g}(D_p(R_g)(v), D_p(R_g)(w))\sigma\\ =& \eta_p(v, w)\end{aligned}$$

so η is G invariant. □

Combining the preceding proposition with **Example 2.1.6** demonstrates the desired result.

As promised, we now turn to dealing with connections in a less abstract manner, as Lie algebra valued one forms on P. We begin with the following definition:

Definition 2.1.19. Let $\pi : P \to M$ be a principal bundle with structure group G. A **connection one form**, or **connection** on P is a one form $A \in \Omega(P, \mathfrak{g})$[17] on the total space P satisfying:

a) $R_g^* A = \mathrm{Ad}_{g^{-1}} \circ A$

b) $A(\tilde{X}) = X$ for all $X \in \mathfrak{g}$ where $\tilde{X}$ is the fundamental vector field on P associated to X.

A connection one form is also called a **gauge field**.

At the point $p \in P$ the connection one form is then a linear map:

$$A_p : T_pP \to \mathfrak{g}$$

In particular, as we shall see shortly, the kernel of this map determines an Ehresmann connection.

Theorem 2.1.4. *Let $P \to M$ be a principal bundle with structure group G. Then, there is a bijective correspondence between Ehresmann connections on P and connection one forms on P.*

Proof. Let H be an Ehresmann connection, then every $v \in T_pP$ can be decomposed into a horizontal part $Y_p \in H_p$ and a vertical component $\tilde{X}_p \in V_p$. We define a connection one form on P by:

$$A_p(\tilde{X}_p + Y_p) = X$$

where $X \in \mathfrak{g}$ is the element of the Lie algebra associated to the fundamental vector field $\tilde{X}$. We need to check that A satisfies the conditions of **Definition 2.1.19**. Any vector field Z can be written as $Z^V + Z^H$, where Z^V is a vertical component of Z, and Z^H is the horizontal component. Note that since V is the trivial vector bundle isomorphic to $P \times \mathfrak{g}$, any vertical vector field may be written a $(p, X(p))$, where $X : P \to \mathfrak{g}$ is a smooth function. Let $Z_p^V = (p, X(p))$, it follows that:

$$A_p(Z_p) = A(Z_p^V) = X(p)$$

so A is smooth, and in particular an element in $\Omega^1(P, \mathfrak{g})$. It is then clear that:

$$A(\tilde{X}) = X$$

Now we see that for $Z = \mathrm{Ad}_{g^{-1}}(X)$:

$$\begin{aligned}
(R_g^* A)_p(\tilde{X}_p + Y_p) =& A_{p\cdot g}(R_{g*}\tilde{X}_p + R_{g*}Y_p) \\
=& A_{p\cdot g}(\tilde{Z}_{p\cdot g} + R_{g*}Y_p) \\
=& Z \\
=& \mathrm{Ad}_{g^{-1}}(X) \\
=& \mathrm{Ad}_{g^{-1}} \circ A_p(\tilde{X}_p + Y_p)
\end{aligned}$$

so A is a connection one form on the total space.

We now want to show that given a connection one form, we can obtain obtain an Ehresmann connection. Choosing a basis for the Lie algebra $\{T_i\}$, we can write A as:

$$A = A^i \otimes T_i$$

where $A^i \in \Omega(P)$. Furthermore, we see that by definition:

$$A^i(\tilde{T}_j) = \delta^i_j$$

implying that the one forms are linearly independent at all $p \in P$. Let g be a Riemannian metric on P, and Z_i the vector fields dual to A^i under the musical isomorphism, i.e.

$$g(Z_i, W) = A^i(W)$$

[17] Here $\mathfrak{g}$ is thought of as the trivial vector bundle $P \times \mathfrak{g} \cong V$, hence connection one forms take values in the Lie algebra of G.

for all $W \in \mathfrak{X}(P)$. We then see that $\{Z_i\}$ are linearly independent at each T_pP, and span a subbundle S of TP of rank $\dim \mathfrak{g}$. At each point $p \in P$, we now define a subbundle H of TP by:

$$H_p = \ker A_p$$

It is clear that this is a distribution as it is the orthogonal compliment of S. Indeed, if $Y_p \in H_p$, we have that:

$$A^i(Y_p) = 0$$

for all i. For all $p \in P$ and all $v \in S_p$ we can write v as:

$$v = a^i Z_{pi}$$

for $a^i \in \mathbb{R}$, thus:

$$\begin{aligned} g(v, Y_p) =& a^i g(Z_{pi}, Y_p) \\ =& \sum_{i=1}^{\dim \mathfrak{g}} a^i A_p^i(Y_p) \\ =& 0 \end{aligned}$$

so Y_p is in the orthogonal compliment of S_p. Conversely, if Y_p is in the orthogonal compliment of S_p we see that:

$$g_p(Z_{pi}, Y_p) = A_p^i(Y_p) = 0$$

for all i, therefore:

$$A(Y_p) = A_p^i(Y_p)T_i = 0$$

so $Y_p \in H_p = \ker A_p$, and H is thus a distribution of TP as desired.

We now need to show that H is indeed an Ehresmann connection. We first show that H is horizontal. Let $Y_p \in H_p \cap V_p$, then $Y_p \in V_p$ hence for some fundamental vector field $\tilde{X}$ associated to $X \in \mathfrak{g}$ we have:

$$Y_p = \tilde{X}_p$$

But $Y_p \in \ker A_p$, so:

$$A_p(Y_p) = 0 = X$$

so $Y_p = 0$. Therefore $H_p \cap V_p = \{0\}$, so by rank nullity we have:

$$\dim T_pP = \dim \ker A_p + \dim \mathfrak{g} = \dim H_p + \dim V_p$$

hence:

$$T_pP = H_p \oplus V_p$$

so H is horizontal. Finally, for all $p \in P$, $Y_p \in H_p$, and $g \in G$ we have that:

$$\begin{aligned} A_{p\cdot g}(R_{g*}Y_p) =& (R_g^*A)_p(Y_p) \\ =& \mathrm{Ad}_{g^{-1}} \circ A_p(Y_p) \\ =& 0 \end{aligned}$$

so $R_{g*}Y_p \in \ker A_{p\cdot g}$, implying that $R_{g*}H_p = H_{p\cdot g}$ so H is right invariant as well. Thus H is a right invariant horizontal distribution TP, and therefore an Ehresmann connection, as desired. □

Example 2.1.8. Continuing with **Example 2.1.5**, let $P = M \times G$ and let an Ehresmann connection on P be defined by:

$$H_{x,g} = T_{x,g}(M \times \{g\}) \cong T_xM$$

It is clear that the pullback of the Maurer cartan form by the projection on to G, $A = \pi_G^* \mu_G$, then satisfies $H_p = \ker A_p$, since π_{G*} is identically zero on $T_{x,g}(M \times \{g\})$. Furthermore, the fundamental vector fields on P satisfy:

$$\begin{aligned}\tilde{X}_{(x,g)} =& \frac{d}{dt}\Big|_{t=0} (x,g) \cdot \exp(tX) \\ =& \frac{d}{dt}\Big|_{t=0} (x, g \cdot \exp(tX)) \\ =& (0, L_{g*}X) \in T_xU \oplus T_gG\end{aligned}$$

for $\tilde{X}$ associated to $X \in \mathfrak{g}$. Hence we see that:

$$\begin{aligned}(\pi_G^*\mu_G)(\tilde{X}_{(x,g)}) =& L_{g^{-1}*} \circ L_{g*}(X) \\ =& X\end{aligned}$$

Finally, we have that for all $g \in G$, and all $(x,h) \in M \times G$ that:

$$\pi_G((x,g) \cdot h) = g \cdot h = \pi_G(x,g) \cdot h$$

hence[18]:

$$\pi_G \circ R_g = R_g \circ \pi_G$$

So by **Proposition 1.2.10** we have that:

$$\begin{aligned}R_g^*(\pi_G^*\mu_G) =& \mu_G \circ (\pi_G \circ R_g)_* \\ =& (\mu_G \circ R_{g*}) \circ \pi_{G*} \\ =& \mathrm{Ad}_{g^{-1}} \circ (\mu_G \circ \pi_{G*}) \\ =& \mathrm{Ad}_{g^{-1}} \circ (\pi_G^*\mu_G)\end{aligned}$$

so $\pi_G^*\mu_G$ is indeed a connection one form on $M \times G$.

Example 2.1.9. Recall from **Example 2.1.1** that $\mathbb{S}^3$ is a $\mathbb{S}^1$ principal bundle over $\mathbb{S}^2$. We want to find a connection one form on $\mathbb{S}^3$. We first identify the Lie algebra of $\mathbb{S}^1$ with $i\mathbb{R}$, and recall that the tangent spaces of $\mathbb{S}^3$ are defined by:

$$T_{(z_1,z_2)}\mathbb{S}^3 = \{(X_1, X_2) \in \mathbb{C}^2 : \mathrm{Re}(\bar{z}_1X_1 + \bar{z}_2X_2) = 0\}$$

Define one forms on $\mathbb{S}^3$ by:

$$\begin{aligned}\alpha_j(X_0, X_1) =& X_j \\ \bar{\alpha}_j(X_0, X_1) =& X_j^*\end{aligned}$$

We claim that:

$$A_{(z_1,z_2)} = \frac{1}{2}\left(\bar{z}_1\alpha_1 - z_1\bar{\alpha}_1 + \bar{z}_2\alpha_2 - z_2\bar{\alpha}_2\right)$$

is a connection one form. First note that for all $(X_1, X_2) \in T_{(z_1,z_2)}\mathbb{S}^3$ we have:

$$A_{(z_1,z_2)}(X_1, X_2) = \frac{1}{2}\left(\bar{z}_1X_1 - z_1\bar{X}_1 + \bar{z}_2X_2 - z_2\bar{X}_2\right)$$

of which the complex conjugate is clearly equal to $-A_{(z_1,z_2)}(X_0, X_1)$ so A takes values in $i\mathbb{R}$. Secondly, since $\mathbb{S}^1$ is abelian, we have that $\mathrm{Ad}_{g^{-1}} = \mathrm{Id}$ hence we need to show that:

$$R_g^*(A) = A$$

for all $g \in \mathbb{S}^1$. Let $g \in \mathbb{S}^1$, then:

$$(R_g^*A)_{(z_1,z_2)}(X_1, X_2) = A_{(z_1,z_2)\cdot g} \circ R_{g*}(X_1, X_2)$$

[18] Once again these two R_g maps are technically not the same.

Note that:

$$\begin{aligned}R_{g*}(X_1, X_2) =& D_{(z_1,z_2)}R_g(X_1, X_2)\\ =& \frac{d}{dt}\Big|_{t=0}(z_1(t), z_2(t)) \cdot g\end{aligned}$$

for some curves $z_1, z_2 : I \to \mathbb{C}$ satisfying:

$$|z_1(t)|^2 + |z_2(t)|^2 = 1$$

for all $t \in I$, and:

$$(\dot{z}_1(0), \dot{z}_2(0)) = (X_1, X_2) \in T_{(z_1,z_2)}\mathbb{S}^3$$

Thus:

$$R_{g*}(X_1, X_2) =(X_1 \cdot g, X_2 \cdot g)$$

Furthermore, any $g \in \mathbb{S}^1$ can be written as $e^{i\theta}$ for some $\theta \in \mathbb{R}$, hence:

$$\begin{aligned}A_{(z_1,z_2)\cdot g} \circ R_{g*}(X_1, X_2) =& A_{(z_1\cdot e^{i\theta}, z_2\cdot e^{i\theta})}(X_1 \cdot e^{i\theta}, X_2 \cdot e^{i\theta})\\ =& \frac{1}{2}\left(\bar{z}_1 e^{-i\theta}X_1 e^{i\theta} + z_1 e^{i\theta}\bar{X}_1 e^{-i\theta} + \bar{z}_2 e^{-i\theta}X_2 e^{i\theta} - z_2 e^{i\theta}\bar{X}_1 e^{-i\theta}\right)\\ =& \frac{1}{2}\left(\bar{z}_1 X_1 - z_1\bar{X}_1 + \bar{z}_2 X_2 - z_2\bar{X}_2\right)\\ =& A_{(z_1,z_2)}(X_1, X_2)\end{aligned}$$

so:

$$(R_g^* A) = A$$

as desired. Finally, we need to show that:

$$A(\tilde{Y}) = Y$$

for $\tilde{Y} \in V_p$. Let $Y = i\theta \in i\mathbb{R}$, we then see that:

$$\begin{aligned}\tilde{Y}_{(z_1,z_2)} =& \frac{d}{dt}\Big|_{t=0}(z_1, z_2)e^{i\theta t}\\ =& (i\theta z_1, i\theta z_2)\end{aligned}$$

Hence:

$$\begin{aligned}A_{(z_1,z_2)}(\tilde{Y}) =& A_{(z_1,z_1)}(i\theta z_1, i\theta z_2)\\ =& \frac{1}{2}\left(2|z_1|^2 i\theta + 2|z_2|^2 i\theta\right)\\ =& i\theta\left(|z_1|^2 + |z_2|^2\right)\\ =& i\theta\\ =& Y\end{aligned}$$

so A is a connection one form as desired.

Connection one forms can also be viewed as local Lie algebra valued one forms on the base by pulling A back to an open set U via a local gauge.

Definition 2.1.20. Let $P \to M$ be a principal bundle with structure group G, A a connection one form, and $s : U \to P_U$ a local gauge. Then we define a **local connection one form**, some times called a local **gauge field** by:

$$A_s = s^*A = A \circ Ds$$

If we have a local coordinate frame $\{\partial_\mu\}$ for $U \subset M$, then we set:

$$A_\mu = A_s(\partial_\mu)$$

In addition we can choose a basis $\{T_a\}$ for $\mathfrak{g}$ and write:

$$A_\mu = A_\mu^a T_a$$

hence we can write A_s as:

$$A_s = A_\mu^a T_a \otimes dx^\mu$$

where $A_\mu^a \in C^\infty(U)$. In particular, if ϕ is the local trivialization corresponding to s, with inverse given by:

$$\begin{aligned} \phi^{-1} : U \times G &\longrightarrow P_U \\ (x, g) &\longmapsto s(x) \cdot g = \Phi(s(x), g) \end{aligned}$$

then pulling A back by ϕ^{-1} yields for $(X, Y) \in T_{(x,g)}(U \times G)$:

$$\begin{aligned} ((\phi^{-1})^* A)_{(x,g)}(X, Y) =& A \circ D_{(s(x),g)}\Phi(D_x s(X), Y)) \\ =& A\left(D_{s(x)}R_g(D_x s(X))\right) + A\left(\widetilde{\mu_G(Y)}_{s(x)\cdot g}\right) \\ =& (R_g^* A)_{s(x)}(D_x s(X)) + \mu_G(Y) \\ =& \mathrm{Ad}_{g^{-1}} \circ A_{s(x)}(D_x s(X)) + \mu_G(Y) \\ =& \mathrm{Ad}_{g^{-1}} \circ A_s(X) + \mu_G(Y) \end{aligned}$$

Hence the connection form in a local trivialization is given by:

$$((\phi^{-1})^* A)_{(x,g)} = \mathrm{Ad}_{g^{-1}} \circ (\pi_M^* A_s) + \pi_G^* \mu_G \tag{2.1.5}$$

If we wish to examine matrix Lie groups, we can instead write this as:

$$((\phi^{-1})^* A)_{(x,g)} = g^{-1}(A_s)g + g^{-1}(dg) \tag{2.1.6}$$

where we have dropped the pullbacks of the projection map to avoid clutter. It should be clear that A_s takes is only non zero on the $T_x U$ subspace of $T_{(x,g)}(U \times G)$.

Finally, we want to see that the set of connection is an affine space. We need the following definition:

Definition 2.1.21. Let $\omega \in \Omega^k(P, \mathfrak{g})$ be a differential k form on P with values in $\mathfrak{g}$. Then, ω is said is to be a **of type** Ad if for all $g \in G$:

$$R_g^* \omega = \mathrm{Ad}_{g^{-1}} \circ \omega$$

and is **horizontal** if:

$$\omega_p(X_1, \cdots, X_k) = 0$$

when at least one $X_i \in V_p$. We denote the set of all horizontal k-forms of type Ad by on P with values in $\mathfrak{g}$ by $\Omega^k_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$.

Proposition 2.1.17. *Let $P \to M$ be a principal G bundle, and A, A' connection one forms P. Then:*

$$A - A' \in \Omega^1_{hor}(P, \mathfrak{g})^{Ad}$$

Moreover, for any $\omega \in \Omega^1_{hor}(P, \mathfrak{g})^{Ad}$, we have that $A + \omega$ is a connection on P.

Proof. We see that if A and A' are connections, then for any vertical vector field $\tilde{X}$ associated to $X \in \mathfrak{g}$:

$$A(\tilde{X}_p) - A'(\tilde{X}_p) = X - X = 0$$

hence $A - A'$ is horizontal. Furthermore for all $g \in G$:

$$\begin{aligned} R_g^*(A - A') =& R_g^* A - R_g^* A' \\ =& \mathrm{Ad}_{g^{-1}} \circ A - \mathrm{Ad}_{g^{-1}} \circ A' \\ =& \mathrm{Ad}_{g^{-1}} \circ (A - A') \end{aligned}$$

so $A - A' \in \Omega^1_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$.

Let $\omega \in \Omega^1_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$, then for any vertical vector field $\tilde{X}$ associated to $X \in \mathfrak{g}$, we have:

$$\begin{aligned} (A + \omega)(\tilde{X}_p) =& A(\tilde{X}_p) + \omega(\tilde{X}_p) \\ =& X \end{aligned}$$

Furthermore, for all $g \in G$:

$$R_g^*(A + \omega) = \mathrm{Ad}_{g^{-1}} \circ (A + \omega)$$

so $A + \omega$ is a new connection one form on P. □

Corollary 2.1.4. *The set of all connection one forms on P is an affine space over $\Omega^1_{hor}(P, \mathfrak{g})^{Ad}$. A base point is given by any connection one form on P.*

2.1.5 Gauge Transformations

As discussed earlier, a local gauge transformation amounts to a change of local trivialization, however, there is also a notion of a global gauge transformation, i.e. a specific type of diffeomorphism on the total space P. In this section, we discuss the consequences of both local and global gauge transformations, beginning with the latter. We start with the following definition:

Definition 2.1.22. Let $\pi : P \to M$ be a principal bundle with structure group G. A **global gauge transformation** is a **bundle automorphism**, i.e. a diffeomorphism $f : P \to P$ such that the following hold:

a) $\pi \circ f = \pi$

b) $f(p \cdot g) = f(p) \cdot g$.

In other words, f preserves the fibres and is G equivariant. The set of global bundle automorphism is denoted by $\mathscr{G}(P)$.

Proposition 2.1.18. *The set of a global bundle automorphisms forms a subgroup of Diff(P).*

Proof. By definition the bundle automorphisms are a subset of Diff(P). Recall the the group action of Diff(P) is composition, we then see that for all $f, g \in \mathscr{G}(P)$, and all $p \in P$, $g \in G$:

$$g \circ f(p \cdot g) = g(f(p) \cdot g) = (g \circ f(p)) \cdot g$$

hence $g \circ f$ is G equivariant. Furthermore:

$$\begin{aligned} \pi \circ (g \circ f) =& (\pi \circ g) \circ f \\ =& \pi \circ f \\ =& \pi \end{aligned}$$

so $g \circ f$ preserves the fibres, demonstrating that $g \circ f \in \mathscr{G}(P)$ for all $g, f \in \mathscr{G}(P)$. Furthermore let $f \in \mathscr{G}(P)$, we then have a $f^{-1} \in \mathrm{Diff}(P)$ satisfying:

$$f^{-1} \circ f = \mathrm{Id}_P$$

Note that for $p \in P$ there exists a $q \in P$ such that:

$$f^{-1}(p) = q \text{ and } f(q) = p$$

It follows that q and p are in the same fibre hence $\pi \circ f^{-1} = \pi$. Furthermore, for some $g \in G$ we have that $p = q \cdot g$. We see that:

$$f(q \cdot g) = f(q) \cdot g = p \cdot g$$

Applying f^{-1} to the left most side and right most side yields:

$$f^{-1}(p \cdot g) = q \cdot g = f^{-1}(p) \cdot g$$

hence f^{-1} is G equivariant as well, so $f^{-1} \in \mathscr{G}(P)$. Clearly $\mathscr{G}(P)$ contains the identity, so $\mathscr{G}(P)$ is indeed a subgroup of $\mathrm{Diff}(P)$. $\square$

We often call $\mathscr{G}(P)$ the **gauge group** of P. Furthermore, we can then think of a **local gauge transformation** as an element of $\mathscr{G}(P_U)$, where P_U is the trivial principal bundle over an open set $U \subset M$, as that will surmount to a local change of trivialization. We would now like to take an alternative approach to gauge transformations, by viewing them as G valued maps on P.

Definition 2.1.23. Let $P \to M$ be a principal bundle with structure group G. We denote by $C^\infty(P, G)^G$ the following set of maps $P \to G$:

$$C^\infty(P, G)^G = \{\sigma : P \to G \text{ smooth} : \sigma(p \cdot g) = g^{-1}\sigma(p)g\}$$

The set is a group under pointwise multiplication:

$$(\sigma' \cdot \sigma)(p) = \sigma'(p) \cdot \sigma(p)$$

where the neutral element is the constant map on P with value $e \in G$.

Proposition 2.1.19. *The map:*

$$\begin{aligned} \mathscr{G}(P) &\longrightarrow C^\infty(P, G)^\infty \\ f &\longmapsto \sigma_f \end{aligned}$$

with σ_f defined by:

$$f(p) = p \cdot \sigma_f(p)$$

is a well defined group isomorphism.

Proof. Since p and $f(p)$ are in the same fibre, there exists a unique g such that $f(p) = p \cdot g$. Define σ_f pointwise by:

$$\sigma_f(p) = g$$

Let (U, ϕ) be a local bundle chart containing the point $\pi(p) = x$, then we see that for some $h \in G$:

$$\phi(p) = (x, h)$$

since ϕ is a G equivariant map:

$$\phi(f(p)) = \phi(p \cdot g) = (x, h \cdot g)$$

hence we see that locally:

$$\sigma_f(p) = (\mathrm{pr}_G \circ \phi(p))^{-1} \cdot (\mathrm{pr}_G \circ \phi(f(p)))$$

which is clearly smooth. Since smoothness is a local criterion, it follows that σ_f is a smooth map $P \to G$. For all $g \in G$ we have that:

$$\begin{aligned} f(p \cdot g) =& f(p) \cdot g \\ =& p \cdot \sigma_f(p) \cdot g \end{aligned}$$

However:

$$f(p \cdot g) = p \cdot g \cdot \sigma_f(p \cdot g)$$

hence:

$$p \cdot g \cdot \sigma_f(p \cdot g) = p \cdot \sigma_f(p) \cdot g$$

Since the action is free we obtain:

$$g \cdot \sigma_f(p \cdot g) = \sigma_f(p) \cdot g \Longrightarrow \sigma_f(p \cdot g) = g^{-1} \cdot \sigma_f(p) \cdot g$$

so $\sigma_f \in C^\infty(P, G)^G$.

An inverse for this map is obtained by:

$$\begin{aligned} C^\infty(P, G)^G &\longrightarrow \mathscr{G}(P) \\ \sigma &\longmapsto f_\sigma \end{aligned}$$

where f_σ is defined by:

$$f_\sigma(p) = p \cdot \sigma_f(p)$$

It is clear that f_σ is a smooth, and preserves the fibres, hence it is a bundle map. We further see that f_σ is G equivariant as:

$$\begin{aligned} f_\sigma(p \cdot g) =&(p \cdot g) \cdot \sigma(p \cdot g) \\ =&(p \cdot \sigma(p)) \cdot g \\ =&f(p) \cdot g \end{aligned}$$

Finally it is clear that:

$$f_\sigma^{-1} = f_{\sigma^{-1}}$$

as:

$$\begin{aligned} f_{\sigma^{-1}} \circ f_\sigma(p) =&f_{\sigma^{-1}}(p \cdot \sigma(p)) \\ =&(p \cdot \sigma(p)^{-1}) \cdot \sigma(p) \\ =& p \end{aligned}$$

so $f \in \mathscr{G}(P)$. It is clear that these two maps are inverses of one another, hence we need only check that:

$$\sigma_{f \circ f'} = \sigma_f \cdot \sigma_{f'} \tag{2.1.7}$$

We see that:

$$\begin{aligned} f \circ f'(p) =&f(p \cdot \sigma_{f'}(p)) \\ =&f(p) \cdot \sigma_{f'}(p) \\ =&p \cdot (\sigma_f(p) \cdot \sigma_{f'}(p)) \end{aligned}$$

so (2.1.7) holds and $\mathscr{G}(P) \cong C^\infty(P, G)^G$ as groups. □

In physics, gauge transformations are often viewed as maps from the base manifold to the structure group G. We define these types of gauge transformations below:

Definition 2.1.24. Let $\pi : P \to M$ be a principal bundle with structure group G. A **physical gauge transformation** is a smooth map $\tau : U \to G$, defined on an open subset U. We denote the set of physical gauge transformations by $C^\infty(U, G)$, and note that it is a group under pointwise multiplication.

We see that $\mathscr{G}(P_U) \cong C^\infty(P_U, G)^G$ as groups, we would then like to show that $C^\infty(P_U, G)^G \cong C^\infty(U, G)$ to demonstrate that these are all equivalent notions of a gauge transformation.

Proposition 2.1.20. *Let $s : U \to P$ be a local section, then s determines a group isomorphism:*

$$\begin{aligned} C^\infty(P_U, G)^G &\longrightarrow C^\infty(U, G) \\ \sigma &\longmapsto \tau_\sigma = \sigma \circ s \end{aligned}$$

with inverse given by:

$$\begin{aligned} C^\infty(U, G) &\longrightarrow C^\infty(P_U, G)^G \\ \tau &\longmapsto \sigma_\tau \end{aligned}$$

where:

$$\sigma_\tau(s(x) \cdot g) =g^{-1}\tau(x)g, \forall x \in U, g \in G \tag{2.1.8}$$

Proof. It is clear that τ_σ is a smooth map $U \to G$, and thus an element of $C^\infty(U, G)$. Furthermore, we see that σ_τ is smooth, as on P_U we have that (2.1.9) is equivalent to the composition:

$$\sigma_\tau(p) = (\mathrm{pr}_G \circ \phi(p))^{-1} \cdot \tau(\mathrm{pr}_M \circ \phi(p)) \cdot (\mathrm{pr}_G \circ \phi(p))$$

where ϕ is the bundle chart corresponding to the section s, hence $\sigma_\tau \in C^\infty(P_U, G)$. By (2.1.9):

$$\begin{aligned}\tau_{\sigma_\tau}(x) =& \sigma_\tau \circ s(x)\\ =& \tau(x)\end{aligned}$$

Furthermore, for all $x \in U, g \in G$:

$$\begin{aligned}\sigma_{\tau_\sigma}(s(x) \cdot g) =& \sigma_{\sigma \circ s}(s(x) \cdot g)\\ =& g^{-1}\sigma(s(x))g\\ =& \sigma(s(x) \cdot g)\end{aligned}$$

so:

$$\sigma_{\tau_\sigma} = \sigma \qquad \text{and} \qquad \tau_{\sigma_\tau} = \tau$$

thus the maps are inverses of one another. Finally, the map is a group isomorphism as for $\sigma, \sigma' \in C^\infty(P_U, G)^G$, we have:

$$(\sigma \cdot \sigma')(p) = \sigma(p) \cdot \sigma'(p)$$

hence:

$$\begin{aligned}\tau_{\sigma \cdot \sigma'}(s(x)) =& (\sigma \cdot \sigma') \circ s(x)\\ =& \sigma(s(x)) \cdot \sigma'(s(x))\\ =& \tau_\sigma \cdot \tau_{\sigma'}\end{aligned}$$

so the groups are isomorphic as desired. □

Gauge transformations on the principal bundle induce gauge transformations on associated vector bundles. In the following two theorems we examine two cases of this induced transformation: one corresponding to global gauge transformations, and another corresponding to physical gauge transformations.

Theorem 2.1.5. *Let $P \to M$ be a principal bundle with structure group G, and $E = P \times_\rho V$ a vector bundle associated to P. The group of bundle automorphisms $\mathscr{G}(P)$ then acts on E through bundle isomorphisms via:*

$$\begin{aligned}\mathscr{G}(P) \times E &\longrightarrow E\\ (f, [p, v]) &\longmapsto f \cdot [p, v] = [f(p), v] = [p \cdot \sigma_f(p), v]\end{aligned}$$

Proof. We need to see that the action is well defined, let $[q, w] = [p, v]$, then for some $g \in G$:

$$\begin{aligned}f \cdot [q, w] =& [f(q), w]\\ =& [f(p \cdot g), \rho(g)^{-1}v]\\ =& [f(p) \cdot g, \rho(g)^{-1}v]\\ =& [f(p), v]\\ =& f \cdot [p, v]\end{aligned}$$

so the action is well defined. It is then clear that f restricts to a vector space isomorphism on the fibres of E, and is thus a vector bundle isomorphism. □

Theorem 2.1.6. *Let $s : U \to P$ be a local gauge and $\Phi : U \to E$ a local section. We write the section with respect to the gauge as:*

$$\Phi(x) = [s(x), \phi(x)]$$

for some smooth map $\phi : U \to V$. Suppose f is a local bundle automorphism, and $\tau_f : U \to G$ the associated physical gauge transformation. Then:

$$(f \cdot \Phi)(x) = [s(x), \rho(\tau_f(x))\phi(x)]$$

Proof. We see that:

$$\begin{aligned}(f \cdot \Phi)(x) =&[f(s(x)), \phi(x)]\\ =&[s(x) \cdot \sigma_f(s(x)), \phi(x)]\\ =&[s(x) \cdot \tau_f(x), \phi(x)]\\ =&[s(x), \rho(\tau_f(x))\phi(x)]\end{aligned}$$

□

Now we would like to see how connections change under gauge transformations. We begin by noting that for any physical gauge transformation $h : U \to G$, and any local $s : U \to P$, that $s \cdot h$ is another local gauge. With this in mind, We se that if $\phi : P_U \to U \times G$ is a bundle chart corresponding to s, then the map:

$$\begin{aligned}\phi_h^{-1} : U \times G &\longrightarrow P_U\\ (x, g) &\longmapsto (s \cdot h)(x) \cdot g\end{aligned}$$

is the inverse map of the bundle chart corresponding $s \cdot h$. From this, we see that the change in trivialization obtained by the gauge transformation h amounts to the substitution $g \mapsto h(x)g$. Recalling (2.1.6) we see that:

$$\begin{aligned}((\phi_h^{-1})^*A)_{(x,g)} =&(h(x)g^{-1})A_s(h(x)g) + (h(x)g)^{-1}d(h(x)g)\\ =&g^{-1}h(x)^{-1}A_sh(x)g + g^{-1}(h(x)^{-1}dh(x))g + g^{-1}dg\\ =&g^{-1}A_{sh}g^{-1} + g^{-1}dg\end{aligned}$$

where:

$$A_{sh} = h(x)^{-1}A_sh(x) + h(x)^{-1}dh(x)$$

We see that A_{sh} is the local connection one form obtained from A_s by a gauge transformation. We make this argument precise with the following theorem:

Theorem 2.1.7. *Let $P \to M$ be a principal G bundle, A be a connection one form on P. Furthermore, let $s_i : U_i \to P_{U_i}$, and $s_j : U_j \to P_{U_j}$ be two local gauges related on the overlap, $U_i \cap U_j \neq \emptyset$, by the physical gauge transformation $h : U_i \cap U_j \to G$, i.e.*

$$s_i(x) = s_j(x) \cdot h(x)$$

Then:

$$A_{s_i} = Ad_{h^{-1}} \circ A_{s_j} + h^*\mu_G$$

Or if G is a matrix Lie group then:

$$A_{s_i} = h^{-1}A_{sj}h + h^{-1}dh$$

Proof. First, for brevity set $A_i = A_{s_i}$, and $A_j = A_{s_j}$. Then we see that:

$$\begin{aligned}A_{s_i} =&s_i^*A\\ =&(\Phi(s_j, h))^*A\end{aligned}$$

where Φ is the right group action of G on P. Let $X \in T_x(U_i \cap U_j)$, then:

$$\begin{aligned}A_i(X) =&A(D_{(s_j(x),h(x))}\Phi(D_xs_j(X), D_xh(X)))\\ =&A(D_{s(x)}R_h(D_xs_j(X))) + A\left(\mu_G(\widetilde{D_xh}(X))_{s(x)\cdot h}\right)\\ =&\mathrm{Ad}_{h^{-1}} \circ A(D_xs_j(X)) + \mu_G(D_xh(X))\\ =&\mathrm{Ad}_{h^{-1}} \circ A_j(X) + h^*\mu_G(X)\end{aligned}$$

Thus:

$$A_{s_i} = \mathrm{Ad}_{h^{-1}} \circ A_{s_j} + h^*\mu_G$$

as desired. Furthermore, if G is a matrix Lie group then:

$$\mathrm{Ad}_{h^{-1}} \circ A_j = h^{-1} A_j h$$

and:

$$\mu_G(D_x h(X)) = h^{-1} \cdot (D_x h(X)) = h^{-1} dh(X)$$

since for all $g \in G$, and all $v \in T_g G$, $\mu_G(v) = g^{-1} \cdot v$. Therefore, for matrix Lie groups

$$A_{s_i} = h^{-1} A_{s_j} h + h^{-1} dh$$

as desired. □

Via a similar argument we also obtain the following theorem for global gauge transformation.

Theorem 2.1.8. *Let $P \to M$ be a principal bundle with structure group G, A a connection one form, and f a global bundle automorphism. Then:*

$$f^* A = Ad_{\sigma_f^{-1}} \circ A + \sigma_f^* \mu_G$$

Proof. We see that:

$$f = \Phi(p, \sigma_f(p))$$

hence for all $X \in T_p P$:

$$\begin{aligned}
(f^* A)(X) =& A(D\Phi(X, D_p \sigma_f(X))) \\
=& A(D_p R_{\sigma_f}(X)) + A(\mu_G(\widetilde{D_p \sigma_f}(X))_{p \cdot \sigma_f(p)}) \\
=& \mathrm{Ad}_{\sigma_f^{-1}} \circ A(X) + \mu_G(D_p \sigma_f(X)) \\
=& \mathrm{Ad}_{\sigma_f^{-1}} \circ A(X) + \sigma_f^* \mu_G(X)
\end{aligned}$$

Therefore:

$$f^* A = \mathrm{Ad}_{\sigma_f^{-1}} \circ A + \sigma_f^* \mu_G$$

as desired. □

We end with an example from physics.

Example 2.1.10. We begin by guessing that Electromagnetism is a $U(1)$ gauge theory over $\mathbb{R}^{1,3}$[19] , and identify the Lie algebra of $U(1)$ with $i\mathbb{R}$. Let V, and $\mathbf{M}$ denote the electric and magnetic potentials respectively. In the introduction we asserted that for any $\lambda \in C^\infty(\mathbb{R}^4)$, the new potentials:

$$V' = V - \frac{\partial \lambda}{\partial t} \qquad \text{and} \qquad \mathbf{M}' = \mathbf{M} + \nabla \lambda$$

gave the same physical fields. Such a transformation can be better encoded by defining a four potential on $\mathbb{R}^{1,3}$ by:

$$A^i \partial_i = (V, M_x, M_y, M_z)$$

Then, under the musical isomorphism we have:

$$A_i dx^i = -V dt + M_i dx^i$$

so the transformation is given by:

$$A' = A + d\lambda \tag{2.1.9}$$

[19] i.e $\mathbb{R}^4$ with the a minkowski metric of signature $(- + ++)$

Now since $P = \mathbb{R}^{1,3} \times U(1)$ is a trivial principal bundle, let iA_j be a connection one form on the base manifold corresponding to a global section s_j. In coordinates we write that:

$$iA_j = i\left(-V dt + M_i dx^i\right)$$

for some smooth functions V, M^i on $\mathbb{R}^{1,3}$. Let s_i be another global section, then:

$$s_i = s_j \cdot e^{i\lambda(x)}$$

for some $\lambda : \mathbb{R}^{1,3} \to \mathbb{R}$. By **Theorem 2.1.6** we obtain:

$$\begin{aligned} iA_i =& e^{-i\lambda(x)} A_j e^{i\lambda(x)} + e^{-i\lambda(x)} de^{i\lambda(x)} \\ =& i\left(A_j + d\lambda\right) \end{aligned}$$

hence:

$$A_i = A_j + d\lambda$$

which exactly matches (2.1.10), so from our work on the general case of gauge transformations we have obtained the quintessential example of a gauge transformation in physics. In Chapter 3.1 we will continue to look at Electromagnetism as a classical gauge theory from this perspective.

2.1.6 Curvature

Let $P \to M$ be a principal G bundle, and A a connection one form on P. We have that a horizontal subbundle $H \subset TP$ defined by the kernel of A such that:

$$TP = H \oplus V$$

Let π_H denote the projection map:

$$\pi_H : TP \longrightarrow V$$

We then define curvature as follows:

Definition 2.1.25. The two form $F \in \Omega^2(P, \mathfrak{g})$ defined by:

$$F_p(X, Y) = dA(\pi_H(X), \pi_H(Y)) \tag{2.1.10}$$

for all $X, Y \in T_pP$, and $p \in P$ is called the **curvature two form** or **curvature** of A. We sometimes denote the curvature by F^A to emphasize dependence on A.

The curvature of a connection as the following properties:

Proposition 2.1.21. *Let $P \to M$ be a principal G bundle, A a connection one form on P, and F the curvature of said connection. Then then the following identities hold:*

a) $R_g^* F = Ad_{g^{-1}} \circ F$

b) $\tilde{X} \lrcorner F = 0$ *for all fundamental vector fields* $\tilde{X}$.

Proof. As both H_p and V_p are right invariant, we see that:

$$R_{g*} \circ \pi_H = \pi_H \circ R_{g*}$$

Therefore for $X, Y \in T_pP$:

$$\begin{aligned} (R_g^* F)_p(X, Y) =& dA\left(\pi_H \circ R_{g*}(X), \pi_H \circ R_{g*}(Y)\right) \\ =& d(R_g^* A)_p(\pi_H(X), \pi_H(Y)) \\ =& d(\mathrm{Ad}_{g^{-1}} \circ A)(\pi_H(X), \pi_H(Y)) \\ =& \mathrm{Ad}_{g^{-1}} \circ F_p(X, Y) \end{aligned}$$

Furthermore, for any fundamental vector field $\tilde{X}$, we have that $\tilde{X}_p \in V_p$, so $\pi_H(\tilde{X}_p) = 0$. Therefore, for all $Y \in T_pP$:

$$\begin{aligned} \tilde{X} \lrcorner F(Y) =& F(\tilde{X}, Y) \\ =& dA(0, Y) \\ =& 0 \end{aligned}$$

so $\tilde{X} \lrcorner F$ is identically zero. $\square$

The proposition above demonstrates that F is a horizontal two form of type Ad, hence $F \in \Omega^2_{\text{hor}}(P,\mathfrak{g})^{\text{Ad}}$.

Despite the simplicity of (2.1.10), this form of the curvature is not so useful in practice. Ideally, we would like to be able to write F^A solely in terms of A and the elementary operations on forms we have encountered. As it turns out, we will be able to do this once we define a new wedge type product for differential forms valued in a Lie algebra.

Definition 2.1.26. Let $\omega \in \Omega^k(P,\mathfrak{g})$ and $\eta \in \Omega^l(P,\mathfrak{g})$, then we define:

$$[\omega,\eta]_p(X_1,\dots,X_{k+l}) = \frac{1}{k!l!}\sum_{\sigma\in S_{k+l}} \text{sgn}(\sigma)\left[\omega(X_{\sigma(1),\dots},X_{\sigma(k)}),\eta(X_{\sigma(k+1)},\dots,X_{\sigma(k+l)})\right]$$

which lies in $\Omega^{k+l}(P,\mathfrak{g})$.

By choosing a basis $\{T_a\}$ for $\mathfrak{g}$, ω and η can be written as:

$$\omega = \omega^a \otimes T_a, \qquad \eta = \eta^b \otimes T_b$$

Hence:

$$\begin{aligned}
[\omega,\eta]_p(X_1,\dots,X_{k+l}) =& \frac{1}{k!l!}\sum_{\sigma\in S_{k+l}} \text{sgn}(\sigma)\left[\omega^a(X_{\sigma(1),\dots},X_{\sigma(k)})T_a,\eta^b(X_{\sigma(k+1)},\dots,X_{\sigma(k+l)})T_b\right]\\
=&\frac{1}{k!l!}\sum_{\sigma\in S_{k+l}} \text{sgn}(\sigma)\omega^a(X_{\sigma(1),\dots},X_{\sigma(k)})\eta^b(X_{\sigma(k+1)},\dots,X_{\sigma(k+l)})\,[T_a,T_b]\\
=&\omega^a\wedge\eta^b(X_1,\cdots,X_{k+1})\,[T_a,T_b]
\end{aligned}$$

so we obtain that:

$$[\omega,\eta] = \omega^a\wedge\eta^b\otimes[T_a,T_b]$$

which implies:

$$[\omega,\eta] = (-1)^{kl+1}[\eta,\omega]$$

In particular for one forms we have that:

$$[\omega,\eta](X_1,X_2) = [\omega(X_1),\eta(X_2)] - [\omega(X_2),\eta(X_1)]$$

so:

$$[\omega,\omega](X_1,X_2) = 2[\omega(X_1),\omega(X_2)] \tag{2.1.11}$$

hence $[\cdot,\cdot]$ is not identically zero like the bracket is for a regular Lie algebra, or like the wedge is for a regular $2k+1$ form. With (2.1.11) in mind, we seek to prove the following theorem, also known as the structure equation:

Theorem 2.1.9. *Let $P \to M$ be a principal G bundle, and A a connection one form on P. The curvature two form then satisfies:*

$$F^A = dA + \frac{1}{2}[A,A] \tag{2.1.12}$$

We first need the following lemmas:

Lemma 2.1.2. *Let $\tilde{X}$ be a fundamental vector field, i.e. a vertical vector field on P, and Y a horizontal vector field on P. Then $[\tilde{X},Y]$ is horizontal.*

Proof. We know that $\tilde{X}$ is associated to some $X \in \mathfrak{g}$, so the flow of X is given by $R_{\exp(tX)}$. We then see by the definition of the Lie derivative that for all $p \in P$:

$$[\tilde{X},Y]_p = \frac{d}{dt}\Big|_{t=0} R_{\exp(-tX)*}Y_{p\cdot\exp(tX)}$$

since $Y_{p\cdot\exp(tX)} \in H_p$, and H is right invariant, we see that $[\tilde{X},Y]_p \in H_p$. □

Lemma 2.1.3. *Let X be a vector field on M, then there exists a unique horizontal vector field X^H on P satisfying $D_p\pi(X_p^H) = X_{\pi(p)}$ for all $p \in P$. In particular, for all $p \in P$ any $X \in H_p$ can be extended to a horizontal vector field X^H on P_U such that $X_p^H = X$.*

Proof. Let X be a vector field on M, then since $D_p\pi : H_p \to T_{\pi(p)}M$ is an isomorphism the following assignment on H is uniquely determined:

$$X_p^H = (D_p\pi)^{-1}(X_{\pi(p)})$$

We need to show that X^H is a smooth vector field. Let $\pi(p) = x$, then for some bundle chart (U, ϕ) such that $\phi(p) = (x, g)$, we have that $D_p\phi : T_pP \to T_xU \oplus T_gG$ is an isomorphism. We want to find a vector Z field on $U \times G$ such that:

$$D_{(x,g)}\pi_U(Z_{(x,g)}) = X_x$$

We define a map Z by:

$$\begin{aligned} Z : (U \times G) &\longrightarrow (\pi_U^* TU) \\ (x, g) &\longmapsto ((x, g), X_x) \in (\pi_U^* TU)_{(x,g)} \end{aligned}$$

which is clearly smooth, and takes values in $(\pi_U^* TU) \subset (\pi_U^* TU) \oplus (\pi_G^* TG) \cong T(U \times G)$, and thus Z is clearly a smooth section of $T(U \times G)$. Furthermore, we see that for any $v \in (\pi_U^* TU)_{(x,g)} = T_xM$:

$$v = (v, 0) \in T_xM \oplus T_gG$$

Then for some smooth curve $\gamma : I \to U \times G$, we have smooth curves $x : I \to U$ and $g : I \to G$ such that:

$$\gamma(t) = (x(t), g(t))$$

Let γ satisfy:

$$\gamma(t) = (x, g) \qquad \text{and} \qquad \dot\gamma(t) = (\dot x(0), \dot g(0)) = (v, 0)$$

So we have that:

$$\begin{aligned} D_{(x,g)}\pi_U(v) =& \frac{d}{dt}\Big|_{t=0} \pi(\gamma(t)) \\ =& \frac{d}{dt}\Big|_{t=0} x(t) \\ =& v \end{aligned}$$

Hence for all $(x, g) \in U \times G$ we have that:

$$D_{(x,g)}\pi_U(Z_{(x,g)}) = X_x$$

as desired. Since ϕ is a diffeomorphism it follows that there exists a unique vector field Y such that $(\phi_* Y) = Z$, hence for all $p \in P_U$:

$$D_p\phi(Y_p) = Z_{\phi(p)}$$

So:

$$\begin{aligned} D_{\phi(p)}\pi_U \circ D_p\phi(Y_p) =& D_{\phi(p)}\pi_U(Z_{\phi(p)}) \\ =& X_{\pi(p)} \end{aligned}$$

However:

$$\begin{aligned} D_{\phi(p)}\pi_U \circ D_p\phi(Y_p) =& D_p(\pi_U \circ \phi)(Y_p) \\ =& D_p\pi(Y_p) \end{aligned}$$

hence:

$$D_p\pi(Y_p) = X_{\pi(p)}$$

Thus the horizontal component of Y must be equal to X^H, as:

$$\begin{aligned} D_p\pi(X_p^H) =& D_p\pi \circ (D_p\pi)^{-1}(X_x) \\ =& X_x \end{aligned}$$

while the vertical component is 0 under $D_p\pi$. Since Y is a smooth vector field, it follows that the horizontal component of Y must be smooth as well hence X^H is the unique horizontal vector field on P satisfying $D_p\pi(X_p^H) = X_\pi(p)$ for all $p \in P$.

Let $X_p \in H_p$ be horizontal, then there exists a unique $Y \in T_{\pi(p)}M$ such that:

$$D_p\pi(X_p) = Y$$

Let (U, ψ) be a a coordinate chart on M containing the point $\pi(p) = x$. Then define a constant vector field on $\psi(U)$ by:

$$Z_y = (y, D_x\psi(Y))$$

for all $y \in \psi(U)$. It then follows that $\phi_*^{-1}Z$ is a vector field on U satisfying:

$$(\phi_*^{-1}Z)_x = Y$$

Thus there exists a horizontal vector field on P_U satisfying:

$$D_p\pi X_p^H = (\phi^{-1}Z)_x = Y$$

so since $D_p\pi : H_p \to T_xM$ is an isomorphism we have that:

$$\begin{aligned} X_p^H =& (D_p\pi)^{-1}Y \\ =& (D_p\pi)^{-1} \circ D_p\pi(X_p) \\ =& X_p \end{aligned}$$

as desired. $\square$

We now prove **Theorem 2.1.8**:

Proof. We first show that (2.1.13) is right invariant. First we see that

$$\begin{aligned} (R_g^*F^A) =& d(R_g^*A) + \frac{1}{2}R_g^*[A, A] \\ =& \mathrm{Ad}_{g^{-1}} \circ dA + \frac{1}{2}R_g^*[A, A] \end{aligned}$$

Examining the right most term, we find that for $X_p, Y_p \in T_pP$:

$$\begin{aligned} \frac{1}{2}(R_g^*[A, A])(X_p, Y_p) =& [A(R_{g*}X_p), A(R_{g*}Y_p)] \\ =& [\mathrm{Ad}_{g^{-1}} \circ A(X_p), \mathrm{Ad}_{g^{-1}} \circ A(Y_p)] \\ =& \mathrm{Ad}_{g^{-1}} \circ [A(X_p), A(Y_p)] \end{aligned}$$

so:

$$\begin{aligned} (R_g^*F^A) =& \mathrm{Ad}_{g^{-1}} \circ dA + \frac{1}{2}\mathrm{Ad}_{g^{-1}} \circ [A, A] \\ =& \mathrm{Ad}_{g^{-1}} \circ \left(dA + \frac{1}{2}[A, A] \right) \end{aligned}$$

We now need to check that F^A and (2.1.13) are equal. We proceed by cases, let $X, Y \in H_p$ then we see that :

$$F(X, Y) = dA(X, Y)$$

while:

$$dA(X,Y) + \frac{1}{2}[A(X), A(Y)] = dA(X,Y)$$

Now let $X \in V_p$, and $Y \in H_p$, then for some $V \in \mathfrak{g}$ we have that:

$$X = \tilde{V}_p$$

so:

$$F(X,Y) = dA(0,Y) = 0$$

By **Lemma 2.1.3** extend Y to a horizontal vector field Z on an open neighborhood p such that $Z_p = Y$, then by (1.1.26):

$$dA(X,Y) + \frac{1}{2}[A,A](X,Y) = \mathscr{L}_{\tilde{V}}(A(Z))_p - \mathscr{L}_Z(A(\tilde{V}))_p - A([\tilde{V}, Z]_p) + [A(\tilde{V}_p), A(Z_p)]$$

Since $A(Z)$ is identically zero, and $A(\tilde{V})$ is constant we obtain:

$$dA(X,Y) + \frac{1}{2}[A,A](X,Y) = -A([\tilde{V}, Z]_p)$$

However, $[\tilde{V}, Z]_p$ is horizontal by **Lemma 2.1.12** so:

$$dA(X,Y) + \frac{1}{2}[A,A](X,Y) = 0$$

Finally, let $X, Y \in V_p$, then for some $V, W \in \mathfrak{g}$, we have $X = \tilde{V}_p$ and $Y = \tilde{W}_p$. Clearly:

$$F(X,Y) = dA(0,0) = 0$$

while:

$$\begin{aligned}
dA(X,Y) + \frac{1}{2}[A,A](X,Y) =& \mathscr{L}_{\tilde{V}}(A(\tilde{W}))_p - \mathscr{L}_{\tilde{W}}(A(\tilde{W}))_p - A([\tilde{V}, \tilde{W}])_p + [A(\tilde{V}_p), A(\tilde{W}_p)] \\
=& -A([\tilde{V}, \tilde{W}]_p) + [V, W] \\
=& -A(\widetilde{[V, W]}_p) + [V, W] \\
=& -[V, W] + [V, W] \\
=& 0
\end{aligned}$$

Therefore:

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

as desired. $\square$

Example 2.1.11. We continue from **Example 2.1.8**. Let $P = M \times G$, then a connection one form is on $M \times G$ is given by $\pi_G^* \mu_G$. We want to show that F vanishes identically on P. Since, F is zero whenever a vector is vertical, we only need to check $F(X,Y) = 0$ for $X, Y \in H_{(x,g)} = (\pi_M^* TM)_{(x,g)}$. Let $X, Y \in H_{(x,g)}$, then:

$$F(X,Y) = dA(\pi_{G*} \circ \pi_H(X), \pi_{G*} \circ \pi_H(Y)) = dA(\pi_{G*}(X), \pi_{G*}(Y)) = dA(0,0) = 0$$

so F vanishes identically on $M \times G$. This then implies that:

$$d(\pi_G^* \mu_G) + \frac{1}{2}[\pi_G^* \mu_G, \pi_G^* \mu_G] = 0$$

Then, pulling out π_G^* we see:

$$\pi_G^* \left(d\mu_G + \frac{1}{2}[\mu_G, \mu_G] \right) = 0$$

and since π_{G*} is not identically zero on TP, we obtain:

$$d\mu_G + \frac{1}{2}[\mu_G, \mu_G] = 0 \tag{2.1.13}$$

The equation above is known as the **Maurer-Cartan equation**.

With the above we see that if F vanishes on P, then for all $X, Y \in \Gamma(H)$:

$$F^A(X,Y) = dA(X,Y) = -A([X,Y]) = 0 \Longrightarrow [X,Y] \in \Gamma(H)$$

Hence if F^A is identically zero then $\Gamma(H)$ is a Lie subalgebra of $\mathfrak{X}(P)$, so in essence F^A can be thought of has measuring how much $\Gamma(H)$ fails to be closed under the Lie bracket.

The curvature form also satisfies the following important property, often known as the Bianchi-Identity:

Theorem 2.1.10. *Let $P \to M$ be a principal G bundle, A a connection one form on P, and F^A the curvature of A. Then:*

$$\pi_H^* dF = 0$$

Proof. This follows from the structure equation. Let $X, Y, Z \in T_pP$ then:

$$\begin{aligned}\pi_H^* dF(X,Y,Z) =& d\left(dA + \frac{1}{2}[A,A]\right)(\pi_H(X), \pi_H(Y), \pi_H(Z)) \\ =& \frac{1}{2} d[A,A](\pi_H(X), \pi_H(Y), \pi_H(Z))\end{aligned}$$

From equation (1.1.27) we have that by extending $\pi_H(X), \pi_H(Y), \phi_H(Z)$ to horizontal vector fields X^H, Y^H, Z^H on an open neighborhood of p:

$$\begin{aligned}d[A,A](\pi_H(X), \pi_H(Y), \pi_H(Z)) =& \mathscr{L}_{X^H}([A(Y^H), A(Z^H)])_p + \mathscr{L}_{Y^H}([A(Z^H), A(X^H)])_p + \mathscr{L}_{Z^H}([A(X^H, Y^H)])_p \\ & - [A(\mathscr{L}_{X^H} Y^H), A(Z^H)]_p - [A(\mathscr{L}_{Y^H} Z^H), A(X^H)]_p - [A(\mathscr{L}_{Z^H} X^H), A(Y^H)]_p\end{aligned}$$

However this is identically zero since $A(X^H) = A(Y^H) = A(Z^H) = 0$. Thus:

$$\pi_H^* dF = 0$$

as desired. □

We will revisit the theorem above in the sections on covariant derivative, as there will be more convenient ways of expressing this property.

Similarly to the connection one form we can pull F^A back to the base manifold M if we are given a local gauge s.

Definition 2.1.27. Let $P \to M$ be a principal G bundle, A a connection one form on P, and F^A the curvature of A. For a local gauge $s : U \to P_U$, the **local curvature form** is defined by:

$$F_s = s^* F^A$$

with:

$$F_{\mu\nu} = F_s(\partial_\mu, \partial_\nu)$$

for some coordinate frame $\{\partial_\mu\}$ on U.

Proposition 2.1.22. *Let $P \to M$ be a principal G bundle, A a connection one form on P, and F^A the curvature of A. For a local gauge $s : U \to P_U$, the local curvature form satisfies the following local structure equation:*

$$F_s = dA_s + \frac{1}{2}[A_s, A_s]$$

In particular, given a coordinate frame $\{\partial_\mu\}$ for U, we have that the components of F are given by:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial \nu A_\mu + [A_\mu, A_\nu]$$

Proof. We see that:

$$\begin{aligned} s^*F =& s^*(dA) + \frac{1}{2}s^*[A,A] \\ =& d(s^*A) + \frac{1}{2}[s^*A, s^*A] \\ =& dA_s + \frac{1}{2}[A_s, A_s] \end{aligned}$$

Furthermore, we see that:

$$\begin{aligned} F_{\mu\nu} =& dA_s(\partial_\mu, \partial_\nu) + \frac{1}{2}[A_s, A_s](\partial_\mu, \partial_\nu) \\ =& \partial_\mu(A_s(\partial_\nu)) - \partial_\nu(A_s(\partial_\mu)) - A([\partial_\mu, \partial_\nu]) + [A_s(\partial_\mu), A_s(\partial_\nu)] \\ =& \partial_\mu A_\nu - \partial\nu A_\mu + [A_\mu, A_\nu] \end{aligned}$$

as desired. □

We would like to see how F transforms under physical gauge transformations, and global gauge transformations. We will proceed similarly to our work with connection one forms.

Theorem 2.1.11. *Let $P \to M$ be a principal G bundle, A a connection one form on P, and F^A the curvature of A. Furthermore, let $s_i : U_i \to P_{U_i}$ and $s_j : U_j \to P_{U_j}$ be two local gauges related on the overlap by the physical gauge transformation $h : U_i \cap U_j \to G$, i.e.*

$$s_i(x) = s_j(x) \cdot h(x)$$

Then:

$$F_{s_i} = Ad_{h^{-1}} \circ A_{s_j}$$

or if G is a matrix Lie group:

$$F_{s_i} = h^{-1} \cdot F_{s_j} \cdot h$$

Proof. For brevity we denote F_{s_i} and F_{s_j} by F_i and F_j respectively. We have that:

$$\begin{aligned} F_i =& s_i^* F_i \\ =& (\Phi(s_j, h))^* F \end{aligned}$$

From our work in **Theorem 2.1.7** we have that:

$$D_{(s_j(x), h(x))}\Phi(D_x s_j(X), D_x h(X)) = D_{s(X)} R_h(D_x s_j(X)) + \mu_G(\widetilde{D_x h}(x))_{s_j(x)\cdot h}$$

F is identically zero on fundamental vector fields, hence:

$$\begin{aligned} F_i(X,Y) =& F(D_{s(X)}R_h(D_x s_j(X)), D_{s(X)}R_h(D_x s_j(Y))) \\ =& (R_h^* F)(D_x s_j(X), D_x s_j(Y)) \\ =& \mathrm{Ad}_{h^{-1}} \circ F(D_x s_j(X), D_x s_j(Y)) \\ =& \mathrm{Ad}_{h^{-1}} \circ F_j(X,Y) \end{aligned}$$

Thus:

$$F_{s_i} = \mathrm{Ad}_{h^{-1}} \circ A_{s_j}$$

or, if G is a matrix Lie group:

$$F_{s_i} = h^{-1} \cdot F_{s_j} \cdot h$$

□

An entirely analogous proof demonstrates that:

Theorem 2.1.12. *Let $P \to M$ be a principal G bundle, A a connection one form on P, and F^A the curvature of A. Let f be a global bundle automorphism, then:*

$$f^* F^A = Ad_{\sigma_f^{-1}} \circ F^A$$

Furthermore:

$$F^{f^* A} = Ad_{\sigma_f^{-1}} \circ F^A$$

Proof. We see that $f(p) = \Phi(p, \sigma_f(p))$, hence for $X, Y \in T_p P$:

$$(f^* F)(X, Y) = F(D_p R_{\sigma_f}(X), D_p R_{\sigma_f}(Y))$$

where the fundamental vector fields portion of $D_{p,\sigma_f}\Phi$ has vanished by the same argument as before. It is then clear that:

$$(f^* F) = (R^*_{\sigma_f} F) = \mathrm{Ad}_{\sigma_f^{-1}} \circ F$$

Finally, we have that:

$$\begin{aligned} F^{f^* A} =& df^* A + \frac{1}{2}[f^* A, f^* A] \\ =& f^*(dA) + \frac{1}{2} f^*[A, A] \\ =& f^* F^A \\ =& \mathrm{Ad}_{\sigma_f^{-1}} \circ F \end{aligned}$$

as desired. □

Example 2.1.12. We continue with **Example 2.1.9** by calculating the curvature of the connection one form:

$$A_{(z_1, z_2)} = \frac{1}{2} \left(\bar{z}_1 \alpha_1 - z_1 \bar{\alpha}_1 + \bar{z}_2 \alpha_2 - z_2 \bar{\alpha}_2 \right)$$

on principal $\mathbb{S}^1$ bundle $\mathbb{S}^3 \to \mathbb{S}^2$. First we note that in the (z_1, z_2) coordinates for $\mathbb{C}^2$:

$$dz_i(X_0, X_1) = X_i \qquad \text{and} \qquad d\bar{z}_j(X_0, X_1) = \bar{X}_j$$

so we write A as:

$$A = \frac{1}{2} \left(\bar{z}_1 dz_1 - z_1 d\bar{z}_1 + \bar{z}_2 dz_2 - z_2 d\bar{z}_2 \right)$$

Then since $\mathbb{S}^1$ is abelian it follows that:

$$\begin{aligned} F^A =& dA \\ =& \frac{1}{2} \left(d\bar{z}_1 \wedge dz_1 - dz_1 \wedge d\bar{z}_1 + d\bar{z}_2 \wedge dz_2 - dz_2 \wedge d\bar{z}_2 \right) \\ =& - \left(dz_1 \wedge d\bar{z}_1 + dz_2 \wedge d\bar{z}_2 \right) \end{aligned}$$

We now want to pull F^A back by a local gauge s. We begin by quickly noting that since $\mathbb{S}^1$ is abelian, if we have another local gauge $s_j = s \cdot h(x)$, then:

$$s_j^* F = \mathrm{Ad}_{h^{-1}} \circ F_s = F_s$$

so this $s^* F^A$ form on $\mathbb{S}^2$ is independent of choice of gauge, and thus globally defined. We see that by **Example 1.2.15**, for all $(z_1, z_2) \in \mathbb{S}^3$:

$$\pi(z_1, z_2) = (2 z_1 \bar{z}_2, 2|z_1|^2 - 1) \in \mathbb{S}^2$$

Hence we define a map on $U = \mathbb{S}^2 \setminus (0,1) \in \mathbb{C} \times \mathbb{R}$ with image in $\mathbb{S}^3$ by:

$$s(w,z) = \left(i\sqrt{\frac{1}{2}(1-z)}, \frac{1}{2}\frac{i\bar{w}}{\sqrt{\frac{1}{2}(1-z)}} \right)$$

From our work in **Example 1.2.15** we know that this map satisfies $\pi \circ s = \mathrm{Id}_U$, and is thus a local gauge $U \to \mathbb{S}^3_U$.

We now calculate:

$$\begin{aligned} s^*(dz_1) =& d(z_1 \circ s) \\ =& i\sqrt{\frac{1}{2}} d\left(\sqrt{1-z}\right) \\ =& \frac{-idz}{2\sqrt{2}\sqrt{1-z}} \end{aligned}$$

Furthermore:

$$d(\bar{z}_1 \circ s) = \frac{idz}{2\sqrt{2}\sqrt{1-z}}$$

Hence:

$$s^*(dz_1 \wedge d\bar{z}_1) = 0$$

For the other component of F we have that:

$$d(z_2 \circ s) = \frac{i\bar{w}}{2\sqrt{2}(1-z)^{3/2}} dz + \frac{i}{\sqrt{2}\sqrt{1-z}} d\bar{w}$$

while:

$$d(\bar{z}_2 \circ s) = \frac{-iw}{2\sqrt{2}(1-z)^{3/2}} dz + \frac{-i}{\sqrt{2}\sqrt{1-z}} dw$$

so:

$$d(z_2 \circ s) \wedge d(\bar{z}_2 \circ s) = \frac{\bar{w}}{4(1-z)^2} dz \wedge dw + \frac{w}{4(1-z)^2} d\bar{w} \wedge dy + \frac{1}{2(1-z)} d\bar{w} \wedge dw$$

with $w = x + iy$ we obtain:

$$d(z_2 \circ s) \wedge d(\bar{z}_2 \circ s) = \frac{-ix}{2(1-z)^2} dz \wedge dy + \frac{-iy}{2(1-z)^2} dz \wedge dx + \frac{-i}{(1-z)} dx \wedge dy$$

hence:

$$\begin{aligned} s^*F =& \frac{ix}{2(1-z)^2} dx \wedge dy + \frac{iy}{2(1-z)^2} dz \wedge dx + \frac{-i}{(1-z)} dy \wedge dx \\ =& \frac{i}{2(1-z)^2} \left(xdz \wedge dy + ydy \wedge dz + 2(z-1) dy \wedge dx \right) \\ =& \frac{i}{2(1-z)^2} \left(xdz \wedge dy + ydy \wedge dz + zdy \wedge dx + (z-2) dy \wedge dx \right) \end{aligned}$$

We can pull s^*F back by the angle parameterization of $\mathbb{S}^2$:

$$\phi^{-1} = (\sin\theta\cos\phi, \sin\theta\cos\phi, \cos\theta)$$

by first noting that by **Example 1.1.26**:

$$\phi^{-1*}(xdz \wedge dy + ydy \wedge dz + zdy \wedge dx) = \sin\theta d\theta \wedge d\phi$$

so:

$$\begin{aligned}\phi^{-1*}(s^*F) =& \frac{i}{2(1-\cos\theta)^2}\left(\sin\theta + (\cos\theta - 2)\cos\theta\sin\theta\right)d\theta\wedge d\phi \\ =& \frac{i}{2(1-\cos\theta)^2}(1-\cos\theta)^2\sin\theta d\theta\wedge d\phi \\ =& \frac{i}{2}\sin\theta d\theta\wedge d\phi\end{aligned}$$

which is the standard volume form for $\mathbb{S}^2$ in the angle coordinates scaled by $i/2$. Therefore we deduce that the global curvature form $F_{\mathbb{S}^2}$ is:

$$F_{\mathbb{S}^2} = \frac{i}{2}\mathrm{dvol}_g$$

where g is the standard round metric on $\mathbb{S}^2$.

2.1.7 Covariant Derivatives

In order to reproduce physical theories in our geometric formalism, we will need a way of differentiating matter fields along vector fields on the base manifold. In particular, as matter fields are represented by sections of a vector bundle E associated to a principal bundle over the space time M, we will need a $\mathbb{K}$-linear map:

$$\Gamma(E) \longrightarrow \Omega^1(M, E)$$

that satisfies some type of Leibniz law. Though in general there are a great variety of such maps[20], if we are given a connection on our principal bundle, we can write one down for free.

Definition 2.1.28. Let $P \to M$ be a principal G bundle, $E = P \times_\rho V$ a vector bundle associated to P, and A a connection on P. For any smooth section $\Phi \in \Gamma(E)$, and any local gauge $s : U \to P$, we have that

$$\Phi = [s, \phi]$$

for some smooth map $\phi : M \to V$. We define the **covariant derivative** of Φ induced by A as:

$$\nabla^A\Phi = [s, d\phi + \rho_*(A_s)\phi] \tag{2.1.14}$$

For some vector field $X \in \mathfrak{X}(M)$ we write:

$$\nabla^A_X\Phi = [s, d\phi(X) + \rho_*(A_s)\phi]$$

In order to ensure this is map is globally defined, we need to show that it is independent of our choice of local gauge, as then with an adapted bundle atlas we can glue together the local sections of $T^*M \otimes E$ defined by (2.1.14) with a partition of unity to obtain a global section of $T^*M \otimes E$.

Theorem 2.1.13. *Let $P \to M$ be a principal G bundle, $E = P \times_\rho V$ and A a connection on P. Then, the covariant derivative ∇^A is independent of the choice of local gauge $s : U \to P_U$.*

Proof. Let $s' : U' \to P_{U'}$ be another local section of P such that on the overlap $U' \cap U \neq \emptyset$:

$$s' = s \cdot h$$

for some physical gauge transformation h. If $\Phi \in \Gamma(E)$ satisfies :

$$\Phi = [s, \phi]$$

in the local gauge s for some smooth map $\phi : U \to V$, then on the over lap $U \cap U'$ we have that for some $\phi' : U \cap U' \to V$:

$$\Phi = [s', \phi'] = [s \cdot h, \phi'] = [s, \rho(h)\phi']$$

[20] Any map $\Gamma(E) \to \Omega^1(M, E)$ which satisfies the properties of **Theorem 2.1.14** is a covariant derivative, and any of the following proofs which do not explicitly make use of a local connection one form A_s extend to this more general notion.

hence:

$$\phi' = \rho(h^{-1})\phi$$

We also have that under this gauge transformation:

$$A_{s'} = \mathrm{Ad}_{h^{-1}} \circ A_s + h^*\mu_G$$

We would thus like to show that:

$$[s', d\phi' + \rho_*(A_{s'})\phi'] = [s, d\phi + \rho_*(A_s)\phi]$$

For all $x \in M$ and $X_x \in T_xM$, we have that:

$$d\phi'(X_x) = \rho(h^{-1})d\phi(X_x) + (D_x\rho(h^{-1})(X_x))\phi \tag{2.1.15}$$

and that:

$$\rho_*(A_{s'}(X_x))\phi' = \rho_*\left(\mathrm{Ad}_{h^{-1}} \circ A_s(X_x)\right)\rho(h^{-1})\phi + \rho_*\left(\mu_G(D_xh(X_x))\right)\rho(h^{-1})\phi \tag{2.1.16}$$

Looking a the right most term of (2.1.15) we see that:

$$D_x\rho(h^{-1})(X_x) = D_{h^{-1}}\rho \circ D_hi \circ \circ D_xh(X_x)$$

where $i : G \to G$ is the inversion map. Let $D_xh(X_x) = Z \in T_{h(x)}G$, then we see that for any $g \in G$:

$$\begin{aligned}
D_gi(Z) =& \frac{d}{dt}\Big|_{t=0} (g \cdot \exp(t(\mu_G(Z))))^{-1} \\
=& \frac{d}{dt}\Big|_{t=0} \exp(-t(\mu_G(Z))) \cdot g^{-1} \\
=& - D_eR_{g^{-1}}(\mu_G(Z))
\end{aligned}$$

Thus we obtain that:

$$\begin{aligned}
D_x\rho(h^{-1})(X_x) =& D_{h^{-1}}\rho \circ D_eR_{h^{-1}}(-\mu_G(Z)) \\
=& - D_{\rho(e)}R_{\rho(h^{-1})} \circ D_e\rho(\mu_G(Z)) \\
=& - D_e(\rho\mu_G(Z)) \cdot \rho(h^{-1}) \\
=& - \rho_*(\mu_G(D_xh(X_x))) \cdot \rho(h^{-1})
\end{aligned}$$

Therefore (2.1.15) becomes:

$$d\phi'(X_x) = \rho(h^{-1})d\phi(X_x) - \rho_*(\mu_G(D_xh(X_x))) \cdot \rho(h^{-1})\phi \tag{2.1.17}$$

We now examine the first term of (2.1.16):

$$\begin{aligned}
\rho_*(\mathrm{Ad}_{h^{-1}} \circ A_s(X_x)) =& \rho_* \circ c_{h^{-1}*}(A_s(X_x)) \\
=& (\rho \circ c_{h^{-1}})_*(A_s(X_x)) \\
=& (R_{\rho(h)} \circ L_{\rho(h^{-1})} \circ \rho)_*(A_s(X_x)) \\
=& \rho(h^{-1}) \cdot \rho_*(A_s(X_x)) \cdot \rho(h)
\end{aligned}$$

Thus (2.1.16) becomes:

$$\rho_*(A_{s'}(X_x)) = \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi + \rho_*(\mu_G(D_xh(X_x))) \cdot \rho(h^{-1})\phi \tag{2.1.18}$$

Adding (2.1.17) and (2.1.18) gives:

$$d\phi'(X_x) + \rho_*(A_{s'}(X_x)) = \rho(h^{-1})d\phi(X_x) + \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi$$

hence:

$$\begin{aligned}
[s', d\phi'(X_x) + \rho_*(A_{s'})(X_x)] =& [s \cdot h, \rho(h^{-1})d\phi(X_x) + \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi] \\
=& [s, d\phi(X_x) + \rho_*(A_s(X_x))\phi]
\end{aligned}$$

Thus we conclude that:

$$[s', d\phi' + \rho_*(A_{s'})\phi'] = [s, d\phi + \rho_*(A_s)\phi]$$

so the map $\nabla^A : \Gamma(E) \to \Omega^1(M, E)$ is independent of a choice of local gauge. $\square$

We now check ∇^A satisfies the expected properties of a covariant derivative.

Theorem 2.1.14. *Let A be a connection in a principal G bundle over M, and $E = P \times_\rho V$ an associated vector bundle. The map $\nabla^A : \Gamma(E) \to \Omega^1(M, E)$ is $\mathbb{K}$-linear in both entries, and satisfies:*

$$\nabla^A_{fX}\Phi = f\nabla^A_X\Phi$$

for all $f \in C^\infty(M, \mathbb{R})$, and the Leibniz law:

$$\nabla^A_X(\lambda\Phi) = (\mathscr{L}_X\lambda)\Phi + \lambda\nabla^A_X\Phi$$

for all smooth functions $\lambda \in \mathbb{C}^\infty(M, \mathbb{K})$

Proof. Taking $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, the first two claims are clear from the definitions. Let $\lambda \in C^\infty(M, \mathbb{K})$, then:

$$\begin{aligned}\nabla^A_X(\lambda\Phi) =&[s, d(\lambda\phi)(X) + \rho_*(A_s(\lambda X))]\\ =&[s, d\lambda(X)\phi + \lambda d\phi(X) + \lambda\rho_*(A_s(X))]\\ =&(\mathscr{L}_X\lambda)[s, \phi] + \lambda[s, d\phi(X) + \rho_*(A_s(X))]\\ =&(\mathscr{L}_X\lambda)\Phi + \lambda\nabla^A_X\Phi\end{aligned}$$

as desired. □

Note that the gauge field A only acts on sections of E when the induced representation ρ_* is non trivial. Physically, we interpret this as the gauge field interacting with matter. For example, if A is the electromagnetic four potential, then A should interact with electrically charged matter such as electrons and positrons, hence the induced representation of $U(1)$ must be non trivial to recover Maxwell's field equations with matter. We call such non trivial representations **charged**.

We would now like to show that the covariant derivative is compatible with the natural bundle metrics discussed in **Proposition 2.1.13**.

Proposition 2.1.23. *Let $\langle\cdot,\cdot\rangle$ be a scalar product on the vector space V, ρ a (pseudo) orthogonal representation of G on V, and $P \to M$ a principal G bundle over M with connection A. The covariant derivative is then compatible with the induced bundle metric from **Proposition 2.1.13**, $\langle\cdot,\cdot\rangle_E$ in the sense that:*

$$\mathscr{L}_X\langle\Phi, \Phi'\rangle_E = \langle\nabla^A_X\Phi, \Phi'\rangle_E + \langle\Phi, \nabla^A_X\Phi'\rangle_E$$

Proof. By **Proposition 1.2.19**, we see that for all $X \in \mathfrak{g}$, and all $v, w \in V$:

$$\langle\rho_*(X)v, w\rangle + \langle v, \rho_*(X)w\rangle = 0$$

It then follows that for any local gauge $s : U \to P$, and maps $\phi, \phi' : U \to V$:

$$\begin{aligned}\langle\nabla^A_X\Phi, \Phi'\rangle_E + \langle\Phi, \nabla^A_X\Phi'\rangle_E =&\langle d\phi(X) + \rho_*(A_s(X))\phi, \phi'\rangle_V + \langle\phi, d\phi'(X) + \rho_*(A_s(X))\phi'\rangle_V\\ =&\langle d\phi(X), \phi'\rangle_V + \langle\phi, d\phi'(X)\rangle_V + \langle\rho_*(A_s(X))\phi, \phi'\rangle_V + \langle\phi, \rho_*(A_s(X))\phi'\rangle_V\\ =&\langle d\phi(X), \phi'\rangle_V + \langle\phi, d\phi'(X)\rangle_V\\ =&\mathscr{L}_X\langle\Phi, \Phi'\rangle_E\end{aligned}$$

□

Recall that when we first introduced the exterior derivative $d : \omega^k(M) \to \omega^{k+1}(M)$, we defined it first for zero forms, i.e. smooth functions on M. If we notice that smooth sections of E are essentially smooth functions on M with values in E, we that the covariant derivative is a map:

$$\nabla^A_X : \Omega^0(M, E) \longrightarrow \Omega^1(M, E)$$

Our goal is now to extend this map as we did for the exterior derivative, in order to obtain a map $d_A : \Omega^k(M, E) \to \Omega^{k+1}(M, E)$, which we call the **exterior covariant derivative**. We begin with the following definition:

Definition 2.1.29. There is a well defined wedge product $\wedge$ between differential forms on M valued in $\mathbb{K}$ and forms twisted in a $\mathbb{K}$-linear vector bundle E:

$$\wedge : \Omega^k(M) \times \omega^l(M, E) \longrightarrow \omega^{k+l}(M, E)$$

The definition above simply alludes to the standard wedge product, as if $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M, E)$, then we write that:

$$(\omega \wedge \eta)_p(X_1, \ldots, X_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \mathrm{sgn}(\sigma)\omega_p(X_{\sigma(1)}, \ldots, X_{\sigma(k)})\eta_p(X_{\sigma(k+1)}, \ldots, X_{\sigma(k+l)})$$

which is well defined as ω_p will take values in $\mathbb{R}$ or $\mathbb{C}$, and η_p will take values in a $\mathbb{R}$ or $\mathbb{C}$ linear vector space. If E is complex, and M is a real smooth manifold, then the product above is still well defined, as we can still think of ω_p as taking values in $\mathbb{C}$.

To define the exterior covariant derivative, let ω be an element of $\Omega^k(M, E)$, and $e_1, \ldots, e_n$ be a local frame for E over an open set $U \subset M$. Then ω can be written as:

$$\omega = \omega^i \otimes e_i$$

where $\omega^i \in \Omega^k(U)$.

Definition 2.1.30. Let A be a connection one form on a principal G bundle $P \to M$, and $E = P \times_\rho V$ an associated vector bundle. Then we define the **exterior covariant derivative** or **covariant differential**:

$$d_A : \Omega^k(M, E) \longrightarrow \Omega^{k+1}(M, E)$$

by:

$$d_A\omega = (d\omega^i) \otimes e_i + (-1)^k \omega^j \wedge (\nabla^A e_j) \tag{2.1.19}$$

$$\nabla : \Gamma(\otimes^k T^*M \otimes E) \longrightarrow \Gamma(\otimes^{k+1} T^*M \otimes E)$$

$$\nabla(\omega \otimes \Psi) = \nabla\omega \otimes \Psi + \omega \otimes \nabla\Psi$$

This definition superficially depends on a choice of a local frame for E, we need to show that this choice actually does not matter.

Lemma 2.1.4. *The definition of d_A is independent of a choice of local frame for E.*

Proof. Suppose $\omega \in \Omega^m(M, E)$, then for a local frame $e_1, \ldots, e_n$ of E_U we have:

$$\omega = \omega^i \otimes e_i$$

for $\omega^i \in \Omega^m(U)$. Let $f_1, \ldots, f_n$ be another frame for E_U, then there exists an matrix of smooth functions A^i_j on U such that:

$$f_j = A^i_j e_i$$

Let η^i be a family of one forms such that:

$$\eta^i = \left(A^{-1}\right)^i_j \omega^j$$

Thus:

$$\begin{aligned}
\eta^i \otimes f_i &= \left(A^{-1}\right)^i_k \omega^k \otimes A^j_i e_j \\
&= \left(A^{-1}\right)^i_k A^j_i \omega^k \otimes e_j \\
&= \delta^j_k \omega^k \otimes e_j \\
&= \omega^j \otimes e_j \\
&= \omega
\end{aligned}$$

We now want to calculate the exterior covariant derivative of $\eta^i \otimes f_i$ and show it is equal to (2.1.20):

$$\begin{aligned} d_A(\eta^i \otimes f_i) =& d(\eta^i) \otimes f_i + (-1)^m \eta^i \otimes \nabla^A f_i \\ =& d\left(\left(A^{-1}\right)^i_k \omega^k\right) \otimes (A^j_i e_j) + (-1)^m \left(A^{-1}\right)^i_k \omega^k \wedge \nabla^A (A^j_i e_j) \end{aligned}$$

We see that the first term can be written:

$$d\omega^j \otimes e_j + d\left(A^{-1}\right)^i_k \wedge \omega^k \otimes (A^j_i e_j) \tag{2.1.20}$$

while the second term can be written as:

$$(-1)^k \omega^j \wedge \nabla^A e_j + (-1)^m (A^{-1})^i_k \omega^k \wedge d(A^j_i) \otimes e_j \tag{2.1.21}$$

Passing the the differential one form $d(A^j_i)$ through η^j we obtain:

$$(-1)^m \omega^j \wedge \nabla^A e_j + (A^{-1})^i_k d(A^j_i) \wedge \omega^k \otimes e_j \tag{2.1.22}$$

Adding (2.1.21) and (2.1.23) together we get:

$$d_A(\eta^i \otimes f_i) = d_A(\omega^j \otimes e_j) + \left(d\left(A^{-1}\right)^i_k A^j_i + \left(A^{-1}\right)^i_k d(A^j_i)\right) \wedge \omega^k \otimes e_j$$

However:

$$\begin{aligned} d\left(A^{-1}\right)^i_k A^j_i + \left(A^{-1}\right)^i_k d(A^j_i) =& d\left(\left(A^{-1}\right)^i_k A^j_i\right) \\ =& d(\delta^j_k) \\ =& 0 \end{aligned}$$

so:

$$d_A(\eta^i \otimes f_i) = d_A(\omega^j \otimes e_j)$$

as desired. □

With the lemma above, we can prove the following:

Proposition 2.1.24. *Let $P \to M$ be a principal G bundle with connection A, and $E = P \times_\rho V$ an associated vector bundle. Then for all $\omega, \eta \in \Omega^k(M, E)$, $\sigma \in \Omega^l(M)$, and $e \in \Gamma(E)$, the exterior covariant derivative satisfies the following properties:*

a) $d_A(\omega + \eta) = d_A\omega + d_A\omega'$

b) $d_A(\sigma \otimes e) = d\sigma \otimes e + (-1)^l \sigma \wedge \nabla^A e$

c) $d_A(\sigma \wedge \omega) = d\sigma \wedge \omega + (-1)^l \sigma \wedge d_A\omega$

Proof. We begin with *a*). For some local frame $e_1, \ldots, e_n$ of E let:

$$\omega = \omega^i \otimes e_i \qquad \text{and} \qquad \eta = \eta^i \otimes e_i$$

where $\omega^i, \eta^i \in \Omega(U)$. Then:

$$\omega + \eta = \omega^i \otimes e_i + \eta^i \otimes e_i = (\omega^i + \eta^i) \otimes e_i$$

hence:

$$\begin{aligned} d_A(\omega + \eta) =& d(\omega^i + \eta^i) \otimes e_i + (-1)^k (\omega^i + \eta^i) \wedge \nabla^A e_i \\ =& (d\omega^i + d\eta^i) \otimes e_i + (-1)^k \omega^i \wedge \nabla^A e_i + (-1)^k \eta^i \wedge \nabla^A e_i \\ =& d\omega^i \otimes e_i + (-1)^k \omega^i \wedge \nabla^A e_i + d\eta^i \otimes e_i + (-1)^k \eta^i \wedge \nabla^A e_i \\ =& d_A\omega + d_A\eta \end{aligned}$$

To prove *b*), note that if e is a section of $\Gamma(E)$, then on U we can write e:

$$e = f^i e_i$$

where $f^i \in \mathbb{C}^\infty(U)$. We then see that:

$$\sigma \otimes e = \sigma \otimes f^i e_i = f^i \sigma \otimes e_i$$

hence:

$$\begin{aligned} d_A(\sigma \otimes e) =& d_A(f^i \sigma \otimes e_i) \\ =& d(\sigma f^i) \otimes e_i + (-1)^l f^i \sigma \wedge \nabla^A e_i \\ =& f^i d(\sigma) \otimes e_i + (-1)^l \left(\sigma \wedge d(f^i) \otimes e_i + f^i \sigma \wedge \nabla^A e_i\right) \\ =& d\sigma \otimes e + (-1)^l \sigma \wedge \nabla^A e \end{aligned}$$

Finally, we see that $\sigma \wedge \omega \in \Omega^{k+l}(M, E)$, so:

$$\sigma \wedge \omega = (\sigma \wedge \omega^i) \otimes e_i$$

We calculate:

$$\begin{aligned} d_A(\sigma \wedge \omega) =& d(\sigma \wedge \omega^i) \otimes e_i + (-1)^{k+l} \sigma \wedge \omega \wedge \nabla^A e_i \\ =& d\sigma \wedge \omega^i \otimes e_i + (-1)^l \sigma \wedge d\omega^i \otimes e_i + (-1)^{k+l} \sigma \wedge \omega^i \wedge \nabla^A e_i \\ =& d\sigma \wedge \omega + (-1)^l \sigma \wedge \left(d\omega^i \otimes e_i + (-1)^k \omega^i \wedge \nabla^A e_i\right) \\ =& d\sigma \wedge \omega + (-1)^l \sigma \wedge d_A \omega \end{aligned}$$

□

Let $\sigma \in \Omega^0(M) = C^\infty(M)$ be the constant function identically equal to 1 on M, and e be a section of E, then by condition $b)$

$$\begin{aligned} d_A(\sigma \otimes e) =& d(\sigma) e + \sigma \nabla^A e \\ =& \nabla^A e \end{aligned}$$

so d_A acting on $\Omega^0(M, E) = \Gamma(E)$ is just the usual covariant derivative, as expected.

We can also describe the exterior covariant derivative in a purely local way. Let ρ be a representation of G on a vector space V.

Definition 2.1.31. We define the wedge product:

$$\begin{aligned} \Omega^k(M, \mathfrak{g}) \times \Omega^l(M, V) &\longrightarrow \Omega^{k+l}(M, V) \\ (\eta, \omega) &\longmapsto \eta \wedge \omega \end{aligned}$$

by choosing a basis $\{v_i\}$ for V such that $\omega = \omega^i \otimes v_i$ and setting:

$$\eta \wedge \omega = \rho_*(\eta) v_i \wedge \omega^i$$

The definition above is clearly indepedent of our choice of basis. Furthermore, with respect to a local section $s : U \to P_U$ of some principal G bundle over M, and a basis $\{v_i\}$ for V, we obtain a local frame for $E = P \times_\rho V$ by:

$$e_i = [s, v_i] \tag{2.1.23}$$

Every $\omega \in \Omega^l(M, E)$ then defines an l form on M with values in V by:

$$\omega_s = \omega^i \otimes v_i$$

We now wish to prove the following theorem.

Theorem 2.1.15. *Let $P \to M$ be a principal G bundle with a connection one form A, and $E = P \times_\rho V$ a vector bundle associated to P. With respect to a local gauge $s : U \to P$ we can write:*

$$(d_A \omega)_s = d\omega_s + A_s \wedge \omega_s$$

for all $\omega \in \Omega^l(M, E)$.

Proof. Let $s : U \to P_U$ be a local gauge, and $\{v_i\}$ be a basis for V. Then we obtain local frame for E_U by (2.1.24), and with this local frame we have that:

$$\omega = \omega^i \otimes e_i$$

and:

$$\omega_s = \omega^i \otimes v_i$$

for any $\omega \in \Omega^l(M, E)$. We see that:

$$\begin{aligned} d_A\omega =& d\omega^i \otimes e_i + (-1)^l \omega^i \wedge \nabla^A e_i \\ =& d\omega^i \otimes e_i + \nabla^A e_i \wedge \omega^i \end{aligned}$$

Note that:

$$\nabla^A e_i = [s, \rho_*(A_s)v_i]$$

so:

$$\left(\nabla^A e_i\right)_s = \rho_*(A_s)v_i$$

Hence:

$$\begin{aligned} (d_A\omega)_s =& d\omega^i \otimes v_i + \rho_*(A_s)v_i \wedge \omega^i \\ =& d\omega_s + A_s \wedge \omega_s \end{aligned}$$

as desired. □

We now wish to study curvature in an associated vector bundle. We begin with the following definition:

Definition 2.1.32. Let E be vector bundle associated to a principal G bundle $P \to M$, and A a connection one form on P. The curvature of ∇^A is defined as:

$$F(X, Y)\Phi = \nabla^A_X \nabla^A_Y \Phi - \nabla^A_Y \nabla^A_X \Phi - \nabla_{[X,Y]}\Phi$$

for all $X, Y \in \mathfrak{X}(M)$ and all $\Phi \in \Gamma(E)$.

As it turns out, the curvature in an associated vector bundle is intimately related to the curvature form F^A on the principal bundle.

Theorem 2.1.16. *Let $P \to M$ be a principal G bundle with a connection one form A on P, $E = P \times_\rho V$ a vector bundle associated to P, and Φ a smooth section of E. Then, for any local gauge $s : U \to P_U$, and a smooth map $\phi : U \to V$ such that:*

$$\Phi = [s, \phi]$$

the curvature of Φ satisfies:

$$F(X, Y)\Phi = [s, \rho_*(F^A_s(X, Y))\phi]$$

Proof. We see for some smooth functions ϕ^i on U, and a basis $\{v_i\}$ for V that:

$$\Phi = [s, \phi^i v_i] = [s, \phi]$$

implying that for $Y \in \mathfrak{X}(M)$:

$$d\phi(Y) = d\phi^i(Y)v_i = \mathscr{L}_Y \phi^i v_i = \mathscr{L}_Y \phi$$

hence:

$$\nabla^A_Y \Phi = [s, \mathscr{L}_Y \phi + \rho_*(A_s(Y))\phi] \tag{2.1.24}$$

Choosing a basis $\{T_a\}$ for $\mathfrak{g}$ we have that:

$$\rho_*(A_s(Y)) = A_s^a(Y) \otimes \rho_*(A_s)$$

so:

$$\begin{aligned} d\rho_*(A_s(Y))(X) =& d(A_s^a(Y))(X) \otimes \rho_*(T_a) \\ =& \mathscr{L}_X(A_s^a(Y)) \otimes \rho_*(T_a) \\ =& \rho_*(\mathscr{L}_X(A_s(Y))) \end{aligned} \tag{2.1.25}$$

With (2.1.24) and (2.1.25) at hand we obtain:

$$\begin{aligned} \nabla_X^A \nabla_Y^A = [s, & \mathscr{L}_X(\mathscr{L}_Y\phi) + \rho_*(\mathscr{L}_X(A_s(Y)) + A_s(X)A_s(Y))\phi \\ & \rho_*(A_s(X))d\phi(Y) + \rho_*(A_s(Y))d\phi(X)] \end{aligned}$$

Deducing the form of $\nabla_Y^A \nabla_X^A \Phi$ by symmetry, and recalling that:

$$A_s(X)A_s(Y) - A_s(Y)A_s(X) = \frac{1}{2}[A_s, A_s](X, Y)$$

we see that:

$$\begin{aligned} (\nabla_X^A \nabla_Y^A - \nabla_Y^A \nabla_X^A)\Phi = [s, & \mathscr{L}_X(\mathscr{L}_Y\phi) - \mathscr{L}_Y(\mathscr{L}_X\phi) \\ & + \rho_*(\mathscr{L}_X(A_s(Y)) - \mathscr{L}_Y(A_s(X)) + \frac{1}{2}[A_s, A_s](X, Y))\phi] \end{aligned} \tag{2.1.26}$$

Let $f \in C^\infty(M)$, then it is easy to see that:

$$\mathscr{L}_X(\mathscr{L}_Y f) - \mathscr{L}_Y(\mathscr{L}_X f) = \mathscr{L}_{[X,Y]} f$$

So we can rewrite (2.1.26) as:

$$\begin{aligned} (\nabla_X^A \nabla_Y^A - \nabla_Y^A \nabla_X^A)\Phi = [s, \mathscr{L}_{[X,Y]}\phi + \rho_*(& \mathscr{L}_X(A_s(Y)) - \mathscr{L}_Y(A_s(X)) \\ & + \frac{1}{2}[A_s, A_s](X, Y))\phi] \end{aligned} \tag{2.1.27}$$

We now examine the term:

$$\nabla_{[X,Y]}^A \Phi = [s, \mathscr{L}_{[X,Y]}\phi + \rho_*(A_s([X, Y]))\phi] \tag{2.1.28}$$

Recall that:

$$dA_s(X, Y) = \mathscr{L}_X(A_s(Y)) - \mathscr{L}_Y(A_s(Y)) - A_s([X, Y])$$

so by subtracting (2.1.28) from (2.1.27) we obtain that:

$$\begin{aligned} F(X, Y)\phi =& [s, \rho_*(dA_s(X, Y) + \frac{1}{2}A_s(X, Y))\phi] \\ =& [s, \rho_*(F_s^A(X, Y)\phi)] \end{aligned}$$

as desired. $\square$

From the proceeding theorem it is clear that $F\Phi$ is an element of $\Omega^2(M, E)$, thus motivating our next result.

Theorem 2.1.17. *Let $P \to M$ be a principal G bundle with connection one form A on P, $E = P \times_\rho V$ a vector bundle associated to P, and Φ a smooth section of E. Then, as two forms with values in E:*

$$d_A d_A \Phi = F\Phi$$

Proof. We present two separate proofs, one which makes explicit use of a local connection one form A_s and **Theorem 2.1.16**, and another which follows from **Theorem 2.1.14** and **Proposition 2.1.24**; we begin with the latter.

First, from **Proposition 2.1.24** it follows that:

$$d_A\Phi = \nabla^A\Phi$$

With a local frame e_i for U we can write this as:

$$d_A\Phi = \left(\nabla^A\Phi\right)^i \otimes e_i$$

where $\left(\nabla^A\Phi\right)^i \in \Omega^1(U)$. Then again from **Proposition 2.1.14**, it follows that:

$$d_Ad_A\Phi = d\left(\nabla^A\Phi\right)^i \otimes e_i - \left(\nabla^A\Phi\right)^i \wedge \nabla^A e_i$$

For all $X, Y \in \mathfrak{X}(M)$ we then obtain the following:

$$\begin{aligned}(d_Ad_A\Phi)(X,Y)\Phi =& \mathscr{L}_X\left(\nabla^A_Y\Phi\right)^i \otimes e_i - \mathscr{L}_Y\left(\nabla^A\Phi_x\right)^i \otimes e_i - \nabla^A_{[X,Y]}\Phi \\ &- \left(\nabla^A_X\Phi\right)^i \nabla^A_Y e_i + \left(\nabla^A_Y\Phi\right)^i \nabla^A_X e_i\end{aligned}$$

By **Theorem 2.1.14** we notice that:

$$\nabla^A_X\left(\nabla^A_Y\Phi\right)^i e_i = \mathscr{L}_X\left(\nabla^A_Y\Phi\right)^i e_i + \left(\nabla^A_Y\Phi\right)^i \nabla^A_X e_i$$

thus:

$$(d_Ad_A\Phi)(X,Y) = \nabla^A_X\nabla^A_Y\Phi - \nabla^A_Y\nabla^A_X\Phi - \nabla_{[X,Y]}\Phi$$

implying the claim.

For the former method, let $s : U \to P$ be a local gauge, and $\{v_i\}$ be a basis for V, then:

$$\Phi = [s, \phi^i v_i]$$

for some $\phi^i \in C^\infty(U)$. We take the exterior covariant derivative:

$$d_A\Phi = [s, d\phi^i \otimes v_i + \phi^i\rho_*(A_s)v_i]$$

Taking the exterior covariant derivative again we obtain:

$$\begin{aligned}d_Ad_A\Phi =& [s, -d\phi^i \wedge \rho_*(A_s)v_i + d\phi^i \wedge \rho_*(A_s)v_i + \phi^i\rho_*(dA_s)v_i + \phi^i A_s \wedge (\rho_*(A_s)v_i)] \\ =& [s, \phi^i\rho_*(dA_s)v_i + \phi^i A_s \wedge (\rho_*(A_s)v_i)]\end{aligned}$$

We see that after choosing a basis $\{T_a\}$ for $\mathfrak{g}$:

$$\rho_*(A_s)v_i = A^a \otimes \rho_*(T_a)v_i$$

so:

$$\phi^i A_s \wedge (\rho_*(A_s)v_i) = \phi^i(\rho_*(A_s)\rho_*(T_a)v_i) \wedge A^a_s$$

We now insert the vector fields $X, Y \in \mathfrak{X}(M)$ into the twisted two form above:

$$\begin{aligned}\phi^i A_s \wedge (\rho_*(A_s)v_i)(X,Y) =& \phi^i(\rho_*(A_s(X))\rho_*(T_a)v^i A^a_s(Y) - \rho_*(A_s(Y))\rho_*(T_a)v^i A^a_s(YX)) \\ =& \phi^i(\rho_*(A_s(X))\rho_*(A_s(Y)) - \rho_*(A_s(Y))\rho_*(A_s(X)))v_i \\ =& \phi^i[\rho_*(A_s(X)), \rho_*(A_s(Y))]v_i \\ =& \rho_*([A_s(X), A_s(Y)])\phi \\ =& \rho_*\left(\frac{1}{2}[A_s, A_s](X,Y)\right)\phi\end{aligned}$$

Thus:

$$\begin{aligned} d_A d_A \Phi = & \left[s, \rho_*(dA_s) + \rho_* \left(\frac{1}{2}[A_s, A_s]\right) \phi\right] \\ = & [s, \rho_*(F_s)\phi] \end{aligned}$$

which by **Theorem 2.1.16** implies that:

$$d_A d_A \Phi = F\Phi$$

as desired. □

The theorem above demonstrates two important facts; first and foremost, the exterior covariant derivative does not in general satisfy:

$$d_A \circ d_A = 0$$

This is in start contrast to the standard exterior derivative on the usual forms on $\Omega^k(M)$. Secondly, by the theorem above, we can interpret the non-vanishing of $d_A \circ d_A$ as the measurement of the curvature of the covariant derivative ∇^A.

2.1.8 Forms With Values in Ad(P)

Recall from the previous sections that both the curvature form, and the difference between two connections were realized as Ad-invariant horizontal forms on P. Our main goal in this section is to realize these two objects as global fields on the spacetime, specifically as elements of $\Omega^l(M, \mathrm{Ad}(P))$, where $\mathrm{Ad}(P)$ is the adjoint bundle from **Example 2.1.15**, and then state two more useful, but equivalent, forms of the Bianchi identity. It is easiest to see that horizontal Ad invariant forms are in one to one correspondence with forms twisted with $\mathrm{Ad}(P)$ in the $l = 0$ case. Indeed, we see that:

$$\Omega^0_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}} = \{f \in C^\infty(P, \mathfrak{g}) : f(p \cdot g) = \mathrm{Ad}_{g^{-1}} \circ f(p)\}$$

hence the value of f in any fibre is entirely determined by it's value at a single point in the fibre. Furthermore:

$$\Omega^0(M, \mathrm{Ad}(P)) = \Gamma(\mathrm{Ad}(P))$$

so for any $f \in \Omega^0_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$ we can obtain a unique smooth section $\Phi \in \Gamma(\mathrm{Ad}(P))$ by:

$$\Phi(x) = [p, f(p)]$$

for any $p \in P_x$. This is independent of our choice of p as for all $g \in G$:

$$[p \cdot g, f(p \cdot g)] = [p \cdot g, \mathrm{Ad}_{g^{-1}} \cdot f(p)] = [p, f(p)]$$

so it is well defined. Furthermore, let $\Phi \in \Gamma(\mathrm{Ad}(P))$, then for all $x \in M$, we have that:

$$\Phi(x) = [p, v]$$

for some $p \in P_x$, and $v \in \mathfrak{g}$. We then define f by:

$$f(p) = v$$

We see that for any other $q \in P_x$:

$$f(q) = f(p \cdot g) = \mathrm{Ad}_{g^{-1}} \circ f(p)$$

so f is Ad invariant. Additionally, for some local gauge $s : U \to P$, and some smooth map $\phi : U \to V$:

$$\Phi(x) = [s(x), \phi(x)]$$

Then for any $p \in P_U$ we have that $p = s(x) \cdot g$ hence:

$$f(p) = f(s(x) \cdot g) = \mathrm{Ad}_{g^{-1}} \circ f(s(x)) = \mathrm{Ad}_{g^{-1}} \circ \phi(\pi(p))$$

so f is smooth, and is thus an element of $\Omega^0_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$ such that:

$$[p, f(p)] = \Phi$$

so the assignment $f \to \Phi$ is an isomorphism. With this discussion in mind, we move onwards to the case of general l.

Theorem 2.1.18. *Let P be a principal G bundle over M. Then, the vector spaces $\Omega^l_{hor}(P, \mathfrak{g})^{Ad}$ and $\Omega^l(M, Ad(P))$ are canonically isomorphic.*

Proof. We define the map:

$$F : \Omega^l_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}} \longrightarrow \Omega^l(M, \mathrm{Ad}(P))$$

by:

$$F(\omega)_x(X_1, \ldots, X_l) = [p, \omega_p(Y_1, \ldots, Y_l)]$$

where $x \in P_x$, and $Y_i \in T_pP$ such that $\pi_* Y_i = X_i$ for each i. We first check that this is well defined; let Z_i be another set of vectors in T_pP satisfying $\pi_* Z_i = X_i$. We see that:

$$\pi_*(Y_i - Z_i) = 0$$

hence the difference $Y_i - Z_i$ lies in V_p. Therefore:

$$\begin{aligned} \omega_p(Z_1, \ldots, Z_l) =& \omega_p(Z_1 + (Y_1 - Z_1), \ldots, Z_l + (Y_l - Z_l)) \\ =& \omega_p(Y_1, \ldots, Y_l) \end{aligned}$$

so $F(\omega)_x$ is independent of our choice of Y_i. Secondly, if $q \in P_x$, then for some $g \in G$ we have that $q = p \cdot g$; let $Y_i \in T_qP$, then:

$$\begin{aligned} [q, \omega_q(Y_1, \ldots, Y_l)] =& [p \cdot g, \omega_{p \cdot g}(Y_1, \ldots, Y_l)] \\ =& [p, \mathrm{Ad}_g \circ \omega_{p \cdot g}(Y_1, \ldots Y_l)] \\ =& [p, (R^*_{g^{-1}}\omega)_{p \cdot g}(Y_1, \ldots, Y_l)] \\ =& [p, \omega_p(R_{g^{-1}*}Y_1, \ldots, R_{g^{-1}*}Y_1)] \end{aligned}$$

We see that:

$$\pi_* \circ R_{g^{-1}*}(Y_i) = (\pi \circ R_{g^{-1}})_* Y_i = \pi_* Y_i = X_i$$

so by the independence of our choice of Y_i, we have that $F(\omega)_x$ is independent of our choice of p. To see that $F(\omega)_x$ is smooth, and thus an element of $\Omega^l(M, \mathrm{Ad}(P))$, let $s : U \to P_U$ be a local gauge, and X_i a set of local vector fields on U; then:

$$(s^*\omega)(X_1, \ldots, X_l) = \omega(s_* X_1, \ldots, s_* X_l)$$

is a smooth map $U \to \mathfrak{g}$, so:

$$F(\omega)(X_1, \ldots, X_l)|_U = [s, \omega(s_* X_1, \ldots, s_* X_l)]$$

is a smooth section of $\mathrm{Ad}(P)$, hence $F(\omega) \in \Omega^l(M, \mathrm{Ad}(P))$ as desired.

We clearly see that this assignment is injective, that is $F(\omega)$ is only the zero element if ω is the zero element in $\Omega^l_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$. To see that this map is also surjective, let $\eta \in \Omega^l(M, \mathrm{Ad}(P))$, then:

$$\eta_x(X_1, \ldots, X_l) = [p, v]$$

for some $p \in P_x$, and $v \in \mathfrak{g}$, and a set of vectors $X_i \in T_xM$. We then define an $\omega \in \Omega^l_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$ by:

$$\omega_p(Y_1, \ldots, Y_l) = v$$

for all $Y_i \in T_pP$ satisfying $\pi_* Y_i = X_i$. The condition that this holds for all Y_i which map to X_i forces ω to be horizontal. Furthermore, we see for any other $q = p \cdot g \in P_x$, we have that:

$$\begin{aligned}\omega_q(R_{g*}Y_1, \ldots, R_{g*}Y_l) =& \mathrm{Ad}_{g^{-1}}(v) \\ =& \mathrm{Ad}_{g^{-1}} \circ \omega_p(Y_1, \ldots, Y_l)\end{aligned}$$

so ω is also Ad invariant. Now let $s : U \to P_U$ be a local gauge , with corresponding smooth map $\phi : U \to V$, and X_i be a set of vector fields on U, then:

$$\eta(X_1, \ldots, X_l) = [s, \phi]$$

By our work in **Lemma 2.1.3**, there exist vector fields Y_i on P_U such that for all $p \in P_U$:

$$D_p\pi(Y_i) = X_{i\pi(p)}$$

so:

$$\omega_{s(x)\cdot g}(Y_1, \ldots, Y_l) = \mathrm{Ad}_{g^{-1}} \circ \phi(x)$$

hence ω is smooth. It follows then that $\omega \in \Omega^l_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$, and satisfies:

$$F(\omega) = \eta$$

by construction. The linearity of F is clear, so F is then an isomorphism, which implies the claim. □

From the theorem above, we obtain the following corollary immediately:

Corollary 2.1.5. *Let P be a principal G bundle over M. Then:*

a) *The difference between any two connections can be uniquely identified with an element $\alpha_M \in \Omega^1(M, Ad(P))$. In particular, the set of all connection one forms on P forms an affine space over $\Omega^1(M, Ad(P))$*

b) *The curvature F^A of any connection A can be uniquely identified with an element $F^A_M \in \Omega^2(M, Ad(P))$.*

The above corollary has two important implications. First, recall from **Example 2.1.10** that we realized the four potential of electromagnetism as a connection one form on $\mathbb{R}^4 \times U(1)$, then recall from our knowledge of undergraduate physics that it is only the *difference* of two potentials which has physical meaning. This is exactly what the first part of **Corollary 2.1.5** encodes, as it means we can only view the difference between two potentials as being globally well defined on spacetime.

The second point encodes a more easily digestible statement. For abelian gauge theories, this point offers not much insight, as F_s is a globally well defined form on M. However, for non-abelian gauge theories, this point implies that the curvature form is globally defined on M with values in the adjoint bundle, and though F^A_M is not gauge invariant, we will see that it's L^2 norm is, so long as we we fix an Ad-invariant scalar product on $\mathfrak{g}$.

With the aide of **Theorem 2.1.18**, we turn to proving two new forms of the Bianchi identity.

Theorem 2.1.19. *Let P be a principal G bundle over M, and a A a connection one form. Then:*

$$dF^A + [A, F^A] = 0 \tag{2.1.29}$$

Proof. For brevity we denote the curvature of A by F. We have the structure equation:

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

so:

$$dF = d(dA) + \frac{1}{2}d[A, A] = \frac{1}{2}d[A, A]$$

Expanding A in a local basis $\{T_a\}$ for $\mathfrak{g}$ we obtain:

$$\begin{aligned}
d[A,A] =& d(A^a \wedge A^b \otimes [T_a, T_b]) \\
=& dA^a \wedge A^b \otimes [T_a, T_b] - A^a \wedge dA^b \otimes [T_a, T_b] \\
=& dA^a \wedge A^b \otimes [T_a, T_b] + dA^b \wedge A^a \otimes [T_b, T_a] \\
=& 2[dA, A]
\end{aligned}$$

hence:

$$dF = [dA, A] \tag{2.1.30}$$

Now examine the other term in (2.1.30):

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

For all $p \in P$ and any $X_1, X_1, X_3 \in T_pP$, we have that by **Definition 2.1.26**:

$$\begin{aligned}
[A,[A,A]]_p\,(X_1, X_2, X_2) = & \sum_{\sigma \in S_3} \left[A(X_{\sigma(1)}), [A(X_{\sigma(2)}), A(X_{\sigma(3)})]\right] \\
= & [A(X_1)[A(X_2), A(X_3)]] + [A(X_2)[A(X_3), A(X_1)]] \\
& + [A(X_3)[A(X_1), A(X_2)]] - [A(X_2)[A(X_1), A(X_3)]] \\
& - [A(X_1)[A(X_3), A(X_2)]] - [A(X_3)[A(X_2), A(X_1)]]
\end{aligned}$$

The sum of the three positive terms vanish by the Jacobi identity, and the sum of three negative terms also vanish by the Jacobi identity, hence:

$$[A, [A, A]] = 0$$

Therefore:

$$[A, F^A] = [A, dA] = -[dA, A] \tag{2.1.31}$$

Putting (2.1.31) and (2.1.32) together we see:

$$dF + [A, F^A] = [dA, A] - [dA, A] = 0$$

as desired. □

Corollary 2.1.6. *Let F^A be the curvature of a connection A, then $F^A_M \in \Omega^2(M, Ad(P))$ satisfies:*

$$d_A F^A_M = 0$$

Proof. We again write F^A as F for brevity. From our proof of **Theorem 2.1.18** we have that given a local gauge $s : U \to P_U$:

$$F_M|_U = [s, s^*F] = F^a_s \otimes [s, T_a]$$

where $\{T_a\}$ is a basis for $\mathfrak{g}$. Therefore:

$$F_{Ms} = F^a_s \otimes T_a = F_s$$

where F_s is the local curvature form obtained by pulling F back to U. Since F_s satisfies the local structure structure equations, we see that:

$$(d_A F_M)_s = dF_s + [A_s, F_s] = 0$$

and since this holds for every local gauge:

$$d_A F_M = 0$$

as desired. □

2.2 Spinors

It was a harrowing discovery of the twentieth century that particles carry an intrinsic angular momentum, often referred to as it's *spin.* Without mentioning this fact, we have already studied a classical analogue to this quantum phenomenon. Indeed, gauge fields are the classical counterpart to gauge bosons, which can be viewed as particles which intermediate certain interactions. In the case of electromagnetism, we have seen that the gauge field, or connection A corresponds to the classical electromagnetic four potential. In the quantum setting, after fixing a vacuum gauge field A_0, excitations in the gauge field, which can be viewed as the difference between two connections, are viewed as photons, i.e. a spin 1 particle. In fact, all gauge fields are spin 1 because they are one forms forms, and thus transform as usual under rotations.

However, there exists other particles in physics, which are not spin 1. Indeed, fermions, such as electrons, positrons, neutrinos, and quarks, are all spin $\frac{1}{2}$. These types of particles, do not transform like gauge bosons under rotations, but instead admit a rotational transformation property under a double covering of SO^+. This transformation under the double covering reflects the fact that these types of particles have spin $\frac{1}{2}$. As we shall see, this double covering is precisely the orthochronus spin group Spin^+, which is intimately related to the study of Clifford algebras. In order to write down the classical Lagrangian for QED, which incorporates fermionic matter, we will thus need to have an apt description of spinor fields, i.e. fields which transform under representations of Spin^+, instead of SO^+.

Since Clifford algebras lay the bedrock for studying the spin groups, we begin this chapter by introducing the algebra necessary to study them, and their properties. Once we have done that, we shall see that Clifford algebras are in essence generalizations of the exterior algebra of a vector space, and can thus be thought of as deformations of the wedge product. We then go on to discuss the representations of Clifford algebras, which we will use to naturally induce representations of the spin groups. Once we have constructed the spin groups, we move onwards to discussing various spin structures on pseudo Riemannian manifolds, which is a principal bundle with structure group Spin^+, that admits a double covering of the proper orthochronus frame bundle $SO^+(M)$ compatible with both group structures. Spinor fields, or fermionic matter fields, are then sections of Spinor bundles, i.e. vector bundles associated to the spin structure on a pseudo-Riemannian manifolds. We end the chapter with a discussion of the spin covariant derivative, the associated Dirac operator on spinor fields, and the Dirac operator on twisted spinors, which is necessary to develop the classical QED Lagrangian.

We follow Hamilton's *Mathematical Gauge Theory*, and Lawson and Michelsohn's *Spin Geometry.*

2.2.1 $\mathbb{K}$ -Algebras

We begin this chapter by recalling some basic fact regarding abstract algebra, necessary for our discussion of Clifford algebras. In particular, our goal in this section is to develop the exterior algebra of a vector space V as the quotient of the tensor algebra of V.

Definition 2.2.1. A **$\mathbb{K}$-algebra**, is a vector space A over a field $\mathbb{K}$ that is equipped with a bilinear map:

$$\mu : A \times A \longrightarrow A$$

For all $v, w \in A$ we often denote $\mu(v, w)$ by $v{\cdot}w$, or simply vw, and hence refer to μ as multiplication in A. If the map μ satisfies:

$$\mu(u, \mu(v, w)) = u \cdot (v \cdot w) = (u \cdot v) \cdot w = \mu(\mu(u, v), w)$$

for all $u, v, w \in A$, then A is an **associative $\mathbb{K}$-algebra**. If in addition there exists an element $1 \in A$ such that:

$$\mu(1, v) = 1 \cdot v = v \cdot 1 = \mu(v, 1)$$

then A an **associative $\mathbb{K}$-algebra with unit element** 1.

In this paper we will take $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$.

Example 2.2.1. We see that with $V = \mathbb{R}^n$, then $\mathrm{End}(V) = \mathrm{Mat}_{n\times n}(\mathbb{R}^n)$ is an associative $\mathbb{R}$-algebra with unit element 1, where μ is simply matrix multiplication, and 1 is the identity matrix. Furthermore, the Lie algebras we studied in chapter 1.2, are examples of non associative $\mathbb{R}$-algebras, where μ is the Lie bracket.

Definition 2.2.2. Let A be an associative $\mathbb{K}$-algebra with unit element 1. We say that A is **graded** if it can be written of the form:

$$A = \bigoplus_{n=0}^{\infty} A_n$$

where A_n are vector spaces satisfying $A_n A_m \subset A_{m+n}$. Note that unit element lies in A_0, as for all n, $A_0 A_n \subset A_n$. An element $a \in A$ is said to be **homogenous** if $a \in A_n$ for some n.

We also have a notion of maps between $\mathbb{K}$-algebras:

Definition 2.2.3. A **$\mathbb{K}$-algebra homomorphism** is a $\mathbb{K}$-linear map $\phi : A \to B$ that respects multiplication, i.e. for all $a, a' \in A$:

$$\phi(a \cdot a') = \phi(a) \cdot \phi(a')$$

If A and B are $\mathbb{K}$ algebras with unit element 1, then we also require that $\phi(1_A) = 1_B$, or rather, $\phi(1) = 1$. An **isomorphism**, or **automorphism** of $\mathbb{K}$-algebras is a bijective homomorphism. A **representation** of A on a vector space V is a homomorphism $\phi : A \to \mathrm{End}(V)$. A representation is called **faithful** if $\phi : A \to \mathrm{End}(V)$ is injective.

Definition 2.2.4. Let A and B be associative $\mathbb{K}$ algebras with unit element 1. A map $\phi : A \to B$ is an **antihomomorphism** is a linear map such that $\phi(1) = 1$, and $\phi(ab) = \phi(b)\phi(a)$.

Finally, we can take the tensor product over $\mathbb{K}$ of associative algebras.

Definition 2.2.5. Let A and B be associative $\mathbb{K}$ algebras. The **tensor product of associative $\mathbb{K}$ algebras** is the vector space $A \otimes_{\mathbb{K}} B$, equipped with the bilinear map:

$$\mu : A \otimes_{\mathbb{K}} B \longrightarrow A \otimes_K B$$

given on simple tensors by:

$$(a \otimes b) \cdot (a' \otimes b') = (a \cdot a') \otimes (b \cdot b')$$

Going forward, all $\mathbb{K}$-algebras will be assumed to be associative with unit element 1, unless stated otherwise. Unlike vector spaces, quotients do not exist when talking about $\mathbb{K}$-algebras. That is if B is a subalgebra of A, i.e. a vector subspace of A which is closed under multiplication, and contains the unit element 1, then the quotient vector space A/B does not inherit the structure of algebra. Indeed, one can easily check that if B is a sub algebra of A, then there exists no $\mathbb{K}$-algebra homomorphism $\pi : A \to A/B$. In other words, equivalence classes in A/B will in general fail to satisfy:

$$[a] \cdot [a'] = [a \cdot a']$$

The solution to this is problem is similar to the case of a ring. We first introduce the following object[21]:

Definition 2.2.6. Let A be a $\mathbb{K}$-algebra, then an ideal of A is a vector subspace $I \subset A$ such that I 'swallows multiplication'. In other words, for all $a \in A$, and $i \in I$, we have that $a \cdot i \in I$ and $i \cdot a \in I$.

Proposition 2.2.1. *Let $\phi : A \to B$ be an algebra homomorphism, then $\ker\phi \subset A$ is an ideal of A.*

Proof. It is clear that $\ker\phi$ is a vector subspace of A, so we need only check that it swallows multiplication. The bilinearity of multiplication guarantees that for all $b \in B$, $0 \cdot b = b \cdot 0 = 0$, hence if $a \in \ker\phi$, then for all $c \in A$ we have that:

$$\phi(a \cdot c) = \phi(a) \cdot \phi(c) = \phi(a) \cdot 0 = 0$$

[21] In the literature, this is sometimes called a two sided ideal, however we have no use for one sided ideals, so the definition is unambiguous in the context of this paper.

and similarly that:

$$\phi(c \cdot a) = \phi(c) \cdot \phi(a) = 0 \cdot \phi(a) = 0$$

so $\ker \phi$ is an ideal of A. □

Often times we talk of ideals as *generated by a subset of* A. Indeed, Let B a subset of A, then the ideal generated by B consists of all element $a \in A$ such that a can be written as the finite sum:

$$a = \sum_i a_i b_i c_i$$

where $a_i, c_i \in A$, and $b_i \in B$. It should be clear that the above prescription defines an ideal.

Theorem 2.2.1. *Let A be $\mathbb{K}$-algebra, and I an ideal of A. Then the quotient space A/I has the structure of an associative $\mathbb{K}$ algebra with unit element, and satisfies the universal property that for all homomorphisms $\phi : A \to B$, such that $I \subset \ker \phi$, there exists a unique homomorphism $\psi : A/I \to B$ such that the following diagram commutes:*

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[d, "\pi"'] & B \\ A/I \arrow[ur, "\psi"'] & \end{tikzcd} \tag{$*$}$$

Proof. We know that as vector spaces, A/I has the structure of a $\mathbb{K}$ vector space, as I is a vector subspace of A. We define multiplication in a way that makes π a homomorphism, i.e. for all $[x], [y] \in A/I$ we set:

$$[x] \cdot [y] = [x \cdot y]$$

We check that this well defined. Let $x' = x + i_x$, and $y' = y + i_y$, for some $i_x, i_y \in I$, then $[x'] = [x]$, and $[y'] = [y]$, and we see that:

$$[x'] \cdot [y'] = [x' \cdot y'] = [xy + xi_y + i_x y + i_x i_y]$$

Since I swallows multiplication, we have that xi_y, $i_x y$, and $i_x i_y$ are all contained in I, and furthermore, since I is a vector subspace we have that:

$$xi_y + i_x y + i_x i_y \in I$$

Hence, for $i = xi_y + i_x y + i_x i_y$ we have:

$$[x'] \cdot [y'] = [xy + i] = [xy]$$

so multiplication is well defined. We also see that equivalence class containing 1 also satisfies:

$$[1] \cdot [x] = [1 \cdot x] = [x] = [x \cdot 1] = [x] \cdot [1]$$

so $[1]$ is the unit element of A/I. Furthermore, multiplication is associative as for all $[x], [y], [z] \in A/I$:

$$[x] \cdot ([y] \cdot [z]) = [x \cdot (y \cdot z)] = [(x \cdot y) \cdot z] = ([x] \cdot [y]) \cdot [z]$$

Finally, multiplication is bilinear, as for all $[x], [y], [z] \in A/I$, and all $k \in \mathbb{K}$ we have that:

$$[x] \cdot (k[y] + k[z]) = [x] \cdot [ky + kz] = [x \cdot (ky + kz)] = k([x] \cdot [y]) + k([x] \cdot [z])$$

so A/I is an associative $\mathbb{K}$ algebra, with unit element, and π is clearly a homomorphism, as $\pi(x) = [x]$.

Now let B be another $\mathbb{K}$-algebra, and $\phi : A \to B$ a homomorphism, such that $I \subset \ker \phi$. This then implies that for all $i \in I$, we have $\phi(i) = 0$. We thus define a map $\psi : A/I \to B$ by:

$$\psi([x]) = \phi(x) \tag{2.2.1}$$

In other words, we define $\psi([x])$ by choosing any element $x \in [x]$ and setting $\psi([x])$ equal to $\phi(x)$. Clearly, we must check that this is well defined. Let $x + i$ be any other element in $[x]$, then:

$$\phi(x + i) = \phi(x) + \phi(i) = \phi(x)$$

so ψ independent of the class representative chosen, and thus well defined. It is unique, as it is uniquely defined by (2.2.1), and it is easily seen to make $(*)$ commute. We check existence by demonstrating it is indeed a homomorphism. First, ψ is linear as for all $[x], [y] \in A/I$ and $k_1, k_2 \in \mathbb{K}$:

$$\psi(k_1[x] + k_2[y]) = \psi([k_1 x + k_2 y]) = \phi(k_1 x + k_2 y) = k_1\phi(x) + k_2\phi(y) = k_1\psi([x)] + k_2\psi([y])$$

Secondly, ψ respects multiplication:

$$\psi([x] \cdot [y]) = \psi([x \cdot y]) = \phi(x \cdot y) = \phi(x) \cdot \phi(y) = \psi([x]) \cdot \psi([y])$$

Finally, it maps $[1]$ to 1:

$$\psi([1]) = \phi(1) = 1$$

so ψ exists, is unique, and makes $(*)$ commute, implying the claim. □

We also have a notion of a graded ideal:

Definition 2.2.7. Let A be a graded algebra, and $I \subset A$ an ideal. We say that I is a **graded ideal** if:

$$I = \bigoplus_{n}^{\infty}(I \cap A_n)$$

Lemma 2.2.1. *Let A be a graded algebra, and $I \subset A$ an ideal. I is graded if and only if I generated by homogenous elements of A.*

Proof. Suppose that I is graded, then any $i \in I$ can be written as the finite sum:

$$i = \sum_i a_i$$

where each $a_i \in I \cap A_i$. Each a_i is a homogenous element of A so it follows that I admits a set of generators which are homogenous.

Suppose that I is generated by a set of homogenous elements. Then every element of i can be written as the finite sum:

$$i = \sum_i a_i b_i c_i$$

where $a_i, c_i \in A$, and each b_i lies in $I \cap A_i$. As A is graded we have that for each i:

$$a_i = \sum_j a_{ij} \qquad \text{and} \qquad c_i = \sum_k c_{ik}$$

where $a_{ij} \in A_j$, and $c_{ik} \in A_k$. We then obtain that:

$$i = \sum_{i,j,k} a_{ij} b_i c_{ik}$$

Since A is graded, and I is an ideal it follows that each $a_{ij} b_i c_{ik} \in I \cap A_{i+j+k}$. We can then rewrite i as the finite sum:

$$i = \sum_n \sum_{i+j+k=n} a_{ij} b_i c_{ik}$$

For each n let:

$$d_n = \sum_{i+j+k=n} a_{ij} b_i c_{ik}$$

which lies in $I \cap A_n$ as each $I \cap A_n$ is closed under addition. Therefore every $i \in I$ can be further rewritten as the finite sum:

$$i = \sum_n d_n$$

Therefore, since $(I \cap A_n) \cap (I \cap A_m) = I \cap (A_n \cap A_m)$, and $A_m \cap A_n = \{0\}$, it follows that:

$$I = \bigoplus_{n=0}^{\infty} (I \cap A_n)$$

so I is a graded ideal. $\square$

Lemma 2.2.2. *Let $\phi : V \to W$ be a linear map, then:*

$$im\ \phi \cong V/\ker\phi$$

Proof. It is easy to see by 'forgetting the algebra structure' of A in the proof of **Theorem 2.1.2** that quotients of vector spaces satisfy a similar universal property. That is if $Y \subset V$ is a vector subspace of V, and $Y \subset \ker\phi$, then ϕ descends to a unique linear map $\psi : V/Y \to W$ such that the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "\pi"'] & W \\ V/Y \arrow[ur, "\psi"'] & \end{tikzcd} \qquad (**)$$

We see that im ϕ is a vector subspace of W, so:

$$\phi : V \longrightarrow \text{im } \phi$$

is a surjective linear map. Furthermore, with $Y = \ker\phi$, we trivially have that $Y \subset \ker\phi$, so ϕ descends to a linear map:

$$\psi : V/\ker\phi \longrightarrow \text{im } \phi$$

which satisfies:

$$\phi = \psi \circ \pi$$

Since $\ker\phi = \ker\pi$, it follows that ψ is injective, as if $\psi([v]) = 0$, then $\phi(v) = 0$, implying that $[v] = [0]$. Furthermore, ψ is surjective as if $w \in$ im ϕ, then we have that $\phi(v) = w$ for some $v \in V$, so $[v]$ satisfies $\psi([v]) = w$. Therefore, ψ is a bijective linear map, and thus an isomorphism, implying the claim. $\square$

Proposition 2.2.2. *Let A be a graded algebra, and $I \subset A$ be a graded ideal, with $I_n = I \cap A_n$. Then A/I is a graded $\mathbb{K}$-algebra satisfying :*

$$A/I = \bigoplus_{n=0}^{\infty} A_n/I_n$$

Proof. From **Theorem 2.2.1**, we know that A/I is an algebra, so we show that it is graded. Since the quotient map $\pi : A \to A/I$ is a $\mathbb{K}$-algebra homomorphism it follows that every $[a] \in A/I$ can be written as:

$$[a] = \left[\sum_n a_n\right] = \sum_n [a_n]$$

Clearly, each $[a_n] \in \pi(A_n)$, which we now denote by $(A/I)_n$. We see that if $[a] \in (A/I)_n \cap (A/I)_m$, then there exists $a_n \in A_n$ and a $a_m \in A_m$ such that:

$$[a_n] = [a] = [a_m]$$

This then implies that:

$$a_n = a_m + i$$

Since I is graded, we can decompose i as a finite sum over homogenous elements $b_i \in I \cap A_i$, so:

$$a_n = a_m + \sum_i b_i$$

implying that:

$$\sum_i b_i = a_n - a_m$$

However, a_n and a_m are homogenous, and each b_i is homogenous, so $b_i = 0$ for $i \neq n, m$, $b_n = a_n$, and $b_m = -a_m$. Since each $b_i \in I$, it follows that a_n and $a_m \in I$, so:

$$[a_m] = [a] = [a_n] = 0$$

Therefore:

$$(A/I)_n \cap (A/I)_m = \{0\}$$

and it follows that:

$$A/I = \bigoplus_n (A/I)_n$$

We also see that for $[a_n] \in (A/I)_n$ and $[a_m] \in (A/I)_m$:

$$[a_n] \cdot [a_m] = [a_m \cdot a_n] \in \pi(A_{n+m}) = (A/I)_{n+m}$$

hence $(A/I)_n(A/I)_m \subset (A/I)_{n+m}$ so A/I is graded. Now define the linear map:

$$\begin{aligned} \phi : A_n &\longrightarrow A/I \\ a_n &\longmapsto [a_n] \end{aligned}$$

We see that ϕ is just the restriction of π to A_n, so im $\phi = (A/I)_n$. Furthermore, we have that $\ker \pi = I \cap A_n = I_n$, thus by **Lemma 2.2.2**:

$$A_n/I_n \cong (A/I)_n$$

It follows that:

$$A/I = \bigoplus A_n/I_n$$

as desired. $\square$

Example 2.2.2. Let V be an n dimensional $\mathbb{K}$-linear vector space; in the section on differential forms, we briefly discussed the exterior algebra of the finite dimensional vector space V^* as the direct sum:

$$\Lambda(V^*) = \bigoplus_{i=0}^{n} \Lambda^i(V^*)$$

where $\Lambda^i(V^*)$ was the vector space of all alternating covariant tensors of order i. Importantly, we can do the same thing for V, where elements $\Lambda^k(V)$ are multilinear maps $(V^*)^k \to \mathbb{K}$, and the wedge product is defined similarly[22]. Furthermore, a similar argument to **Proposition 1.1.12** demonstrates that the wedge product of k vectors $v_1, \dots, v_k$ satisfies:

$$v_1 \wedge \cdots \wedge v_k = \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$$

However, there is an algebraically cleaner way of obtaining the exterior algebra of a vector space, which works equally well for V or it's dual. Indeed, we define the tensor algebra $T(V)$ as the infinite direct sum:

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n} = \mathbb{K} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$$

where all tensor products are over the field $\mathbb{K}$. Elements of $T(V)$ can be written as the sum:

$$\sum_i a_i$$

where each $a_i \in V^{\otimes i}$, and only finitely many a_i are non zero. We see that with multiplication in $T(V)$ defined as the tensor product, $T(V)$ is a graded, associative, $\mathbb{K}$ algebra, with unit element 1, where 1 is the unit element of the field $\mathbb{K}$. This follows from the easily verifiable fact that for any $\mathbb{K}$ vector space W, $W \otimes \mathbb{K} \cong W$, so for any $a \in T(V)$ we have that:

$$1 \otimes a = 1 \cdot a = a$$

Now let I be the ideal generated by the subset:

$$\{v \otimes v : v \in V\}$$

We note that this ideal is generated by homogenous elements of I, so by **Lemma 2.2.1**, I is a graded ideal, and by **Proposition 2.2.2**, $T(V)/I$ is a graded $\mathbb{K}$-algebra. We then wish to show that:

$$T(V)_k / I_k = \Lambda^k(V)$$

for all k. We proceed by cases; let $k < 2$, then $T(V)_0 = \mathbb{K}$ and $T(V)_1 = V$. We see that $I \cap \mathbb{K} = \{0\}$, and $I \cap V = \{0\}$, hence:

$$T(V)_0 / I_0 = \mathbb{K} \qquad \text{and} \qquad T(V)_1 / I_1 = V$$

as desired.

Before moving to the next case, we set some notation, and determine some properties of multiplication in $T(V)/I$. Due to the above fact we write:

$$[k] = k \qquad \text{and} \qquad [v] = v$$

for all elements $[k] \in \mathbb{K} = T(V)_0/I_0$ and $[v] \in V = T(V)_1/I_1$. Furthermore we suggestively denote multiplication in $T(V)/I$ with a wedge, and note that for $v \in V$, and $k \in \mathbb{K}$, $\cdot \wedge \cdot$ satisfies:

$$\begin{aligned} k \wedge v =& [k \cdot v] = k \cdot v \\ v \wedge v =& [v \otimes v] = 0 \end{aligned}$$

In addition, let $w \in V$, then:

$$(v + w) \wedge (v + w) = 0 \Longrightarrow v \wedge w = -w \wedge v$$

Now let $k > n$, we wish to show that $T(V)_k / I_k = \{0\}$. Since π is an additive map, and each $a \in T(V)_k$ can be written as a sum over simple tensors, it suffices to check that all simple tensors get mapped to zero. Let:

$$a = v_1 \otimes \cdots \otimes v_k$$

[22] i.e. just replace the arguments with covectors, instead of vectors.

and note that since $k > n$ at least two of the vectors in the tensor product depend linearly on each other, so for some $1 \leq i, j \leq k$:

$$a = v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_k$$

where $v_j = k \cdot v_i$ for some $k \in \mathbb{K}$. We see that:

$$\begin{aligned}[a] =&[v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_k] \\ =&v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_j \wedge \cdots \wedge v_k \\ =&(-1)^{j-i+1} \cdot v_1 \wedge \cdots \wedge v_i \wedge v_j \wedge \cdots \wedge v_k \\ =&k \cdot (-1)^{j-i+1} \cdot v_1 \wedge \cdots \wedge v_i \wedge v_i \wedge \cdots \wedge v_k \\ =&0\end{aligned}$$

so $T(V)_k/I_k = \{0\}$ as desired. Finally, let $2 \leq k \leq n$, and let ϕ be the linear map:

$$\phi^k : T(V)_k \longrightarrow \Lambda^k(V)$$

given on simple tensors by:

$$v_1 \otimes \cdots \otimes v_k \longmapsto v_1 \wedge \cdots \wedge v_k = \sum_{\sigma \in S_k} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$$

where the above wedge product is the product in $\Lambda(V)$. First note that ϕ^k is surjective, as any element $b \in \Lambda^k(V)$ can be written as the sum:

$$b = \sum_{i_1 \cdots i_k} v_{i_1} \wedge \cdots \wedge v_{i_k}$$

thus:

$$a = k! \sum_{i_1 \cdots i_k} v_{i_1} \otimes \cdots \otimes v_{i_k}$$

satisfies:

$$\phi^k(a) = b$$

We also see that any element in $i \in I_k$ can be written as the finite sum:

$$i = \sum_i \sum_{j+l=k-2} a_j \otimes (v_i \otimes v_i) \otimes b_l$$

where each $a_j \in T(V)_j$ and $b_l \in T(V)_l$, and immediately obtain that $i \in \ker \phi$, so $I_k \subset \ker \phi^k$. Therefore, by the universal property of quotients for a vector space, we see that ϕ descends to a surjective linear map $\psi^k : T(V)_k/I_k \to \Lambda^k(V)$. If we can show that:

$$\dim_{\mathbb{K}}(T(V)_k/I_k) = \binom{n}{k}$$

we will have that ψ is an isomorphism. Let $\{e_i\}$ be a basis for V, then we claim that the set:

$$B = \{[e_{i_1} \otimes \cdots \otimes e_{i_k}] : 1 \leq i_1 < \cdots < i_k \leq k\}$$

forms a basis for $T(V)_k/I_k$. We first show that B spans $T(V)_k/I_k$; the quotient map $\pi : T(V)_k \to T(V)_k/I_k$ is a surjection, so for any $[a] \in T(V)_k/I_k$ we have that there exists an $a \in T(V)_k$ such that $\pi(a) = [a]$. Furthermore, any $a \in T(V)_k$ can be written as a sum of simple tensors:

$$a = \sum_{i_1 \cdots i_k} v_{i_1} \otimes \cdots v_{i_k}$$

and since each $v_{i_j} = \sum_{m_j} a_{i_j}^{m_j} e_{m_j}$ for some $a^{i_j} \in \mathbb{K}$ we have:

$$a = \sum_{i_1 \cdots i_k} \sum_{m_1 \cdots m_k} a_{i_1}^{m_1} \cdots a_{i_k}^{m_k} e_{m_1} \otimes \cdots \otimes e_{m_k}$$

Let:

$$a^{m_1\cdots m_k} = \sum_{i_1\cdots i_k} a_{i_1}^{m_1}\cdots a_{i_k}^{m_k}$$

then we can rewrite this:

$$a = \sum_{m_1\cdots m_k} a^{m_1\cdots m_k} e_{m_1}\otimes\cdots\otimes e_{m_k}$$

We relabel and set $m_j = i_j$, so that:

$$a = \sum_{i_1\cdots i_k} a^{i_1\cdots i_k} e_{i_1}\otimes\cdots\otimes e_{i_k}$$

and note that $m_1\cdots m_k = i_1\cdots i_k$ are not necessarily ordered. Hence:

$$[a] = \sum_{i_1\cdots i_k} a^{i_1\cdots i_k}[e_{i_1}\otimes\cdots\otimes e_{i_k}] \tag{2.2.2}$$

First note that if any $i_l = i_j$ for some $1\leq j,l\leq k$, then:

$$[e_{i_1}\otimes\cdots\otimes e_{i_k}] = 0$$

so we can ignore such terms. Furthermore, for any other multi index $i_1\cdots i_k$ which is not ordered, by our work in the $n>k$ case, we can rewrite the corresponding term in the sum as:

$$[e_{i_1}\otimes\cdots\otimes e_{i_k}] = (-1)^l[e_{j_1}\otimes\cdots\otimes e_{j_k}]$$

where $j_1\cdots j_k$ is ordered, and l is the number of swaps necessary to order $i_1\cdots i_k$. It then follows that:

$$[a] = \sum_{j_1<\cdots<j_k} b^{j_1\cdots j_k}[e_{j_1}\otimes\cdots\otimes e_{j_k}]$$

where each $b^{j_1\cdots j_k}$ is the sum over all $(-1)^l a^{i_1\cdots i_k}$, such that $i_1\cdots i_k$ can be ordered into $j_1\cdots j_k$, and l is the number of swaps necessary to order $i_1\cdots i_k$. Since any $[a]$ can be written as a linear combination of elements in B, we have that B spans $T(V)_k/I_k$. To show linear independence, suppose that the sum:

$$\sum_{i_1<\cdots<i_k} a^{i_1\cdots i_k}[e_{i_1}\otimes\cdots\otimes e_{i_k}] = 0$$

then under ψ^k we have that:

$$\sum_{i_1<\cdots<i_k} a^{i_1\cdots i_k} e_{i_1}\wedge\cdots\wedge e_{i_k} = 0$$

Let $\{e^i\}$ be the dual basis for V^*, then since $V\cong (V^*)^*$ we have that for any ordered multi index $j_1\cdots j_k$:

$$\sum_{i_1<\cdots<i_k} a^{i_1\cdots i_k} e_{i_1}\wedge\cdots\wedge e_{i_k}(e^{j_1},\ldots,e^{j_k}) = a^{j_1\cdots j_k} = 0$$

so each $a^{i_1\cdots i_k} = 0$ and B is a linearly independent set. Clearly, the size of B is n choose k, so ψ^k is an isomorphism as desired. Therefore as vector spaces:

$$T(V)/I = \bigoplus_{k=0}^{n} T(V)_k/I_K = \bigoplus_{k=0}^{n}\Lambda^k(V)$$

Now note that every element $a\in T(V)/I$ can be written as the sum:

$$a = \sum_{i=0}^{n} a_i$$

where each $a_i \in T(V)_k/I_k$. We thus construct a $\mathbb{K}$-algebra isomorphism $\psi : T(V)/I \longrightarrow \Lambda(V)$ by:

$$\sum_{i=0}^{n} a_i \longmapsto \sum_{i=0}^{n} \psi^i(a_i)$$

It is clear that ψ is a vector space isomorphism, and that $\psi(1) = 1$, as ψ_0 is the identity map $\mathbb{K} \to \mathbb{K}$. Thus, by linearity, we need only show that for a set of $k+l$ vectors $v_1, \cdots, v_{k+l} \in V$:

$$\psi([v_1 \otimes \cdots \otimes v_k] \wedge [v_{k+1} \otimes \cdots \otimes v_{k+l}]) = \psi([v_1 \otimes \cdots \otimes v_k]) \wedge \psi([v_{k+1} \otimes \cdots \otimes v_{k+l}])$$

We see that by the induced multiplicative structure on $T(V)/I$:

$$[v_1 \otimes \cdots \otimes v_k] \wedge [v_{k+1} \otimes \cdots \otimes v_{k+l}] = [v_1 \otimes \cdots \otimes v_k \otimes v_{k+1} \otimes \cdots \otimes v_{k+l}]$$

hence:

$$\begin{aligned}\psi([v_1 \otimes \cdots \otimes v_k] \wedge [v_{k+1} \otimes \cdots \otimes v_{k+l}]) =& \psi^{k+l}([v_1 \otimes \cdots \otimes v_k \otimes v_{k+1} \otimes \cdots \otimes v_{k+l}]) \\ =& v_1 \wedge \cdots \wedge v_k \wedge v_{k+1} \wedge \cdots \wedge v_{k+1}\end{aligned}$$

While:

$$\begin{aligned}\psi([v_1 \otimes \cdots \otimes v_k]) \wedge \psi([v_{k+1} \otimes \cdots \otimes v_{k+l}]) =& \psi^{k}([v_1 \otimes \cdots \otimes v_k]) \wedge \psi^l([v_{k+1} \otimes \cdots \otimes v_{k+l}]) \\ =& (v_1 \wedge \cdots \wedge v_k) \wedge (v_{k+1} \wedge \cdots \wedge v_{k+l}) \\ =& v_1 \wedge \cdots \wedge v_k \wedge v_{k+1} \wedge \cdots \wedge v_{k+1}\end{aligned}$$

so ψ is a $\mathbb{K}$-algebra homomorphism, and thus $T(V)/I \cong \Lambda(V)$ as associative $\mathbb{K}$-algebras with unit element 1.

2.2.2 Clifford Algebras

In the previous section, we developed the necessary abstract algebra to construct the the exterior algebra of a vector space as a quotient of the tensor algebra. We are now in a position to understand Clifford algebras as a deformations of the exterior algebra. We first need the following definition:

Definition 2.2.8. Let A be an associative $\mathbb{K}$ algebra, not necessarily with unit element 1. Then for all $x, y \in A$ we define the **commutator** by:

$$[x, y] = x \cdot y - y \cdot x$$

and the **anticommutator** by:

$$\{x, y\} = x \cdot y + y \cdot x$$

A Clifford algebra is then defined as follows:

Definition 2.2.9. Let (V, Q) be a $\mathbb{K}$-linear vector space, with a symmetric bilinear form $Q : V \times V \to \mathbb{K}$. A **Clifford Algebra** of (V, Q) is a pair $(\mathrm{Cl}(V, Q), \gamma)$, where:

a) $\mathrm{Cl}(V, Q)$ is an associative $\mathbb{K}$ algebra with unit element 1

b) γ is a linear map $V \to \mathrm{Cl}(V, Q)$ such that for all $v, w \in V$:

$$\{\gamma(v), \gamma(w)\} = -2Q(v, w) \cdot 1$$

c) $(\mathrm{Cl}(V, Q), \gamma)$ satisfies the universal property that for any other associative $\mathbb{K}$-algebra with unit element 1, and any linear map $\delta : V \to A$ such that:

$$\{\delta(v), \delta(w)\} = -2Q(v, w) \cdot 1$$

there exists a unique algebra homomorphism $\phi : \mathrm{Cl}(V, Q) \to A$ such that the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\gamma"] \arrow[dr, "\delta"'] & \mathrm{Cl}(V,Q) \arrow[d, "\phi"] \\ & A \end{tikzcd}$$

We note that the Clifford algebras were initially motivated by Paul Dirac, in search of finding a *square root* of the Laplacian operator Δ on $\mathbb{R}^{t,s}$. In other words, Dirac wished to find an operator D such that:

$$\Delta f = (D \circ D) f = -\sum_{i=1}^{s+t} \eta_{ii} \frac{\partial^2 f}{\partial x_i^2}$$

where $\eta_{ii} = \eta(e_i, e_i)$ in an orthonormal basis of $\{e_i\}$. If we let:

$$D = \sum_{i}^{s+t} \gamma(e_i) \frac{\partial}{\partial x^i}$$

then it follows that:

$$\begin{aligned} D \circ D = & \sum_{i,j=1}^{s+t} \gamma(e_i) \frac{\partial^2}{\partial x_i x_j} \\ = & \frac{1}{2} \sum_{i,j=1}^{s+t} (\gamma(e_i)\gamma(e_j) + \gamma(e_j)\gamma(e_i)) \frac{\partial^2}{\partial x_i x_j} \end{aligned}$$

hence $\gamma(e_i)$ must satisfy:

$$\{\gamma(e_i), \gamma(e_j)\} = -2\gamma(e_j, e_i)$$

This construction is known as the Dirac operator, and we shall discuss it in detail later in this chapter.

Theorem 2.2.2. *Let (V, Q) be $\mathbb{K}$-linear vector, then there exists a Clifford algebra $(Cl(Q, V), \gamma)$ which is unique up to unique isomorphism.*

Proof. Recall the tensor algebra:

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n}$$

and let I be the ideal generated by:

$$I = v \otimes v + 1 \cdot Q(v, v) : v \in V$$

We claim that $T(V)/I$ is a Clifford algebra. Note that if Q is identically zero on $V \times V$, we see that $T(V)/I$ is the exterior algebra, so $\Lambda(V)$ is a special case of a Clifford algebra. This is precisely what we mean when talking of Clifford algebras as deformations of the exterior algebra.

By **Theorem 2.2.1**, $T(V)/I$ is an associative $\mathbb{K}$ algebra with unit element 1. Furthermore, let $i : V \to T(V)$ denote the inclusion map, then we set:

$$\gamma = \pi \circ i$$

where π is the projection map $T(V) \to T(V)/I$. We see then that for all $v, w \in V$:

$$\gamma(v) \cdot \gamma(w) = [v \otimes w] = -[Q(v, w)] = -Q(v, w) \cdot 1$$

where $1 = [1]$ is the unit element in $T(V)/I$. Therefore since Q is symmetric:

$$\{\gamma(v), \gamma(w)\} = -2Q(v, w)$$

Now let $\delta : V \to A$, where A is an associative $\mathbb{K}$ algebra with unit element 1. For all $k > 0$, we have a $\delta^k(V^{\otimes k}) \to A$ given on simple tensors by:

$$\delta^k(v_1 \otimes \cdots \otimes v_k) = \delta(v_1) \cdots \delta(v_k)$$

Clearly $\delta^1 = \delta$; let $\delta^0 : \mathbb{K} \to A$, be the inclusion map[23], then we obtain the algebra homomorphism:

$$\Delta : T(V) \longrightarrow A$$

defined for all $a \in T(V)$ by:

$$\Delta(a) = \sum_i \delta^i(a_i)$$

where each $a_i \in T(V)_i$. Suppose that δ satisfies:

$$\{\delta(v), \delta(w)\} = -2Q(v, w) \cdot 1$$

for all $v, w \in V$. Let $i \in I$, then i can be written as the finite sum

$$a = \sum_i a_i \otimes (v_i \otimes v_i + Q(v, v)) \otimes b_i$$

where $a_i, b_i \in T(V)$. We then see that:

$$\begin{aligned}
\Delta(a) =& \sum_i \Delta(a_i) \otimes (\delta(v_i) \cdot \delta(v_i) + Q(v_i, v_i)) \otimes \Delta(b_i) \\
=& \sum_i \Delta(a_i) \otimes (-Q(v_i, v_i) + Q(v_i, v_i)) \otimes \Delta(b_i) \\
=& 0
\end{aligned}$$

hence $a \in \ker \Delta$. By the universal property of quotients, we have that Δ descends to a unique algebra homomorphism $\phi : T(V)/I \to A$, such that:

$$\Delta = \phi \circ \pi$$

The restriction of Δ to V is δ, and the restriction of π to V is γ, hence we obtain that there exists a unique ϕ: such that the following diagram commutes:

$$\begin{tikzcd}
V \arrow[r, "\gamma"] \arrow[dr, "\delta"'] & T(V)/I \arrow[d, "\phi"] \\
 & A
\end{tikzcd}$$

Therefore, $(T(V)/I, \gamma)$ is a Clifford algebra for V, so Clifford algebras exists.

Let $(\mathrm{Cl}(V, Q)', \gamma')$ be another Clifford algebra for (V, Q). Then by the universal property of Clifford algebras, we have that there exists a unique algebra homomorphism $f : T(V)/I \to \mathrm{Cl}(V, Q)'$ such that:

$$\gamma' = f \circ \gamma$$

However, there also exists a unique algebra homomorphism $g : CL(V, Q)' \to T(V)/I$ such that:

$$\gamma = g \circ \gamma'$$

Now consider the diagram:

[23]If a $\mathbb{K}$ algebra contains a unit element 1, then since A is a vector space we must have that $\mathbb{K}$ is a vector subspace of A. The inclusion map is then $k \mapsto k \cdot 1$.

$$\begin{tikzcd} V \arrow[r, "\gamma'"] \arrow[dr, "\gamma'"'] & \mathrm{Cl}(V,Q)' \arrow[d, "\phi"] \\ & \mathrm{Cl}(V,Q)' \end{tikzcd}$$

then ϕ is uniquely determined and satisfies:

$$\gamma' = \phi \circ \gamma'$$

Clearly $\phi = \mathrm{Id}_{\mathrm{Cl}(V,Q)'}$ satisfies this property, however so does $f \circ g$ as:

$$\gamma' = f \circ \gamma = f \circ g\gamma$$

hence $f \circ g = \mathrm{Id}_{\mathrm{Cl}(V,Q)'}$. The same argument but replacing $\mathrm{Cl}(V,Q)'$ with $T(V)/I$ demonstrates that $g \circ f = \mathrm{Id}_{T(V)/I}$, hence the objects are unique up to unique isomorphism as desired. $\square$

Since $T(V)/I$, with I as defined in the preceding theorem, is uniquely isomorphic to every other Clifford algebra for (V,Q), going forward we denote $T(V)/I$ by $\mathrm{Cl}(V,Q)$. Furthermore, in the course of showing that $f \circ g = \mathrm{Id}_{\mathrm{Cl}(V,Q)}$, we have obtained the following corollary:

Corollary 2.2.1. *Let $(Cl(V,Q),\gamma)$ be the Clifford algebra for (V,Q). Then, the image of γ, $\gamma(V)$, generates $Cl(V,Q)$. In other words, every $c \in Cl(V,Q)$ can be written as the finite sum:*

$$c = \sum_k \sum_{i_1\cdots i_k} \gamma(v_{i_1})\cdots\gamma(v_{i_k})$$

With this in mind, we can find an upper bound for the dimension of $\mathrm{Cl}(V,Q)$, however we first recall the following lemma:

Lemma 2.2.3. *Let V be a $\mathbb{K}$-linear vector space, equipped with a symmetric bilinear form. Then there exists an orthogonal basis for $\{e_i\}$ for V, i.e. a basis such that $Q(e_i,e_j) = 0$ for $i \neq j$.*

Proof. If Q is nondegenerate then we are done. Suppose Q is degenerate, and define the subset:

$$V^\perp = \{v \in V : Q(v,u) = 0 \forall u \in V\}$$

It is clear that $V^\perp$ is a vector subspace of Q; let U be the complimentary vector subspace such that:

$$V = U \oplus V^\perp$$

Suppose $\dim_{\mathbb{K}} U = k$, and $\dim_{\mathbb{K}} V^\perp = l$, then since Q is non degenerate on U, we can find an orthonormal basis $\{e_1,\cdots,e_k\}$ for U. Any basis $\{f_1,\cdots,f_l\}$ for $V^\perp$ is orthogonal, hence $\{e_1,\cdots,e_k,f_1,\cdots,f_l\}$ is an orthogonal basis for V. $\square$

Corollary 2.2.2. *Let $Cl(V,Q)$ be the Clifford algebra for (V,Q), and suppose that* $\dim_{\mathbb{K}} V = n$. *Then, if $\{e_i\}$ is an orthogonal basis for V, the set:*

$$B = \{\gamma(e_{i_1})\cdots\gamma(e_{i_k}) : 0 \leq k \leq n, 1 \leq i_1 < \cdots < i_k \leq k\}$$

where for $k = 0$ the empty product is equal to 1, spans $Cl(V,Q)$ as a vector space. In particular:

$$\dim_{\mathbb{K}} Cl(V,Q) \leq 2^n$$

Proof. Every element $c \in \mathrm{Cl}(V,Q)$ can be written as:

$$\begin{aligned} c &= \sum_k \sum_{i_1\cdots i_k} \gamma(v_{1_i})\cdots\gamma(v_{i_k}) \\ &= \sum_k \sum_{i_1\cdots i_k} a^{i_1\cdots i_k}\gamma(e_{i_1})\cdots\gamma(e_{i_k}) \end{aligned}$$

Note that for any $i_1 \cdots i_k$, such that $k \geq 2$, we can swap any two adjacent pairs via the clifford relation:

$$\gamma(e_{i_j})\gamma(e_{j+1}) = -\gamma(e_{i_{j+1}})\gamma(e_{i_j})$$

hence:

$$\gamma(e_{i_1}) \cdots \gamma(e_j)\gamma(e_{j+1}) \cdots \gamma(e_{i_k}) = -\,\gamma(e_{i_1}) \cdots \gamma(e_{j+1})\gamma(e_j) \cdots \gamma(e_{i_k}))$$

Furthermore, if any two adjacent $\gamma(e_{i_j})\gamma(e_{i_{j+1}})$ satisfy $e_{i_j} = e_{i_{j+1}}$ we can replace it with the scalar $-Q(e_{i_j}, e_{i_j})$, via the Clifford relation. Therefore, for each k, we can iteratively reorder to the multi index $i_1 \cdots i_k$, into the multi index $j_1 \cdots j_k$, satisfying $j_1 \leq \cdots \leq j_k$. If any of the two adjacent indices are equal, then we can replace them with a scalar, and collect the ordered term into the $k-2$ sum. It follows that if $k > n$, then at most n basis vectors in the product are linearly independent, hence we can always rewrite such a term as a product of at most n $\gamma(e_{i_j})$'s, satisfying $i_1 < \cdots < i_k$. Thus, starting with the largest value of k, we can apply the process above to each k, reordering each term, and applying the Clifford relation to get rid of products where adjacent indices are equal, implying that c can be rewritten as the finite sum:

$$c = \sum_{c=0}^{k} \sum_{i_1 < \cdots < i_k} b^{i_1 \cdots i_k} \gamma(e_{i_1}) \cdots \gamma(e_{i_k})$$

so B spans $\mathrm{Cl}(V, Q)$. Furthermore, the number of elements in B is clearly equal to:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

implying the upper bound on the dimension of $\mathrm{Cl}(V, Q)$. $\square$

It is important to note that as $\mathrm{Cl}(V, Q)$ is not a graded algebra. Indeed, in the existence proof of **Theorem 2.2.2**, the ideal we took a quotient of was not homogenous, so there is no reason to assume that the quotient would be graded in the first place. Furthermore, if we attempt to make a grading:

$$\mathrm{Cl}(V, Q) = \bigoplus_{k=0}^{n} \mathrm{Cl}^k(V, Q)$$

where $\mathrm{Cl}(V, Q)_k$ is the span of:

$$B_k = \{\gamma(e_{i_1}) \cdots \gamma(e_{i_k}) : 1 \leq i_1 < \cdots < i_k \leq k\}$$

We then quickly see that $\mathrm{Cl}^k(V, Q) \cdot \mathrm{Cl}^l(V, Q) \not\subset \mathrm{Cl}^{k+l}(V, Q)$. Indeed, $\gamma(e_i) \in \mathrm{Cl}^1(V, Q)$ however:

$$\gamma(e_i)\gamma(e_i) = -Q(e_i, e_i) \in \mathrm{Cl}^0(V, Q)$$

so no such grading exists.

However, we can instead find a $\mathbb{Z}_2$ grading, that is a grading which satisfies $\mathrm{Cl}(V, Q)_k \cdot \mathrm{Cl}(V, Q)_l \subset \mathrm{Cl}(V, Q)_{k+l}$, where $k, l = 0, 1$, and the indices are taken modulo 2. Indeed, let:

$$T(V)^0 = \bigoplus_{n=0}^{\infty} V^{\otimes 2n} = \mathbb{K} \oplus V^{\otimes 2} \oplus V^{\otimes 4} \cdots$$
$$T(V)^1 = \bigoplus_{n=0}^{\infty} V^{\otimes (2+1)n} = V \oplus V^{\otimes 3} \oplus V^{\otimes 5} \cdots$$

We then set:

$$\mathrm{Cl}^0(V, Q) = T(V)^0/(I \cap T(V)^0) \qquad \text{and} \qquad \mathrm{Cl}^1(V, Q) = T(V)^1/(I \cap T(V)^1)$$

then since:

$$I = (I \cap T(V)^0) \oplus (I \cap T(V)^1)$$

we have that:

$$\mathrm{Cl}(V,Q) = \mathrm{Cl}^0(V,Q) \oplus \mathrm{Cl}^1(V,Q)$$

Then clearly:

$$\mathrm{Cl}^j(V,Q) \cdot \mathrm{Cl}^k(V,Q) \subset \mathrm{Cl}^{j+k}(V,Q)$$

where $j+k$ is taken modulo 2, so $\mathrm{Cl}(V,Q)$ is a $\mathbb{Z}_2$ graded associative algebra with unit element 1. In particular, $\mathrm{Cl}^0(V,Q)$ is a subalgebra of $\mathrm{Cl}(V,Q)$, which is spanned by the set:

$$B^0 = \{\gamma(e_{i_1}) \cdots \gamma(e_{i_{2k}}) : 0 \leq 2k \leq n, i_1 < \cdots i_k\}$$

We now wish to show that as vector spaces $\Lambda^k(V) \cong \mathrm{Cl}(V,Q)$, regardless of the bilinear form Q on V. We first need the following two lemma's:

Lemma 2.2.4. *Let (V,Q) be $\mathbb{K}$-linear vector space, with a symmetric bilinear form Q. Then, for all $0 \leq k \leq n$, and all $v \in V$, there is a unique linear map:*

$$v \lrcorner : \Lambda^k(V) \longrightarrow \Lambda^{(k-1)}(V)$$

such that the following conditions hold:

a) If $\sigma \in V$, then $v \lrcorner \sigma = Q(v,\sigma)$

b) If $\sigma \in \Lambda^k(V)$ and $\omega \in \Lambda^l(V)$ then:

$$v \lrcorner(\sigma \wedge \omega) = (v \lrcorner \sigma) \wedge \omega + (-1)^k \sigma \wedge (v \lrcorner \omega)$$

*We call such a map the **contraction** of σ with v*

Proof. We fix a basis $\{e_i\}$ for V, and for all $\sigma \in \Lambda^k(V)$, and all $v \in V$ define $v \lrcorner \omega$ by:

$$v \lrcorner \sigma = \sum_{i_1 < \cdots < i_k} \sum_{j=1}^{k} (-1)^{j+1} Q(v, e_{i_j}) \sigma^{i_1 \cdots i_k} (-1)^i e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k}$$

where $\hat{e}_{i_j}$ denotes contraction. Note that the above is independent of a chosen basis, as Q is independent of basis. We see that this map is clearly linear, and satisfies a). To see that it satisfies b), let $\sigma \in \Lambda^k(V)$, and $\omega \in \Lambda^l(V)$, since $v \lrcorner$ is a linear map, it suffices to assume that $\sigma = e_{i_1} \wedge \cdots e_{i_k}$, and that $\omega = e_{i_{k+1}} \wedge \cdots \wedge e_{i_{k+l}}$. Then:

$$v \lrcorner(\sigma \wedge \omega) = \sum_{j=1}^{k+l} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge \cdots \wedge e_{i_{k+l}}$$

We split the sum into two parts:

$$\begin{aligned} v \lrcorner(\sigma \wedge \omega) = &\sum_{j=1}^{k} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge \cdots \wedge e_{i_{k+l}} \\ &+ \sum_{j=k+1}^{k+l} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_{k+l}} \end{aligned}$$

The first term is precisely $(v \lrcorner \sigma) \wedge \omega$, so by reindexing the second sum by setting $k+1 = 1$ we obtain:

$$\begin{aligned} v \lrcorner(\sigma \wedge \omega) = &(v \lrcorner \sigma) \wedge \omega + \sum_{j=1}^{k} (-1)^{k+j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_{k+l}} \\ = &(v \lrcorner \sigma) \wedge \omega + (-1)^k \sum_{j=0}^{k} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_{k+l}} \\ = &(v \lrcorner \sigma) \wedge \omega + (-1)^k \sigma \wedge (v \lrcorner \omega) \end{aligned}$$

so $v\lrcorner$ satisfies $b)$ as well, implying such a map exists.

Now suppose that f_v is any linear map satisfying $a)$ and $b)$. We then see that for all $\sigma \in \Lambda^k(V)$:

$$f_v(\sigma) = \sum_{i_1 < \cdots < i_k} \sigma^{i_1 \cdots i_k} f_v(e_{i_1} \wedge \cdots \wedge e_{i_k})$$

We wish to show that:

$$f_v(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \sum_{j=1}^{k} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k}$$

We proceed by induction, for the $k = 1$ case then we have that by $a)$:

$$f_v(e_i) = Q(v, e_i)$$

assuming the inductive hypothesis we have that by $b)$:

$$\begin{aligned} f_v(e_i \wedge \cdots \wedge e_{i_{k-1}} \wedge e_{i_k}) &= \sum_{j=1}^{k-1} Q(v, e_j)(-1)^{j+1} e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_{k-1}} \wedge e_{i_k} \\ &\qquad + (-1)^{k-1}(e_i \wedge \cdots \wedge e_{i_{k-1}}) \wedge f_v(e_{i_k}) \\ &= \sum_{j=1}^{k-1} Q(v, e_j)(-1)^{j+1} e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_{k-1}} \wedge e_{i_k} \\ &\qquad + (-1)^{k-1} Q(v, e_k)(e_i \wedge \cdots \wedge e_{i_{k-1}}) \end{aligned}$$

We see that $(-1)^{k+1} = (-1)^{k-1}$, hence:

$$f_v(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \sum_{j=1}^{k} (-1)^{j+1} Q(v, e_j) e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k}$$

as desired. It then follows that for any $\sigma \in \Lambda^k(V)$:

$$\begin{aligned} f_v(\sigma) = &\sum_{i_1 < \cdots < i_k} \sum_{j=1}^{k} (-1)^{j+1} Q(v, e_{i_j}) \sigma^{i_1 \cdots i_k} (-1)^i e_{i_1} \wedge \cdots \wedge \hat{e}_{i_j} \wedge \cdots \wedge e_{i_k} \\ = &v\lrcorner\sigma \end{aligned}$$

so $v\lrcorner$ is unique, implying the claim. □

Lemma 2.2.5. *With $v\lrcorner$ as defined above, we can extend $v\lrcorner$ linearly to $\Lambda(V)$, such that $v\lrcorner(k) = 0$ for all $k \in \mathbb{K}$, and $v\lrcorner(v\lrcorner\sigma) = 0$ for all $\sigma \in \Lambda(V)$.*

Proof. We define $v\lrcorner$ on $\Lambda(V)$ by noting that $\Lambda(V)$ is graded, hence any element $\sigma \in \Lambda(V)$ can be written as:

$$\sigma = \sum_{k=0}^{n} \sigma_k$$

where each $\sigma_k \in \Lambda^i(V)$. Since $v\lrcorner(\sigma_i)$ is linear on $\Lambda^i(V)$, and each $v\lrcorner(\sigma_i) \in \Lambda^{i-1}(V)$, it follows that:

$$v\lrcorner(\sigma) = \sum_{k=0}^{n} v\lrcorner(\sigma_i)$$

defines a linear map $\Lambda(V) \to \Lambda(V)$, so long as $v\lrcorner$ is identically zero on $\mathbb{K}$.

By linearity of $v\lrcorner$, it suffices to check that $v\lrcorner(v\lrcorner\sigma) = 0$ in the case where $\sigma = e_{i_1} \wedge \cdots \wedge e_{i_k}$, for some basis $\{e_i\}$ of V, and a multi index $i_1 < \cdots < i_k$. We proceed by induction on k, the base case $k = 1$ is trivial as $v\lrcorner = Q(v, e_{i_1})$, so $v\lrcorner(v\lrcorner e_{i_1}) = 0$. Assuming the $k-1$th case we have that:

$$v\lrcorner(\sigma) = v\lrcorner(e_{i_1} \wedge \cdots \wedge e_{i_{k-1}}) \wedge e_{i_k} + (-1)^{k-1} Q(v, e_{i_k}) e_{i_1} \wedge \cdots \wedge e_{i_{k-1}}$$

Hence:

$$\begin{aligned}v\lrcorner(v\lrcorner\sigma) =&v\lrcorner(v\lrcorner(e_{i_1}\wedge\cdots\wedge e_{i_{k-1}})\wedge e_{i_k})+(-1)^{k-1}Q(v,e_{i_k})v\lrcorner(e_{i_1}\wedge\cdots\wedge e_{i_{k-1}})\\ =&(-1)^{k-2}Q(v,e_{i_k})v\lrcorner(e_{i_1}\wedge\cdots\wedge e_{i_{k-1}})+(-1)^{k-1}Q(v,e_{i_k})v\lrcorner(e_{i_1}\wedge\cdots\wedge e_{i_{k-1}})\\ =&0\end{aligned}$$

implying the claim.

□

Theorem 2.2.3. *Let (V,Q) be a $\mathbb{K}$-linear vector space equipped with a symmetric bilinear form Q. Then, there exists canonical vector space isomorphism:*

$$\Lambda(V)\longrightarrow Cl(V,Q)$$

In particular:

$$\dim_{\mathbb{K}} Cl(V,Q)=2^n$$

Proof. For all $v\in V$ define the linear map:

$$\begin{aligned}\delta_v:\Lambda(V)&\longrightarrow\Lambda(V)\\ \sigma&\longmapsto v\wedge\sigma-v\lrcorner\sigma\end{aligned}$$

The assignment:

$$v\longmapsto\delta_v$$

is then a linear map:

$$\delta:V\to\mathrm{End}(\Lambda(V))$$

We see that for all $v,w\in V$, and $\sigma\in\Lambda(V)$:

$$\{\delta_v,\delta_w\}(\sigma)=\delta_v\delta_w(\sigma)+\delta_w\delta_v(\sigma)$$

Note that:

$$\begin{aligned}\delta_v(\delta_w(\sigma)) =&\delta_v(w\wedge\sigma-w\lrcorner\sigma)\\ =&v\wedge(w\wedge\sigma-w\lrcorner\sigma)-v\lrcorner(w\wedge\sigma-w\lrcorner\sigma)\\ =&v\wedge w\wedge\sigma-v\wedge(w\lrcorner\sigma)-v\lrcorner(w\wedge\sigma)+v\lrcorner(w\lrcorner\sigma)\\ =&v\wedge w\wedge\sigma-v\wedge(w\lrcorner\sigma)-Q(v,w)\sigma+w\wedge(v\lrcorner\sigma)+v\lrcorner(w\lrcorner\sigma)\end{aligned}$$

and similarly that:

$$\delta_w(\delta_v(\sigma)) =w\wedge v\wedge\sigma-w\wedge(v\lrcorner\sigma)-Q(v,w)\sigma+v\wedge(w\lrcorner\sigma)+w\lrcorner(v\lrcorner\sigma)$$

hence:

$$\{\delta_v,\delta_w\}(\sigma)=-2Q(v,w)\sigma+v\lrcorner(w\lrcorner\sigma)+w\lrcorner(v\lrcorner\sigma)$$

Note that by **Lemma 2.2.5**:

$$(v+w)\lrcorner((v+w)\lrcorner\sigma)=0=v\lrcorner(w\lrcorner\sigma)+w\lrcorner(v\lrcorner\sigma)$$

hence:

$$\{\delta_v,\delta_w\}(\sigma)=-2Q(v,w)\sigma$$

Since this holds for all σ it follows that

$$\{\delta_v,\delta_w\}=-2Q(v,w)$$

Thus, by the universal property Clifford algebras we obtain a homomorphism:

$$\phi : \mathrm{Cl}(V,Q) \longrightarrow \mathrm{End}(\Lambda(V))$$

With $a = \gamma(e_{i_1})\cdots\gamma(e_{i_k})$, where $i_1 < \cdots < i_k$ we have that:

$$\begin{aligned}\phi(a)(\sigma) =& \delta(e_{i_1})\cdots\delta(e_{i_k})(\sigma)\\ =& e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge \sigma + \text{(terms with contraction)}\end{aligned}$$

Hence we consider the map:

$$\begin{aligned} f : \mathrm{Cl}(V,Q) &\longrightarrow \Lambda(V)\\ a &\longmapsto \phi(a)(1)\end{aligned}$$

where $1 \in \mathbb{K} \subset \Lambda(V)$. We see that since contraction on $\mathbb{K}$ is identically zero:

$$f(\gamma(e_{i_1})\cdots\gamma(e_{i_k})) = e_{i_1} \wedge \cdots \wedge e_{i_k}$$

implying that f is surjective. Since $\dim_{\mathbb{K}} \Lambda(V) = 2^n$, by rank nullity, and **Corollary 2.2.2** we have that:

$$\dim_{\mathbb{K}} \ker f + \dim_{\mathbb{K}} \mathrm{im}\ f = \dim_{\mathbb{K}} \ker f + 2^n \leq 2^n$$

so $\dim_{\mathbb{K}} \ker f = 0$, and f is an isomorphism. □

Corollary 2.2.3. *Given any basis $\{e_i\}$ for (V,Q), the set:*

$$B = \{\gamma(e_{i_1})\cdots\gamma(e_{i_k}) : 0 \leq k \leq n, i_1 < \cdots < i_k\}$$

where the empty product is 1*, forms a basis for* $Cl(V,Q)$*.*

Proof. From **Theorem 2.2.3**, we have that f is an isomorphism, so in particular $f^{-1} : \Lambda(V) \to \mathrm{Cl}(V,Q)$ is an isomorphism. We have that:

$$B' = \{e_{i_1} \wedge \cdots \wedge e_{i_k} : 0 \leq k \leq n, 1 \leq i_1 < \cdots < i_k \leq k\}$$

is a basis for $\Lambda(V)$, so since isomorphism's map basis vectors to basis vectors, and:

$$f^{-1}(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \gamma(e_{i_1})\cdots\gamma(e_{i_k})$$

for all $1 \leq k \leq n$ and all $i_1 < \cdots i_k$, it follows that B is a basis for $\mathrm{Cl}(V,Q)$. □

Corollary 2.2.4. *The linear map $\gamma : V \to Cl(V,Q)$ is injective.*

Proof. We see that if $i : V \to \Lambda(V)$ is the natural inclusion:

$$\gamma = f^{-1} \circ i$$

Since the composition of injective maps is injective, the claim follows □

Since γ is now understood to be an injective linear map, and thus an isomorphism onto it's image, we have that V is a vector subspace of $\mathrm{Cl}(V,Q)$. We will then occasionally suppress the map γ, and simply write $e_{i_1}\cdots e_{i_k}$ in place of the product $\gamma(e_{i_1})\cdots\gamma(e_{i_k})$. We note that we can find isomorphism's between $\mathrm{Cl}(V,Q)$ and a $\mathbb{K}$ algebra A as follows:

i) First find a map from $\delta : V \to A$ such that:

$$\{\delta(v), \delta(w)\} = -2Q(v,w) \cdot 1$$

ii) Use the universal property of CLifford algebras to obtain a unique algebra homomorphism $\phi : \mathrm{Cl}(V,Q) \to A$

iii) Let $\{e_i\}$ be an orthogonal basis for V, then show that products $\delta(e_i)$ span A, implying that ϕ is surjective.

iv) Show $\dim_{\mathbb{K}} \mathrm{Cl}(V,Q) = \dim_{\mathbb{K}} A$, then ϕ is a $\mathbb{K}$ algebra isomorphism.

We end the section with two application's of this procedure:

Lemma 2.2.6. *The map:*

$$\begin{aligned} \alpha : V &\longrightarrow V \\ v &\longmapsto -v \end{aligned}$$

induces a Clifford algebra automorphism, which we also denote by α:

$$\alpha : Cl(V,Q) \longrightarrow Cl(V,Q)$$

such that $\alpha \circ \alpha = Id$. The subspace $Cl^j(V,Q)$ then corresponds to the $(-1)^j$ eigenspace of α, and satisfies $\dim_{\mathbb{K}} Cl^j(V,Q) = \frac{1}{2}\dim_{\mathbb{K}} Cl(V,Q)$.

Proof. We see that the map:

$$\delta = \gamma \circ \alpha \tag{2.2.3}$$

satisfies:

$$\{\delta(v), \delta(w)\} = \{\gamma(-v), \gamma(-w)\} = \{\gamma(v)\gamma(w)\} = -2Q(v,w)\cdot 1$$

and thus induces a unique Clifford algebra endomorphism:

$$\alpha : \mathrm{Cl}(V,Q) \longrightarrow \mathrm{Cl}(V,Q)$$

The image $\delta(V)$ clearly spans $\mathrm{Cl}(V,Q)$, thus α is surjective, so by rank nullity α is an automorphism. Let $\{e_i\}$ be an orthogonal basis for V, we see that for any $a \in \mathrm{Cl}(V,Q)$:

$$\begin{aligned} \alpha(a) =& \sum_{k=0}^{n} \sum_{i_1<\cdots<i_k} a^{i_1\cdots i_k}\delta(e_{i_1})\cdots\delta(e_{i_k}) \\ =& \sum_{k} \sum_{i_1<\cdots<i_k} a^{i_1\cdots i_k}\gamma(-e_{i_1})\cdots\gamma(-e_{i_k}) \\ =& \sum_{k} \sum_{i_1<\cdots<i_k} (-1)^k a^{i_1\cdots i_k}\gamma(e_{i_1})\cdots\gamma(e_{i_k}) \end{aligned}$$

□

This second lemma does not utilize the universal property of Clifford algebras as the map we obtain is an antiautomorphism.

Lemma 2.2.7. *Let (V,Q) be a finite dimensional $\mathbb{K}$-linear vector space with a symmetric bilinear form Q, and $t' : T(V) \to T(V)$ be the antihomomorphism given on simple k tensors by:*

$$v_1 \otimes \cdots v_k \to v_k \otimes v_{k-1} \otimes \cdots \otimes v_2 \otimes v_1$$

This map descends to an anti automorphism:

$$t : Cl(V,Q) \longrightarrow Cl(V,Q)$$

called the ***transpose****, and denote $t(x)$ by x^t.*

Proof. Let I be the ideal used in the construction of the Clifford algebra. We see that for any $i \in I$ written as:

$$i = \sum_i a_i \otimes (v_i \otimes v_i - Q(v_i, v_i)) \otimes b_i$$

where $a_i, b_i \in T(V)$, hence :

$$t'(i) = \sum_i b_i \otimes (v_i \otimes v_i - Q(v_i, v_i)) \otimes a_i \in I$$

implying that $t'(i)$ preserves the ideal I. We thus define the map:

$$x^t = [t'(\omega)]$$

where ω is any class representative in the fibre $\pi^{-1}(x)$. This is unique, and well defined as for any $i \in I$:

$$[t'(\omega + i)] = [t'(\omega) + t(i)] = [t'(\omega)] = x^t$$

It remains to check that $t : \mathrm{Cl}(V, Q) \to \mathrm{Cl}(V, Q)$ is anti automorphism. We see that π is a homomorphism, and t is an antihomomorphism. We defined $t : \mathrm{Cl}(V, Q) \to \mathrm{Cl}(V, Q)$ by the identity:

$$t \circ \pi = \pi \circ t'$$

We want to show that $(xy)^t = y^t x^t$. Let ω_x and ω_y lay in the fibres $\pi^{-1}(x)$ and $\pi^{-1}(y)$, then we see that:

$$t \circ \pi(\omega_x \omega_y) = (xy)^t = \pi(t'(y)t'(x)) = \pi(t'(y))\pi(t'(x)) = y^t x^t$$

as desired. To check that this is an anti automorphism, we need only check t' is surjective by rank nullity. Let $\{e_i\}$ be an orthogonal basis for V, then we see that any basis element of $\mathrm{Cl}(V, Q)$:

$$e_{i_1} \cdots e_{i_k}$$

with $1 \le i_1 < \cdots < i_k \le \dim V$, then:

$$(e_{i_1} \cdots e_{i_k})^t = [t'(e_{i_1} \otimes \cdots \otimes e_{i_k})] = e_{i_k} \cdots e_{i_1}$$

However we can reorder this and pick up some number of minus signs hence:

$$(e_{i_1} \cdots e_{i_k})^t = \pm e_{i_1} \cdots e_{i_k}$$

so t takes basis vectors to basis vectors, and thus the image of a basis under t spans $\mathrm{Cl}(V, Q)$, implying that t is surjective, and thus the claim. □

2.2.3 Clifford Algebras for $\mathbb{R}^{t,s}$ and $\mathbb{C}^n$

For $(V, Q) = (\mathbb{R}^{t,s}, \eta)$, where η is the standard pseudo Euclidean inner product of signature (t, s), we denote the Clifford algebra by $\mathrm{Cl}(t, s)$. Similarly, if $(V, Q) = (\mathbb{C}^n, q)$, where q is the standard complex inner product[24], we denote the Clifford algebra by $\mathbb{C}\mathrm{l}(n)$.

Lemma 2.2.8. *There exists an isomorphism of complex associative algebras:*

$$\mathbb{C}l(t + s) \cong Cl(t, s) \otimes_{\mathbb{R}} \mathbb{C}$$

Complex representations of $Cl(t, s)$ are equivalent to complex representations of $\mathbb{C}l(t + s)$

Proof. Consider the map:

$$\begin{aligned} \delta : \mathbb{R}^{t,s} \otimes_{\mathbb{R}} \mathbb{C} &\longrightarrow \mathrm{Cl}(t, s) \otimes_{\mathbb{R}} \mathbb{C} \\ v \otimes z &\longmapsto \gamma(v) \otimes z \end{aligned}$$

We see that:

$$\begin{aligned} \delta(v \otimes z)\delta(u \otimes z) =& (\gamma(v) \otimes z) \cdot (\gamma(u) \otimes w) \\ =& (\gamma(v)\gamma(u) \otimes zw) \end{aligned}$$

hence:

$$\begin{aligned} \{\delta(v \otimes z), \delta(u \otimes z)\} =& (\gamma(v)\gamma(u) + \gamma(u)\gamma(v)) \otimes (2zw) \\ =& -2\eta(v, u)wz \\ =& -2q(zv, wu) \end{aligned}$$

[24] i.e. in the standard basis for $\mathbb{C}^n$, $q(z^i e_i, w^j e_j) = z^i w^j \delta_{ij} \neq z^i \bar{w}^j \delta_{ij}$

where $zv, wu \in \mathbb{C}^{t+s}$. Since $\mathbb{R}^{t,s} \otimes_R \mathbb{C} \cong \mathbb{C}^{t+s}$, we have that by the universal property of Clifford algebras, there exists a unique algebra homomorphism $\phi : \mathrm{Cl}(t+s) \to \mathrm{Cl}(t,s) \otimes \mathbb{C}$. The dimension of $\mathrm{Cl}(t,s) \otimes \mathbb{C}$ over $\mathbb{C}$ is 2^{t+s}, so $\dim_{\mathbb{C}} \mathrm{Cl}(t,s) \otimes_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{C}} \mathbb{C}\mathrm{l}(t+s)$. Furthermore, the image of δ clearly multiplicatively spans $\mathrm{Cl}(t,s) \otimes_{\mathbb{R}} \mathbb{C}$, so the ϕ is surjective and thus an algebra isomorphism, hence $\mathbb{C}\mathrm{l}(t+s) \cong \mathrm{Cl}(t,s) \otimes_{\mathbb{R}} \mathbb{C}$ □

Lemma 2.2.9. *Let $n \geq 1$ then:*

$$\mathbb{C}\ell^0(n) \cong \mathbb{C}l(n-1)$$

Proof. Let $\{e_i\}$ be an orthonormal basis for $\mathbb{C}^n$, and let V be the one dimensional subspace spanned by e_n. Then, the orthogonal complement of V, denoted $V^\perp$ is $\mathbb{C}^{n-1}$, and is spanned by $\{e_1, \dots, e_{n-1}\}$. Consider the map:

$$\begin{aligned} \mathbb{C}^{n-1} &\longrightarrow \mathbb{C}\mathrm{l}^0(n) \\ w &\longmapsto w \cdot e_n \end{aligned}$$

By construction we have that $q(w, e_n) = 0$, hence for any $u, w \in V^\perp$:

$$\begin{aligned} \delta(u)\delta(w) =& u \cdot e_n \cdot w \cdot e_n \\ =& -u \cdot e_n \cdot e_n \cdot w \\ =& u \cdot w \end{aligned} \tag{2.2.4}$$

hence:

$$\{\delta(u), \delta(w)\} = u \cdot w + w \cdot u = -2q(u, w)$$

Thus, by the universal property of Clifford algebras we have that δ descends to a map $\phi : \mathbb{C}\mathrm{l}(n-1) \to \mathbb{C}\mathrm{l}^0(n)$. By **Lemma 2.2.6**, and **Theorem 2.2.3** we have that $\dim_{\mathbb{K}} \mathbb{C}\mathrm{l}(n-1) = \dim_{\mathbb{K}} \mathbb{C}\mathrm{l}^0(n)$. To conclude that ϕ is an isomorphism, we need to show this map is surjective. Let $a \in \mathbb{C}\mathrm{l}^0(n)$, by linearity it suffices to assume that a is the product:

$$e_{i_1} \cdots e_{i_{2k}}$$

where $i_1 < \cdots < i_{2k}$. Suppose that no $e_{i_j} = e_n$, then, since there by (2.2.4) we have that:

$$\phi(e_{i_1} \cdots e_{i_k}) = \delta(e_{i_1}) \cdots \delta(e_{i_{2k}}) = e_{i_1} \cdots e_{i_{2k}}$$

Now suppose that for some j, $e_{i_j} = e_n$, then:

$$e_{i_1} \cdots e_{i_{j-1}} \cdot e_n \cdot e_{i_{j+1}} \cdots e_{2k} = (-1)^{j-1} e_n \cdot e_1 \cdots e_{i_{j-1}} \cdot e_{i_{j+1}} \cdots e_{2k}$$

hence, since the sequence $i_2, \dots, i_{j-1}, i_{j+1}, \dots, i_{2k}$ has an even amount of terms:

$$\begin{aligned} \phi((-1)^j e_{i_1} \cdots \hat{e}_{i_j} \cdots e_{2k}) =& (-1)^j e_{i_1} \cdot e_n \phi(e_2 \cdots \hat{e}_j \cdots e_{2k}) \\ =& (-1)^{j-1} e_n \cdot e_{i_1} \cdots e_{j-1} \cdot e_{j+1} \cdots e_{2k} \\ =& e_{i_1} \cdots e_{i_{j-1}} \cdot e_n \cdot e_{i_{j+1}} \cdots e_{2k} \end{aligned}$$

so ϕ is surjective, and thus an isomorphism, implying the claim. □

Definition 2.2.10. Let $(V, Q) = (\mathbb{R}^{t,s}, \eta)$, and suppose that:

$$\rho : \mathrm{Cl}(V, Q) \longmapsto \mathrm{End}(\Sigma)$$

is a representation of $\mathrm{Cl}(V, Q)$ on a $\mathbb{K}$ linear vector space $\Sigma = \mathbb{K}^N$. Then for the standard basis $\{e_i\}$ of $\mathbb{R}^{t,s}$, we define the **mathematical gamma matrices**:

$$\gamma_a = \rho \circ \gamma(e_a)$$

for all $1 \leq a \leq t+s$. We define the **physical gamma matrices by**

$$\Gamma_a = (-i)\gamma_a$$

It is easily seen from the definition of an algebra homomorphism that:

$$\{\gamma_a, \gamma_b\} = \rho(\{\gamma(e_a), \gamma(e_b)\}) = \rho(-2\eta_{ab}) = -2\eta_{ab}I_N$$

where I_N is the identity matrix on $\mathbb{K}^N$. Similarly we have that:

$$\{\Gamma_a, \Gamma_b\} = 2\eta_{ab}I_N$$

We also set the following notation:

$$\gamma_{ab} = \frac{1}{2}[\gamma_a, \gamma_b]$$
$$\Gamma_{ab} = \frac{1}{2}[\Gamma_a, \Gamma_b]$$

Similarly to tensors, we raise the index of a gamma matrix by:

$$\gamma^a = \eta^{ab}\gamma_b \qquad \text{and} \qquad \Gamma^a = \eta^{ab}\Gamma_b$$

Furthermore, we set:

$$\gamma^{ab} = \frac{1}{2}[\gamma^a, \gamma^b] = -\frac{1}{2}[\Gamma^a, \Gamma^b] = -\Gamma^{ab}$$

Definition 2.2.11. Let $s+t=n$ be even, the **chirality element** for $\mathrm{Cl}(t,s)$ is a Clifford element of the form:

$$\omega = \lambda e_1 \cdots e_n \in \mathrm{Cl}(t,s) \otimes \mathbb{C}$$

where $\{e_i\}$ is an oriented orthonormal basis for $\mathbb{R}^{t,s}$

Lemma 2.2.10. *The chirality element is independent of a chosen oriented orthonormal basis.*

Proof. Let $\{f_i\}$, and $\{e_i\}$ be two oriented orthonormal basis's. Then there exists an $A \in SO(t,s)$ such that:

$$f_i = A_i^j e_j$$

Let ϕ denote the isomorphism $\mathrm{Cl}(t,s) \to \Lambda(\mathbb{R}^{t,s})$ from **Theorem 2.2.3**. Then:

$$\begin{aligned}\phi(f_1 \cdots f_n) =& f_1 \wedge \cdots \wedge f_n \\ =& (A_1^{j_1} e_{j_1}) \wedge \cdots \wedge (A_n^{j_n} e_{j_n}) \\ =& A_1^{j_1} \cdots A_n^{j_n} e_{j_1} \wedge \cdots \wedge e_{j_n}\end{aligned} \tag{2.2.5}$$

Any term with a repeated index vanishes, and the rest of the terms are permutations of one another, so we can rewrite (2.2.5) as a sum over permutations. Note that:

$$e_{\sigma(j_1)} \wedge \cdots \wedge e_{\sigma(j_n)} = \mathrm{sgn}(\sigma) e_{j_1} \wedge \cdots \wedge e_{j_n}$$

hence (2.2.5) becomes:

$$\begin{aligned}\phi(f_1 \cdots f_n) =& \sum_{\sigma \in S_n} A_1^{\sigma(1)} \cdots A_n^{\sigma(n)} e_1 \wedge \cdots \wedge e_n \\ =& \det(A) e_1 \wedge \cdots \wedge e_n \\ =& e_1 \wedge \cdots \wedge e_n\end{aligned}$$

where there is not implied summation in the first line. Since ϕ is an isomorphism, and thus injective it follows that:

$$e_1 \cdots e_n = f_1 \cdots f_n$$

implying the claim. □

Lemma 2.2.11. *Every chirality element ω satisfies:*

$$\begin{aligned}\{\omega, e_a\} =&0\\ [\omega, e_a \cdot e_b] =&0\end{aligned}$$

Proof. We have that:

$$\begin{aligned}\{\omega, e_a\} =&\omega \cdot e_a + e_a \cdot \omega\\ =&\lambda(e_1 \cdots e_n \cdot e_a + e_a \cdot e_1 \cdots e_n)\\ =&\lambda\left((-1)^{(n-a)} + (-1)^{(a-1)}\right) e_1 \cdots e_{a-1} \cdot e_a \cdot e_a \cdot e_{a+1} \cdots e_n\end{aligned}$$

If a is even then $n-a$ is even and $a-1$ is odd, so the claim follows. If a is odd then $n-a$ is odd, and $a-1$ is even so the claim follows. We then see that:

$$\begin{aligned}[\omega, e_a \cdot e_b] =&\omega \cdot e_a \cdot e_b - e_a \cdot e_b \cdot \omega\\ =&\omega \cdot e_a \cdot e_b + e_b \cdot e_a \cdot \omega\\ =&\omega \cdot e_a \cdot e_b - e_b \cdot \omega \cdot e_a\\ =&\omega \cdot e_a \cdot e_b + \omega \cdot e_b \cdot e_a\\ =&\omega \cdot e_a \cdot e_b - \omega \cdot e_a \cdot e_b\\ =&0\end{aligned}$$

□

Lemma 2.2.12. *If $\lambda^2 = (-1)^{n/2+t}$ then:*

$$\omega^2 = 1$$

Proof. We see that:

$$\begin{aligned}\omega^2 =&\lambda^2 e_1 \cdots e_n \cdot e_1 \cdots e_n\\ =&(-1)^{n/2+t} e_1 \cdots e_n \cdot e_1 \cdots e_n\end{aligned}$$

It takes $n-1$ swaps to reorder $e_1 \cdots e_n$ into $e_n \cdot e_1 \cdots e_{n-1}$. It then takes $n-2$ swaps to reorder into $e_n \cdot e_{n-1} \cdots e_{n-2}$, and so on. Hence the number of swaps necessary to re order $e_1 \cdots e_n$ into $e_n \cdots e_1$ is:

$$\sum_{k=1}^{n} (n-k) = \frac{n(n-1)}{2}$$

hence:

$$\begin{aligned}\omega^2 =&(-1)^{n+2t} e_1 \cdots e_n \cdot e_1 \cdots e_n\\ =&(-1)^n\end{aligned}$$

We have that $n = 2k$ for some k, and:

$$\omega^2 = 1$$

□

Motivated by the preceding lemma, we set the following definition:

Definition 2.2.12. If $s+t=n$ is even, then the **standard chirality element** is given by:

$$\omega = -i^{n/2+t} e_1 \cdots e_n \in \mathrm{Cl}(t,s) \otimes \mathbb{C}$$

Note tht if $n/2+t$ is even, then $\omega \in \mathrm{Cl}(t,s)$. We also set the **mathematical chirality operator** as:

$$\gamma_{n+1} = -i^{n/2+t} \gamma_1 \cdots \gamma_n$$

and the **physical chirality operator** as:

$$\Gamma_{n+1} = -i^{n/2+t} \Gamma_1 \cdots \Gamma_n$$

The chirality element can also be extended to $\mathbb{C}\mathrm{l}(n)$ for even n, by noting that the standard orthonormal basis for $\mathbb{R}^{t,s}$:

$$\{e_1, \ldots, e_t, e_{t+1}, \ldots, e_{t+s}\}$$

can be turned into the standard orthonormal basis for $\mathbb{C}^{t+s}$ via:

$$\{ie_1, \ldots, ie_t, e_{t+1}, \ldots, e_{t+s}\}$$

We see that with this identification η becomes the standard inner product on $\mathbb{C}^{t+s}$. Let $t+s=n$ be even, then the chirality operator is given by:

$$\omega = -i^{n/2}(ie_1)\cdots(ie_t)\cdot(e_{t+1})\cdots e_n$$

Hence, if $\{e_i\}$ is the standard orthonormal basis for $(\mathbb{C}^n, q)$, and n is even, the chirality operator is given by:

$$\omega = -i^{n/2}e_1\cdots e_n$$

We turn to three important examples:

Example 2.2.3. Denote the Pauli spin matrices by:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

It is easy to check that for all $i = 1, 2, 3$:

$$\sigma_i^2 = I_2$$

and that:

$$\sigma_i\sigma_{i+1} = -\sigma_{i+1}\sigma_i = i\sigma_{i+2}$$

Furthermore, the set $\{I, \sigma_1, \sigma_2, \sigma_3\}$ forms a basis for $\mathrm{End}(\mathbb{C}^2)$. With this in mind, consider the Clifford algebra $\mathrm{Cl}(1,1)$, i.e. the Clifford algebra for $\mathbb{R}^{1,1}$ where in the standard basis $\{e_1, e_2\}$:

$$\eta(e_1, e_1) = -1 \qquad \text{and} \qquad \eta(e_2, e_2) = 1$$

We then define the mathematical gamma matrices by:

$$\begin{aligned} \gamma_1 =& \sigma_1 \\ \gamma_2 =& i\sigma_2 \end{aligned}$$

We check that these are indeed mathematical gamma matrices:

$$\begin{aligned} \{\gamma_1, \gamma_2\} =& i(\sigma_1\sigma_2 + \sigma_2\sigma_1) = 0 \\ \gamma_1\gamma_1 =& \sigma_1^2 = I_2 \\ \gamma_2\gamma_2 =& -\sigma_2^2 = -I_2 \end{aligned}$$

We also see that:

$$\gamma_1\gamma_2 = -\sigma_3$$

hence since $\{1, e_1, e_2, e_1e_2\}$ form a basis for $\mathrm{Cl}(1,1)$, it follows that the assignment:

$$\begin{aligned} 1 &\longmapsto I \\ e_1 &\longmapsto \gamma_1 \\ e_2 &\longmapsto \gamma_2 \\ e_1e_2 &\longmapsto \gamma_1\gamma_2 \end{aligned}$$

generates a faithful representation of $\mathrm{Cl}(1,1)$ on $\mathbb{R}^2$, as well as $\mathbb{C}^2$. Importantly, we see that the assignments $1 \to I_2$, and $e_1e_2 \to \gamma_1\gamma_2$ are consequences of any representation being a homomorphism. This representation is also an algebra isomorphism as $\dim_{\mathbb{R}} \mathrm{End}(\mathbb{R}^2) = \dim_{\mathbb{R}} \mathrm{Cl}(1,1)$, so $\mathrm{Cl}(1,1) \cong \mathrm{End}(\mathbb{R}^2)$ Furthermore, since the aforementioned set $\{I, \sigma_1, \sigma_2, \sigma_3\}$ is a basis for $\mathrm{End}(\mathbb{C}^2)$, it follows by **Lemma 2.2.7** that:

$$\mathbb{C}\mathrm{l}(2) \cong \mathrm{Cl}(1,1) \otimes_{\mathbb{R}} \mathbb{C} \cong \mathrm{End}(\mathbb{C}^2)$$

as associative $\mathbb{C}$ algebras.

Example 2.2.4. Consider the clifford algebra $\mathrm{Cl}(1,3)$, that is the Clifford algebra for Minkowski spacetime $\mathbb{R}^{1,3}$ with signature $(-,+,+,+)$. We denote the standard orthonormal basis for $\mathbb{R}^{1,3}$ by $\{e_0, e_1, e_2, e_3\}$, where:

$$\eta(e_0, e_0) = -1 \qquad \text{and} \qquad \eta(e_i, e_i) = 1 \text{ for } 1 \leq i \leq 3$$

Now, let $\mathbb{H}$ denote the space of the quaternions. Recall that the quaternions are an $\mathbb{R}$ vector space spanned by $\{1, i, j, k\}$ which satisfy:

$$ijk = -1 \qquad i^2 = -1 \qquad j^2 = -1 \qquad k^2 = -1 \tag{2.2.6}$$

This multiplicative structure turns the quaternions into an associative $\mathbb{R}$ algebra with unit element one. We can take a direct sum $\mathbb{H}^2 = \mathbb{H} \oplus \mathbb{H}$, which is the set of column 'vectors'[25] with entries in the quaternions. The set of endomorphisms on $\mathbb{H}^2$, denoted $\mathrm{End}(\mathbb{H}^2)$ is the set of all two by two matrices with quaternion entries, and as a real associative algebra satisfies $\dim_{\mathbb{R}} \mathrm{End}(\mathbb{H}^2) = 16$. We thus define mathematical gamma matrices by:

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \gamma_1 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$
$$\gamma_2 = \begin{pmatrix} -j & 0 \\ 0 & j \end{pmatrix} \qquad \gamma_3 = \begin{pmatrix} -k & 0 \\ 0 & k \end{pmatrix}$$

It is then clear that:

$$\gamma_0\gamma_0 = I_2, \qquad \text{and} \qquad \gamma_i\gamma_i = -I_2, \ \forall i = 1, 2, 3$$

Furthermore, from (2.2.6) one can derive that:

$$ij = k = -ji, \qquad jk = i = -kj \qquad ik = j = -ki$$

implying that:

$$\{\gamma_i, \gamma_j\} = 0$$

for all $i, j = 1, 2, 3, 4$ such that $i \neq j$, so the mathematical gamma matrices generate a representation $\rho : \mathrm{Cl}(1,3) \to \mathrm{End}(\mathbb{H}^2)$. We calculate the other terms of the form $\gamma_i\gamma_j$ for $i < j$:

$$\gamma_0\gamma_1 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \qquad \gamma_0\gamma_2 = \begin{pmatrix} 0 & j \\ -j & 0 \end{pmatrix} \qquad \gamma_0\gamma_3 = \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}$$
$$\gamma_1\gamma_2 = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \qquad \gamma_1\gamma_3 = \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix} \qquad \gamma_2\gamma_3 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

For the terms of the form $\gamma_i\gamma_j\gamma_k$ such that $i < j < k$ we have:

$$\gamma_0\gamma_1\gamma_2 = \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix} \qquad \gamma_0\gamma_1\gamma_3 = \begin{pmatrix} 0 & j \\ j & 0 \end{pmatrix}$$
$$\gamma_0\gamma_2\gamma_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad \gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Finally we have that:

$$\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

It is then clear that as a real associative algebra the set:

$$\{\gamma_{i_1} \cdots \gamma_{i_k} : 0 \leq k \leq 4, 0 \leq i_1 < \cdots i_k \leq k\}$$

where with $k = 0$ the empty product is $\rho(1) = I_2$, forms a basis for $\mathrm{End}(\mathbb{H}^2)$. It follows that ρ is a faithful representation of $\mathrm{Cl}(1,3)$ on $\mathbb{H}^2$, and that as real associative algebras:

$$\mathrm{Cl}(1,3) \cong \mathrm{End}(\mathbb{H}^2)$$

[25] $\mathbb{H}^2$ is not a vector space under scalar multiplication by $\mathbb{H}$, as multiplication is noncommunicative. Instead $\mathbb{H}^2$ is an $\mathbb{H}$ module.

Example 2.2.5. We continue with the Clifford algebra $\mathrm{Cl}(1,3)$, but instead seek a faithful representation of $\mathrm{Cl}(1,3)$ on $\mathbb{C}^4$. We define the mathematical gamma matrices by:

$$\gamma_0 = \begin{pmatrix} 0 & -I_2 \\ -I_2 & 0 \end{pmatrix}$$
$$\gamma_i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \qquad \forall i = 1,2,3$$

We see that:

$$\gamma_0\gamma_0 = I_4 \qquad \text{and} \qquad \gamma_i\gamma_i = -I_4$$

while:

$$\{\gamma_i, \gamma_j\} = 0$$

for all $i, j = 0, 1, 2, 3, 4$ such that $i \neq j$, so, as before, the mathematical gamma matrices generate a representation of $\mathrm{Cl}(1,3)$ on $\mathbb{C}^4$. We want to see that this representation is faithful, as before we begin by calculating the $\gamma_i\gamma_j$ terms such that $i < j$:

$$\gamma_0\gamma_1 = \begin{pmatrix} -\sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix} \qquad \gamma_0\gamma_2 = \begin{pmatrix} -\sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix} \qquad \gamma_0\gamma_3 = \begin{pmatrix} -\sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$$
$$\gamma_1\gamma_2 = \begin{pmatrix} -i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix} \qquad \gamma_1\gamma_3 = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix} \qquad \gamma_2\gamma_3 = \begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix}$$

For the $\gamma_i\gamma_j\gamma_k$ such that $i < j < k$ we have:

$$\gamma_0\gamma_1\gamma_2 = \begin{pmatrix} 0 & i\sigma_3 \\ i\sigma_3 & 0 \end{pmatrix} \qquad \gamma_0\gamma_1\gamma_3 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix}$$
$$\gamma_0\gamma_2\gamma_3 = \begin{pmatrix} 0 & i\sigma_1 \\ i\sigma_1 & 0 \end{pmatrix} \qquad \gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 0 & iI_2 \\ -iI_2 & 0 \end{pmatrix}$$

Finally we have that:

$$\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} iI_2 & 0 \\ 0 & -iI_2 \end{pmatrix}$$

With $\rho(1) = I_4$, we see that the set:

$$B = \{\gamma_{i_1} \cdots \gamma_{i_k} : 0 \leq k \leq 4, \}$$

with the empty product equal to I_4, forms a basis for $\mathrm{End}(\mathbb{C}^4)$ as a complex vector space, hence ρ is a faithful representation of $\mathrm{Cl}(1,3)$ on $\mathbb{C}^4$. In particular, it follows that:

$$\mathbb{C}\mathrm{l}(4) \cong \mathrm{Cl}(1,3) \otimes_{\mathbb{R}} \mathbb{C} \cong \mathrm{End}(\mathbb{C}^4)$$

We now turn to describing the structure of the standard Clifford algebras. In the complex case, this is not too difficult, but we need the following two lemmas:

Lemma 2.2.13. *The complex Clifford algebras satisfy the following condition:*

$$\begin{aligned} \mathbb{C}l(n+2) \cong& \mathbb{C}l(n) \otimes_{\mathbb{C}} \mathbb{C}l(2) \\ \cong& \mathbb{C}l(n) \otimes_{\mathbb{C}} End(\mathbb{C}^2) \end{aligned}$$

Proof. We write $\mathbb{C}^{n+2} = \mathbb{C}^n \oplus \mathbb{C}^2$. Let $\{e_{n+1}e_{n+2}\}$ be the standard orthonormal basis for $\mathbb{C}^2$ and $\omega = -ie_{n+1}e_{n+2}$ the chirality element in $\mathbb{C}\mathrm{l}(2)$. Then define the map:

$$\begin{aligned} \delta : \mathbb{C}^n \oplus \mathbb{C}^2 &\longrightarrow \mathbb{C}\mathrm{l}(n) \otimes_{\mathbb{C}} \mathbb{C}\mathrm{l}(2) \\ (v, u) &\longmapsto 1 \otimes u + v \otimes \omega \end{aligned}$$

We see that this map satisfies:

$$\begin{aligned} \delta(v,u)\delta(v,u) =& (1 \otimes u + v \otimes \omega)(1 \otimes u + v \otimes \omega) \\ =& 1 \otimes u \cdot u + v \otimes u \cdot \omega + v \otimes \omega \cdot u + v \cdot v \otimes \omega \cdot \omega \end{aligned}$$

Since $\omega^2 = 1$, and:

$$v \otimes u \cdot \omega + v \otimes \omega \cdot u = v \otimes (\{u, \omega\}) = 0$$

we have that:

$$\delta(v, u)\delta(u, v) = 1 \otimes u \cdot u + v \cdot v \otimes 1 = -(q(u, u) + q(v, v))$$

Therefore, there exists a unique algebra homomorphism $\phi : \mathbb{C}\mathrm{l}(n+2) \to \mathbb{C}\mathrm{l}(n) \otimes_{\mathbb{C}} \mathbb{C}\mathrm{l}(2)$. Note that a basis for $\mathbb{C}\mathrm{l}(n) \otimes_{\mathbb{C}} \mathbb{C}\mathrm{l}(2)$ is given by the union of the following three sets:

$$\begin{aligned}
B_1 &= \{e_{i_1} \cdots e_{i_k} \otimes 1 : 0 \leq k \leq n-2, 1 \leq i_1 < \cdots < i_k \leq k\} \\
B_2 &= \{e_{i_1} \cdots e_{i_k} \otimes e_j : 0 \leq k \leq n-2, 1 \leq i_1 < \cdots < i_k \leq k, j = n+1, n+2\} \\
B_3 &= \{e_{i_1} \cdots e_{i_k} \otimes \omega : 0 \leq k \leq n-2, 1 \leq i_1 < \cdots < i_k \leq k\}
\end{aligned}$$

To show ϕ is surjective, it then suffices to check that for each element b in the above sets there exists an $a \in \mathbb{C}\mathrm{l}(n+2)$ such that $\phi(a) = b$. Suppose that $b \in B_1$ then:

$$b = e_{i_1} \cdots e_{i_k} \otimes 1$$

We see that:

$$\begin{aligned}
\phi(e_1 \cdots e_{i_k} \cdot \omega) =& \delta(e_{i_1}) \cdots \delta(e_{i_k})\delta(-ie_{n+1})\delta(e_{n+2}) \\
=& (e_{i_1} \otimes \omega) \cdots (e_{i_k} \otimes \omega)(-i1 \otimes e_{n+1})(1 \otimes e_{n+2}) \\
=& ((e_{i_1} \cdots e_{i_k}) \otimes \omega) \cdot (1 \otimes \omega) \\
=& e_{i_1} \cdots e_{i_k} \otimes 1
\end{aligned}$$

Now suppose that $b \in B_2$, then:

$$b = e_{i_1} \cdots e_{i_k} \otimes e_j$$

where $j = n+1, n+2$. We thus obtain that:

$$\begin{aligned}
\phi(e_{i_1} \cdots e_{i_k} \cdot e_j \cdot \omega \cdot e_j) =& ((e_{i_1} \cdots e_{i_k}) \otimes \omega) \cdot (1 \otimes \omega) \cdot (1 \otimes e_j) \\
=& e_{i_1} \cdots e_{i_k} \otimes e_j
\end{aligned}$$

Finally, with $b \in B_3$ we have that:

$$b = e_{i_1} \cdots e_{i_k} \otimes \omega$$

hence:

$$\phi(e_{i_1} \cdots e_{i_k}) = e_{i_1} \cdots e_{i_k} \otimes \omega$$

so ϕ is as surjection. Since $\dim_{\mathbb{C}} \mathbb{C}\mathrm{l}(n+2) = \dim_{\mathbb{C}} \mathbb{C}\mathrm{l}(n) \otimes_{\mathbb{C}} \mathbb{C}\mathrm{l}(2)$ the first isomorphism follows by rank nullity. The second isomorphism follows by **Example 2.2.3**. □

Lemma 2.2.14. *Let V and W be finite dimensional $\mathbb{K} = \mathbb{R}, \mathbb{C}$ vector spaces. Then:*

$$End(V) \otimes_{\mathbb{K}} End(W) \cong End(V \otimes_{\mathbb{K}} W)$$

Proof. Consider the linear map:

$$\phi : \mathrm{End}(V) \otimes_{\mathbb{K}} \mathrm{End}(W) \longrightarrow \mathrm{End}(V \otimes_{\mathbb{K}} W) \tag{2.2.7}$$

given on simple tensors by:

$$A \otimes B \longrightarrow \phi(A \otimes B)$$

where $\phi(A \otimes B)$ acts on simple tensors $v \otimes w \in V \otimes_{\mathbb{K}} W$ by:

$$\phi(A \otimes B)(v \otimes w) = A(v) \otimes B(w)$$

Note that this map is homomorphism as $\mathrm{Id}_V \otimes \mathrm{Id}_W$ clearly maps to the identity on $V \otimes_{\mathbb{K}} W$, and for any $A_1 \otimes B_1$, $A_1 \otimes B_2 \in \mathrm{End}(V) \otimes_{\mathbb{C}} \mathrm{End}(W)$ satisfies:

$$\begin{aligned}\phi(A_1 \otimes B_1) \circ \phi(A_2 \otimes B_2)(v,w) =& A_1 \circ A_1(v) \otimes B_1 \circ B_2(v,w)\\ =&\phi(A_1 A_2 \otimes B_1 B_2)\\ =&\phi((A_1 \otimes B_1) \cdot (A_2 \otimes B_2))\end{aligned}$$

Let $\{e_i\}$ and $\{f_i\}$ be basis for V and W respectively. Denote the dual basis for each by $\{e^i\}$ and $\{f^i\}$. Then any elements $A \in \mathrm{End}(V)$ $B \in \mathrm{End}(W)$ can be written as:

$$A = A^i_j e_i \otimes e^j \qquad B = B^l_k f_l \otimes f^k$$

since $\mathrm{End}(V) \cong V \otimes_{\mathbb{K}} V^*$, and likewise for $\mathrm{End}(W)$. Thus we have that any $C \in \mathrm{End}(V) \otimes_{\mathbb{K}} \mathrm{End}(W)$ can be written as:

$$\begin{aligned}C =& \sum_n A_n \otimes B_n\\ =&\sum_n (A^i_{nj} e_i \otimes e^j) \otimes (B^l_{nk} f_l \otimes f^k)\\ =&\sum_n A^i_{nj} B^l_{nk} (e_i \otimes e^j) \otimes (f_l \otimes f^k)\end{aligned}$$

For each i, j, l, k let:

$$A^i_j B^l_k = \sum_n A^i_{nj} B^l_{nk}$$

then:

$$C = A^i_j B^l_k (e_i \otimes e^j) \otimes (f_l \otimes f^k)$$

Now suppose that $\phi(C) = 0$, then:

$$\phi(C)(e_p \otimes f_q) = A^i_p B^l_q e_i \otimes f_l = 0$$

Contracting the first term with e^i and the second with f^m we obtain that:

$$A^i_p B^l_q = 0$$

Repeating this process for all possible combinations of i, j, p, q we see that $C = 0$, as every component is zero, so ϕ is injective. Now, let $\dim_{\mathbb{K}} V = n$ and $\dim_{\mathbb{K}} W = m$. Then, $\dim_{\mathbb{K}} \mathrm{End}(V) = n^2$ and $\dim_{\mathbb{K}} \mathrm{End}(W) = m^2$, so $\dim_{\mathbb{K}} \mathrm{End}(V) \otimes_{\mathbb{K}} \mathrm{End}(W) = n^2 m^2$. Similarly $\dim_{\mathbb{K}} V \otimes_{\mathbb{K}} W = mn$, and $\dim_{\mathbb{K}} \mathrm{End}(V \otimes_{\mathbb{K}} W) = n^2 m^2$. The claim then follows by rank nullity.

Note that we would also prove this via the following chain of isomorphisms:

$$\begin{aligned}\mathrm{End}(V) \otimes_{\mathbb{K}} \mathrm{End}(W) \cong& (V \otimes_{\mathbb{K}} V^*) \otimes_{\mathbb{K}} (W \otimes_{\mathbb{K}} W^*)\\ \cong&(V \otimes_{\mathbb{K}} W) \otimes_{\mathbb{K}} (V^* \otimes_{\mathbb{K}} W^*)\\ \cong&(V \otimes_{\mathbb{K}} W) \otimes_{\mathbb{K}} (V \otimes_{\mathbb{K}} W)^*\\ \cong&\mathrm{End}(V \otimes_{\mathbb{K}} W)\end{aligned}$$

□

We now turn to our first structure theorem for Clifford algebras.

Theorem 2.2.4. *As complex algebras, the Clifford algebras $\mathbb{C}l(n)$ and it's even part are given by the following table of isomorphisms:*

Complex Clifford Algebra Isomorphisms			
n	$\mathbb{C}\mathrm{l}(n)$	$\mathbb{C}\mathrm{l}^0(n)$	N
even	$\mathrm{End}(\mathbb{C}^N)$	$\mathrm{End}\left(\mathbb{C}^{N/2}\right) \oplus \mathrm{End}\left(\mathbb{C}^{N/2}\right)$	$2^{n/2}$
odd	$\mathrm{End}\left(\mathbb{C}^N\right) \oplus \mathrm{End}\left(\mathbb{C}^N\right)$	$\mathrm{End}\left(\mathbb{C}^N\right)$	$2^{(n-1)/2}$

Proof. First note that by **Lemma 2.2.8** we have that:

$$\mathbb{C}\mathrm{l}^0(2) \cong \mathbb{C}\mathrm{l}(1)$$

With $\mathrm{End}(\mathbb{C}) \cong \mathbb{C}$, we have that $\mathbb{C}\oplus\mathbb{C}$ is an associative $\mathbb{C}$ algebra, where the unit element is $(1,1)$. Since $\{1, e_1\}$ is a basis for $\mathbb{C}\mathrm{l}(1)$, it follows that the assignment:

$$\begin{aligned} 1 &\longmapsto (1,1) \\ e_i &\longmapsto (i,-i) \end{aligned}$$

determines an algebra isomorphism $\phi : \mathbb{C}\mathrm{l}(1) \to \mathbb{C}\oplus\mathbb{C}$ as:

$$\rho(e_i)\cdot\rho(e_i) = (i,-i)(i,-i) = (-1,-1) = -q(e_1,e_1)\cdot(1,1)$$

so $\mathbb{C}\mathrm{l}^0 \cong \mathbb{C}\oplus\mathbb{C} \cong \mathrm{End}(\mathbb{C})\oplus\mathrm{End}(\mathbb{C})$.

We proceed by cases, and induction. Suppose that n is even, the first non trivial base case is $n = 2$, then by **Example 2.2.3**, we have that:

$$\mathbb{C}\mathrm{l}(2) \cong \mathrm{End}\left(\mathbb{C}^2\right)$$

Assuming the base case, suppose the nth case holds, where n is even. For $n+2$, we have that by **Lemma 2.2.12** and **Lemma 2.2.13**:

$$\begin{aligned} \mathbb{C}\mathrm{l}(n+2) \cong& \mathbb{C}\mathrm{l}(n)\otimes_{\mathbb{C}}\mathbb{C}\mathrm{l}(2) \\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{n/2}}\right)\otimes_{\mathbb{C}}\mathrm{End}\left(\mathbb{C}^2\right) \\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{n/2}}\otimes_{\mathbb{C}}\mathbb{C}^2\right) \\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{(n+2)/2}}\right) \end{aligned}$$

hence for even n

$$\mathbb{C}\mathrm{l}(n) \cong \mathrm{End}\left(\mathbb{C}^N\right)$$

where $N = 2^{n/2}$, as desired.

We now move to the n odd case before proving the even part of the Clifford algebra. By our work above we have that:

$$\mathbb{C}\mathrm{l}(1) \cong \mathrm{End}(\mathbb{C})\oplus\mathrm{End}(\mathbb{C})$$

Assuming the nth case, we see that by **Lemma 2.2.12** and **Lemma 2.2.13**:

$$\begin{aligned} \mathbb{C}\mathrm{l}(n+2) \cong& \mathbb{C}\mathrm{l}(n)\otimes_{\mathbb{C}}\mathbb{C}\mathrm{l}(2) \\ \cong& \left(\mathrm{End}\left(\mathbb{C}^{2^{(n-1)/2}}\right)\mathrm{End}\left(\mathbb{C}^{2^{(n-1)/2}}\right)\right)\otimes_{\mathbb{C}}\mathrm{End}(\mathbb{C}^2) \\ \cong& \left(\mathrm{End}\left(\mathbb{C}^{2^{(n-1)/2}}\right)\otimes_{\mathbb{C}}\mathrm{End}(\mathbb{C}^2)\right)\oplus\left(\mathrm{End}\left(\mathbb{C}^{2^{(n-1)/2}}\right)\otimes_{\mathbb{C}}\mathrm{End}(\mathbb{C}^2)\right) \\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{(n+1)/2}}\right)\oplus\mathrm{End}\left(\mathbb{C}^{2^{(n+1)/2}}\right) \end{aligned}$$

hence for odd n:

$$\mathbb{C}\mathrm{l}(n) \cong \mathrm{End}\left(\mathbb{C}^N\right)\oplus\mathrm{End}\left(\mathbb{C}^N\right)$$

where $N = 2^{(n-1)/2}$ as desired.

Now we again assume n is even, so $n-1$ is odd. It follows from the preceding result, and **Lemma 2.2.8** that:

$$\begin{aligned} \mathbb{C}\mathrm{l}^0(n) \cong& \mathbb{C}\mathrm{l}(n-1) \\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{(n-2)/2}}\right)\oplus\mathrm{End}\left(\mathbb{C}^{2^{(n-2)/2}}\right) \end{aligned}$$

We see that:

$$2^{(n-2)/2} = 2^{n/2-1} = \frac{1}{2}2^{n/2}$$

hence for even n:

$$\mathbb{C}\mathrm{l}^0(n)\mathrm{End}\left(\mathbb{C}^{N/2}\right) \oplus \mathrm{End}\left(\mathbb{C}^{N/2}\right)$$

with $N = 2^{n/2}$.

If n is odd, then $n-1$ is even. It follows from the even case, and from **Lemma 2.2.8** that:

$$\begin{aligned}\mathbb{C}\mathrm{l}^0(n) \cong& \mathbb{C}\mathrm{l}(n-1)\\ \cong& \mathrm{End}\left(\mathbb{C}^{2^{(n-1)/2}}\right)\end{aligned}$$

hence for odd n:

$$\mathbb{C}\mathrm{l}^0(n) \cong \mathrm{End}(\mathbb{C}^N)$$

where $N = 2^{(n-1)/2}$ □

We have a similar, albeit more complicated, result for real Clifford algebras, which we cite without proof.

Theorem 2.2.5. *Let $p = s - t$, and $s + t = n$. Then as real associative algebras, the Clifford algebras $Cl(t, s)$ and it's even parts are given by the following tables :*

Real Clifford Algebra Isomorphisms		
p mod 8	$\mathrm{Cl}(t,s)$	N
0	$\mathrm{End}\left(\mathbb{R}^N\right)$	$2^{n/2}$
1	$\mathrm{End}\left(\mathbb{C}^n\right)$	$2^{(n-1)/2}$
2	$\mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-2)/2}$
3	$\mathrm{End}\left(\mathbb{H}^N\right) \oplus \mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-3)/2}$
4	$\mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-2)/2}$
5	$\mathrm{End}\left(\mathbb{C}^N\right)$	$2^{(n-1)/2}$
6	$\mathrm{End}\left(\mathbb{R}^N\right)$	2^{n}
7	$\mathrm{End}\left(\mathbb{R}^N\right) \oplus \mathrm{End}\left(\mathbb{R}^N\right)$	$2^{(n-1)/2}$

Real Even Clifford Algebra Isomorphisms		
p mod 8	$\mathrm{Cl}^0(t,s)$	N
0	$\mathrm{End}\left(\mathbb{R}^N\right) \oplus \mathrm{End}\left(\mathbb{R}^N\right)$	$2^{(n-2)/2}$
1	$\mathrm{End}\left(\mathbb{R}^n\right)$	$2^{(n-1)/2}$
2	$\mathrm{End}\left(\mathbb{C}^N\right)$	$2^{(n-2)/2}$
3	$\mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-3)/2}$
4	$\mathrm{End}\left(\mathbb{H}^N\right) \oplus \mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-4)/2}$
5	$\mathrm{End}\left(\mathbb{H}^N\right)$	$2^{(n-3)/2}$
6	$\mathrm{End}\left(\mathbb{C}^N\right)$	$2^{(n-2)/2}$
7	$\mathrm{End}\left(\mathbb{R}^N\right)$	$2^{(n-1)/2}$

2.2.4 Spinor Representation

Definition 2.2.13. The vector space of **Dirac spinors** is $\Delta_n = \mathbb{C}^N$, where N is given by **Theorem 2.2.4**. If n is even, the **(even) Dirac spinor representation** is given by the isomorphism:

$$\rho : \mathbb{C}\mathrm{l}(n) \longrightarrow \mathrm{End}(\Delta_n)$$

If n is odd, the **(odd) Dirac spinor representation** is given by the homomorphism:

$$\rho : \mathbb{C}\mathrm{l}(n) \longrightarrow \mathrm{End}(\Delta_n) \oplus \mathrm{End}(\Delta_n) \longrightarrow \mathrm{End}(\Delta_n)$$

We also have induced representation of $\mathrm{Cl}(t, s)$ on Δ_n, when $s + t = n$, given by restricting ρ, or $\pi_1 \circ \rho$ to the real part of:

$$\mathrm{Cl}(t,s) \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}\mathrm{l}(n)$$

Definition 2.2.14. The bilinear map:

$$\begin{aligned}\mathbb{R}^{t,s} \times \Delta_n &\longrightarrow \Delta_n\\ (v,\psi) &\longmapsto v \cdot \psi = \rho(\gamma(v)) \cdot \psi\end{aligned}$$

is called **mathematical Clifford multiplication** of a spinor and vector. **Physical Clifford multiplication** is given by $(-i)$ times mathematical Clifford multiplication. Via the isomorphism in **Theorem 2.2.3**:

$$f^{-1} : \Lambda(\mathbb{R}^{t,s}) \longrightarrow \mathrm{Cl}(t,s)$$

we define **Clifford multiplication of forms** by:

$$\begin{aligned}\Lambda^k(\mathbb{R}^{t,s}) \times \Delta_n &\longrightarrow \Delta_n \\ (\omega, \psi) &\longmapsto \omega \cdot \psi = \rho(f^{-1}(\omega)) \cdot \psi\end{aligned}$$

We want to see how restriction of ρ to the even subalgebra of $\mathbb{C}\mathrm{l}(n)$ behaves.

Corollary 2.2.5. *If n is odd then the restriction of $\rho : \mathbb{C}l(n) \to \Delta_n$ to $\mathbb{C}l^0(n)$, i.e. the reduced representation of $\mathbb{C}l^0(n)$ on Δ_n, is irreducible:*

$$\mathbb{C}l^0(n) \cong End(\Delta_n)$$

If instead n is even then the induced representation of $\mathbb{C}l^0(n)$ on Δ_n splits into ***left handed (positive)*** *and* ***right handed (negative) Weyl Spinors****:*

$$\mathbb{C}l^0 \cong End(\Delta_n^+) \oplus End(\Delta_n^-)$$

where $\Delta_n^\pm \cong \Delta_{n-1} = \mathbb{C}^{N/2}$.

Proof. This follows from **Theorem 2.2.4**. □

We clarify what we mean by $\Delta_n^\pm$ with the following proposition:

Proposition 2.2.3. *Let n be even, $\rho : \mathbb{C}l(n) \to End(\Delta_n)$ the Dirac spinor representation, and Γ_{n+1} the chirality element. Then the following hold:*

a) $\Delta_n^\pm$ can be identified with the ± 1 eigenspaces of Γ_{n+1} on Δ_n.

b) The induced representation of $\mathbb{C}l^0(n)$ maps $\Delta_n^\pm$ to itself, while elements in $\mathbb{C}l^1(n)$ map $\Delta^\pm$ to $\Delta^\mp$. It follows that:

$$\begin{aligned}\mathbb{C}l^0(n) &\cong Hom(\Delta_n^+, \Delta_n^+) \oplus Hom(\Delta_n^-, \Delta_n^-) \\ \mathbb{C}l^1(n) &\cong Hom(\Delta_n^+, \Delta_n^-) \oplus Hom(\Delta_n^-, \Delta_n^+)\end{aligned}$$

Proof. We see that $\Gamma_{n+1}\Gamma_{n+1} = \mathrm{Id}_{\Delta_n}$, hence Γ_{n+1} has two eigenvalues 1 and -1. We call the -1 eigenspace Δ_n^-, and the $+1$ eigenspace Δ_n^+. Furthermore, we see that Γ_{n+1} satisfies the polynomial:

$$t^2 - 1 = 0$$

Recall that the minimal polynomial of Γ_{n+1} is given by:

$$m_\Gamma = (t-1)^{j_{+1}}(t+1)^{j_{-1}}$$

where j_{+1} and j_{-1} are the size of the Jordan blocks of the eigenvalues $+1$ and -1. The minimal polynomial must divide every polynomial P satisfying $P(\Gamma_{n+1}) = 0$, hence m_Γ divides $t^2 - 1$, but $t^2 - 1 = (t-1)(t+1)$, so the Jordan blocks of the eigenvalues must be 1. This implies that Γ_{n+1} is diagonalizable and admits a full set of linearly independent eigenvectors which span Δ_n. It then follows that:

$$\Delta_n \cong \Delta_n^+ \oplus \Delta_n^-$$

By **Lemma 2.2.10** we have that for all $1 \le i, j \le n$, and all $\psi \in \Delta_n$:

$$[\Gamma_{n+1}, \Gamma_i\Gamma_j]\psi = 0$$

implying that:

$$\Gamma_{n+1}\Gamma_i\Gamma_j\psi = \Gamma_i\Gamma_j\Gamma_{n+1}\psi$$

Suppose that $\psi \in \Delta_n^{\pm}$ then:

$$\Gamma_{n+1}\Gamma_i\Gamma_j\psi = \pm\Gamma_i\Gamma_j\psi$$

so $\Gamma_i\Gamma_j\psi \in \Delta_n^{\pm}$. Let $k-1<n$ be even, and assume that for all $\psi \in \Delta_n^{\pm}$, and all $1 \leq i_1 < \cdots < i_{k-1} \leq n$:

$$\Gamma_{i_1}\cdots\Gamma_{i_{k-1}}\psi \in \Delta_n^{\pm}$$

Denote $\Gamma_{i_1}\cdots\Gamma_{i_{k-1}}\psi$ by φ, then for l,j such that $l \neq j \neq i_1,\ldots,i_{k-1}$:

$$[\Gamma_{n+1},\Gamma_l\Gamma_j]\varphi = 0$$

so:

$$\Gamma_{n+1}\Gamma_l\Gamma_j\varphi = \pm\Gamma_l\Gamma_j\varphi$$

Note that $k+1$ is even, thus by induction it follows that the induced representation of $\mathbb{C}\mathrm{l}^0(n)$ preserves the eigenspaces $\Delta_n^{\pm}$.

Similarly, we have that for all $1 \leq i \leq n$ and all $\psi \in \Delta_n$:

$$\{\Gamma_{n+1},\Gamma_i\}\psi = 0$$

It follows that if $\psi \in \Delta_n^{\pm}$ then:

$$\Gamma_{n+1}\Gamma_i\psi = \mp\Gamma_i\psi$$

so $\Gamma_i\psi \in \Delta_n^{\mp}$. Proceeding by induction, let $k-1<n-1$ be odd, and assume that for all $\psi \in \Delta_n^{\pm}$ and $1 \leq i_1 < \cdots < i_{k-1} \leq n$:

$$\Gamma_{i_1}\cdots\Gamma_{i_{k-1}}\psi \in \Delta_n^{\mp}$$

Denote $\Gamma_{i_1}\cdots\Gamma_{i_{k-1}}\psi$ by φ, then for $l \neq j \neq i_1,\ldots,i_{k-1}$:

$$[\Gamma_{n+1},\Gamma_l\Gamma_j]\varphi = 0$$

so:

$$\begin{aligned}\Gamma_{n+1}\Gamma_l\Gamma_j\varphi &= \mp\,\Gamma_l\Gamma_j\varphi\\ &= \mp\,\Gamma_l\Gamma_j\Gamma_{i_1}\cdots\Gamma_{i_{k-1}}\varphi\end{aligned}$$

Again note that $k+1$ is odd, so it follows that the induced representation of $\mathbb{C}\mathrm{l}^1(n)$ on Δ_n maps $\Delta_n^{\pm}$ to Δ. Combining these two results we obtain that:

$$\begin{aligned}\mathbb{C}\mathrm{l}^0(n) \subset &\mathrm{Hom}(\Delta_n^+,\Delta_n^+) \oplus \mathrm{Hom}(\Delta_n^+,\Delta_n^+)\\ \mathbb{C}\mathrm{l}^1(n) \subset &\mathrm{Hom}(\Delta_n^+,\Delta_n^-) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^+)\end{aligned}$$

However, we have:

$$\mathbb{C}\mathrm{l}^0(n) \oplus \mathbb{C}\mathrm{l}^1(n) \cong \mathbb{C}\mathrm{l}(n) \cong \mathrm{End}(\Delta_n)$$

while:

$$\mathrm{End}(\Delta_n) \cong \mathrm{Hom}(\Delta_n^+,\Delta_n^+) \oplus \mathrm{Hom}(\Delta_n^+,\Delta_n^+) \oplus \mathrm{Hom}(\Delta_n^+,\Delta_n^-) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^+)$$

Thus if $A \in \mathrm{Hom}(\Delta_n^+,\Delta_n^+) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^-)$, then $A \in \mathbb{C}\mathrm{l}^0(n) \oplus \mathbb{C}\mathrm{l}^1(n)$. However, A must lie in $\mathbb{C}\mathrm{l}^0(n) \subset \mathbb{C}\mathrm{l}^0(n) \oplus \mathbb{C}\mathrm{l}^1(n)$ as otherwise $A \in \mathrm{Hom}(\Delta_n^+,\Delta_n^-) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^+)$. Hence:

$$\mathbb{C}\mathrm{l}^0(n) \cong \mathrm{Hom}(\Delta_n^+,\Delta_n^+) \oplus \mathrm{Hom}(\Delta_n^+,\Delta_n^+) \tag{2.2.8}$$

A similar argument demonstrates that if $A \in \mathrm{Hom}(\Delta_n^+,\Delta_n^-) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^+)$, then $A \in \mathbb{C}\mathrm{l}^1(n)$, hence:

$$\mathbb{C}\mathrm{l}^1(n) \cong \mathrm{Hom}(\Delta_n^+,\Delta_n^-) \oplus \mathrm{Hom}(\Delta_n^-,\Delta_n^+)$$

implying b).

Finally, to see that the $\Delta_n^\pm$ are the same as the ones from **Corollary 2.2.5**, let $\dim_\mathbb{C} \Delta_n^+ = a$ and $\dim_\mathbb{C} \Delta_n^- = b$. We see that:

$$a + b = N$$

while by (2.2.8):

$$a^2 + b^2 = \frac{N^2}{2}$$

The first equation implies that $b = N - a$, so:

$$a^2 + (N-a)^2 = \frac{N^2}{2} \Longrightarrow a^2 + \frac{N^2}{4} - Na = 0 \Longrightarrow \left(a - \frac{N}{2}\right)^2 = 0 \Longrightarrow a = \frac{N}{2}$$

thus $b = \frac{N}{2}$ as well, and $\Delta_n^\pm \cong \Delta_{n-1} = \mathbb{C}^{N/2}$, implying a). □

Note that importantly this implies that if n is even, then Clifford multiplication of a spinor $\psi \in \Delta_n^\pm$ with a vector v, has image in $\Delta_n^\mp$. Since we are in general interested in four manifolds of Lorentzian signature, i.e. those equipped with pseudo-Riemannian metric of signature $(- + ++)$, this decomposition of Weyl spinors will play an important role in the chapters to come.

2.2.5 The Spin Group

In chapter 1.2, we constructed Lie groups by first examining the the open subset $GL_n(\mathbb{R}) \subset \mathrm{Mat}_{n\times n}(\mathbb{R}) \cong \mathrm{End}(\mathbb{R}^n)$. In this section, we follow a similar, albeit more complicated path, to construct the groups Spin and Pin, but by first examining open subsets of Clifford algebras. In this section, we mildly deviate from Hamilton's *Mathematical Gauge Theory*, and instead more closely follow Michelsohn and Lawson's *Spin Geometry*, and Atiyah, Bott, and Shapiro's *Clifford Modules*

Definition 2.2.15. The **group of invertible elements** in the Clifford algebras $\mathrm{Cl}(t,s)$ and $\mathbb{C}\mathrm{l}(n)$ are the sets:

$$\begin{aligned} \mathrm{Cl}^\times(t,s) =& \{v \in \mathrm{Cl}(t,s) : \exists u \in \mathrm{Cl}(t,s), v \cdot u = u \cdot v = 1\} \\ \mathbb{C}\mathrm{l}^\times(n) =& \{v \in \mathbb{C}\mathrm{l}(n) : \exists u \in \mathbb{C}\mathrm{l}(n), v \cdot u = v \cdot u = 1\} \end{aligned}$$

It is easily verified that the above sets are indeed groups.

Lemma 2.2.15. *The group of invertible elements of $\mathbb{C}l(n)$ is an open subset of $\mathbb{C}l(n)$.*

Proof. From **Theorem 2.2.4** we have that if n is even:

$$\mathbb{C}\mathrm{l}(n) \cong \mathrm{End}\left(\mathbb{C}^N\right)$$

The group of invertible elements in $\mathrm{End}(\mathbb{C}^N)$ is the set:

$$\{A \in \mathrm{Mat}_{N\times N}(\mathbb{C}) : \det(A) \neq 0\}$$

it then follows that $\mathbb{C}\mathrm{l}^\times(n)$ is an open subset of $\mathbb{C}\mathrm{l}(n)$. If n is odd, we have that by the same theorem:

$$\mathbb{C}\mathrm{l}(n) \cong \mathrm{End}\left(\mathbb{C}^N\right) \oplus End\left(\mathbb{C}^N\right)$$

Recall that multiplicative structure on $\mathrm{End}\left(\mathbb{C}^N\right) \oplus End\left(\mathbb{C}^N\right)$ is given by:

$$(A_1, B_1) \cdot (A_2, B_2) = (A_1 \cdot A_2, B_1 \cdot B_2)$$

hence the group of invertible elements in $\mathrm{End}\left(\mathbb{C}^N\right) \oplus \mathrm{End}\left(\mathbb{C}^N\right)$ is given by:

$$\{(A, B) \in \mathrm{Mat}_{N\times N}(\mathbb{C}) \oplus \mathrm{Mat}_{N\times N}(\mathbb{C}) : \det(A) \neq 0 \text{ and } \det(B) \neq 0\}$$

It then follows that $\mathbb{C}\mathrm{l}^\times(n)$ is the intersection of two open subsets of $\mathrm{End}\left(\mathbb{C}^N\right) \oplus End\left(\mathbb{C}^N\right)$, and thus an open subset of $\mathbb{C}\mathrm{l}(n)$. □

Lemma 2.2.16. *The group of invertible element of Cl(t,s) is an open subset of Cl(t,s), and thus a Lie group.*

Proof. Let $s+t=n$, then by **Lemma 2.2.7**:

$$\mathbb{C}\mathrm{l}(n) \cong \mathrm{Cl}(t,s) \otimes_{\mathbb{R}} \mathbb{C}$$

We can thus decompose any $w \in \mathbb{C}\mathrm{l}(n)$ into $u+iv$, where $u,v \in \mathrm{Cl}(t,s)$. Therefore we have that:

$$\mathrm{Cl}^{\times}(t,s) = \mathrm{Cl}(t,s) \cap \mathbb{C}\mathrm{l}^{\times}(n) \tag{2.2.9}$$

Equipping $\mathrm{Cl}(t,s) \subset \mathbb{C}\mathrm{l}(n)$ with the subspace topology, it follows that $\mathrm{Cl}^{\times}(t,s)$ is an open subset of $\mathrm{Cl}(t,s)$, as by **Lemma 2.2.14** $\mathbb{C}\mathrm{l}^{\times}(n)$ is an open subset of $\mathbb{C}\mathrm{l}(n)$.

The Clifford algebra $\mathrm{Cl}(t,s)$ is a real associative algebra, and thus has the structure of an $\mathbb{R}$-linear vector space. It follows that $\mathrm{Cl}^{\times}(s,t)$ is then an open submanifold of $\mathrm{Cl}(t,s)$. Multiplication by any two elements in $\mathrm{Cl}(t,s)$ is a bilinear map and thus smooth, so multiplication in $\mathrm{Cl}^{\times}(s,t)$ is smooth. By **Lemma 1.2.3**, we then have that $\mathrm{Cl}^{\times}(s,t)$ is a Lie group as desired. □

We see that since $\mathrm{Cl}^{\times}(t,s)$ is open in the associative algebra $\mathrm{Cl}(t,s)$, thus the tangent space for all $x \in \mathrm{Cl}^{\times}(t,s)$ satisfies $T_x\mathrm{Cl}^{\times}(t,s) \cong \mathrm{Cl}(t,s)$. In particular the Lie algebra of $\mathrm{Cl}^{\times}(t,s)$, denoted $\mathfrak{cl}^{\times}(t,s)$ is isomorphic to $\mathrm{Cl}(t,s)$ where the Lie bracket is the commutator in $\mathrm{Cl}(t,s)$:

$$[z,y] = z \cdot y - y \cdot z$$

Recall the adjoint representation of a group G given by:

$$\begin{aligned} G &\longrightarrow \mathrm{Aut}(\mathfrak{g}) \\ g &\longmapsto \mathrm{Ad}_g = (L_g \circ R_{g^{-1}})_* \end{aligned}$$

We define a similar representation of $\mathrm{Cl}^{\times}(t,s)$ with use of the automorphism α from **Lemma 2.2.6**.

Definition 2.2.16. The **twisted adjoint representation** of $\mathrm{Cl}^{\times}(t,s)$ on $\mathfrak{cl}^{\times}(t,s)$, denoted, $\mathrm{Ad}^{\times}$ is defined by it's action on elements $a \in \mathfrak{cl}^{\times}(t,s) \cong \mathrm{Cl}(t,s)$ as follows:

$$\mathrm{Ad}^{\times}_x(a) = \alpha(x) \cdot a \cdot x^{-1}$$

It is clear that if $x \in \mathrm{Cl}^{\times}(t,s)$, then $\alpha(x) \in \mathrm{Cl}^{\times}(t,s)$, so $\mathrm{Ad}^{\times}_x$ is an automorphism. Since α is an automorphism, and thus a smooth map $\mathrm{Cl}^{\times}(t,s) \to \mathrm{Cl}^{\times}(t,s)$, a similar argument to the one in **Theorem 1.2.6** demonstrates that:

$$\begin{aligned} \mathrm{Ad}^{\times} : \mathrm{Cl}^{\times}(t,s) &\longrightarrow \mathrm{Aut}(\mathfrak{cl}^{\times}(t,s)) \\ x &\longmapsto \mathrm{Ad}^{\times}_x \end{aligned}$$

is indeed a representation of $\mathrm{Cl}^{\times}(t,s)$ on $\mathfrak{cl}^{\times}(t,s)$. In particular, it is smooth.

Proposition 2.2.4. *The subset:*

$$Cl^*(t,s) = \{x \in Cl^{\times}(t,s) : \forall v \in \mathbb{R}^{t,s}, Ad^{\times}_x(v) \in \mathbb{R}^{t,s}\}$$

is a subgroup of $Cl^{\times}(t,s)$. Equipped the subspace topology, $Cl^*(t,s)$ is a topological group.

Proof. We first show that $\mathrm{Cl}^*(t,s)$ is a group. It is clearly closed under multiplication, as for any $x,y \in \mathrm{Cl}^*(t)$, and all $v \in \mathbb{R}^{t,s}$ we have that:

$$\mathrm{Ad}^{\times}_{xy}(v) = \mathrm{Ad}^{\times}_x \circ \mathrm{Ad}^{\times}_y(v) \in \mathbb{R}^{t,s}$$

Furthermore, if $x \in \mathrm{Cl}^*(t,s)$, then $x^{-1} \in \mathrm{Cl}^*(t,s)$, as for all $v \in \mathbb{R}^{t,s}$ we have that there exists a unique $u \in \mathbb{R}^{t,s}$ such that:

$$\mathrm{Ad}^{\times}_x(v) = u$$

then:

$$\mathrm{Ad}^{\times}_{x^{-1}}(u) = v$$

Since $\mathrm{Ad}^{\times}$ is an automorphism, it follows that for any $v \in \mathbb{R}^{t,s}$, $\mathrm{Ad}_{x^{-1}}(v) \in \mathbb{R}^{t,s}$, so $x^{-1} \in \mathrm{Cl}^*(t,s)$.

To see that $\mathrm{Cl}^*(t,s)$ is a topological group, note that multiplication and inversion in $\mathrm{Cl}^{\times}(t,s)$ are smooth, and thus continuous. Let μ, and i be the multiplication and inversion maps, then the restriction of multiplication and inversion to $\mathrm{Cl}^*(t,s)$ are continuous in the subspace topology, as for any open $U \subset \mathrm{Cl}^*(t,s)$, we have that:

$$U = V \cap \mathrm{Cl}^*(t,s)$$

for some open $V \subset \mathrm{Cl}^{\times}(t,s)$. We see that if:

$$\mu^{-1}(V) = \bigcup_i Y_i \times V_i$$

for some open open sets $Y_i, V_i \subset \mathrm{Cl}^{\times}(t,s)$. We also have that:

$$\mu^{-1}(\mathrm{Cl}^*(t,s)) = \mathrm{Cl}^*(t,s) \times \mathrm{Cl}^*(t,s)$$

hence:

$$\begin{aligned}
\mu^{-1}(U) =& \mu^{-1}(V \cap \mathrm{Cl}^*(t,s)) \\
=& \mu^{-1}(V) \cap \mu^{-1}(\mathrm{Cl}^*(t,s)) \\
=& \left(\bigcup_i Y_i \times V_i\right) \cap (\mathrm{Cl}^*(t,s) \times \mathrm{Cl}^*(t,s))
\end{aligned}$$

So by definition of the product and subspace topology, $\mu^{-1}(U)$ is open in $\mathrm{Cl}^*(t,s) \times \mathrm{Cl}^*(t,s)$, thus multiplication is continuous. A similar argument shows that inversion is a continuous map, so $\mathrm{Cl}^*(t,s)$ is a topological group. $\square$

Since $\mathrm{Cl}^*(t,s)$ is by definition the group of invertible elements which preserve the subspace $\mathbb{R}^{t,s}$, we have that $\mathrm{Ad}^{\times}$ is a representation of $\mathrm{Cl}^*(t,s)$ on $\mathbb{R}^{t,s}$. In this case, the representation is only continuous, as we have not demonstrated that $\mathrm{Cl}^*(t,s)$ is a Lie group.

Proposition 2.2.5. *The kernel of $Ad^{\times} : Cl^*(t,s) \to Aut(\mathbb{R}^{t,s})$ is equal to $\mathbb{R}^*$, i.e. the subgroup of non zero real scalars in $Cl^*(t,s)$.*

Proof. It is easy to see that if $x \in \mathbb{R}^*$ then $x \in \mathrm{Cl}^*(t,s)$ as for all $v \in \mathbb{R}^{t,s}$:

$$\alpha(x)vx^{-1} = xx^{-1}v = v$$

incidentally this also shows that $\mathbb{R}^* \subset \ker \mathrm{Ad}^{\times}$. Suppose now that $x \in \ker \mathrm{Ad}^{\times}$, then:

$$\alpha(x)vx^{-1} = v \Longrightarrow \alpha(x)v = vx$$

Split x into it's even and odd parts x^0 and x^1 respectively, then:

$$x^0 v = v x^0 \qquad \text{and} \qquad x^1 v = -v x^1 \tag{2.2.10}$$

Let $\{e_i\}$ be the standard basis for V, if $s+t$ is even we can write x^0 as:

$$x_0 = \sum_{k=0}^{(s+t)/2} \sum_{i_1 < \cdots i_{2k}} a^{i_1 \cdots i_k} e_{i_1} \cdots e_{i_{2k}}$$

and if $s+t$ is odd as:

$$x_0 = \sum_{k=0}^{(s+t-1)/2} \sum_{i_1 < \cdots i_k} a^{i_1 \cdots i_{2k}} e_{i_1} \cdots e_{i_{2k}}$$

For some $1 \leq i \leq s_t$, collect all terms that don't contain e_i into a new element $a^0 \in \mathrm{Cl}^0(t,s)$, and all other terms into an element b. We can reorder b such that $b = e_i a^1$ for some element a_1 which clearly will lie in $\mathrm{Cl}^1(t,s)$. We then have that $x^0 = a^0 + e_i a^1$, since (2.2.10) holds for all v, let $v = e_i$ then:

$$(a^0 + e_i a^1)e_i = e_i a^0 - e_i e_i a^1 = e_i + \eta_{ii} a^1$$

while the other side of (2.2.10) satisfies:

$$e_i(a^0 + e_i a^1) = e_i a^0 - \eta_{ii} a^1$$

Since these two expression are equal, we obtain that $a^1 = -a^1$, so x_0 has no terms that contain e_i. Since i was chosen arbitrarily, it follows that x^0 is a scalar. An entirely analogous argument demonstrates that x^1 is zero, as x^1 is odd, implying the claim. □

Proposition 2.2.6. *The map:*

$$\begin{aligned} N : Cl(t,s) &\longrightarrow Cl(t,s) \\ x &\longmapsto x\alpha(x^t) \end{aligned}$$

restricted to $Cl^*(t,s)$ is a continuous group homomorphism:

$$N : Cl^*(t,s) \longrightarrow \mathbb{R}^*$$

satisfying $N(\alpha(x)) = N(x)$, where $\mathbb{R}^*$ is the topological group of nonzero elements in $\mathbb{R}$.

Proof. First note that if $x \in \mathrm{Cl}^*(t,s)$, we have that for all v:

$$\alpha(x)vx^{-1} = u$$

for some $u \in \mathbb{R}^{t,s}$. Apply the transpose to both side, then since it is clear that α and t commute:

$$(x^{-1})^t v\alpha(x^t) = u$$

We also see that:

$$(xx^{-1})^t = (x^{-1})^t x^t = e \Longrightarrow (x^{-1})^t = (x^t)^{-1}$$

So we now have:

$$\alpha(x)vx^{-1} = (x^t)^{-1} v\alpha(x^t)$$

Multiply both sides by x on the right, and $\alpha(x)^{-1}$ on the left:

$$v = \alpha(x)^{-1}(x^t)^{-1} v\alpha(x^t)x$$

Examine the term acting on the left of v, since $\alpha \circ \alpha = \mathrm{Id}_{\mathrm{Cl}(t,s)}$ we have

$$\begin{aligned} \alpha(x^{-1})(x^t)^{-1} =& (x^t\alpha(x))^{-1} \\ =& (\alpha(\alpha(x^t)x))^{-1} \\ =& \alpha\left((\alpha(x^t)x)^{-1}\right) \end{aligned}$$

so:

$$\alpha\left((\alpha(x^t)x)^{-1}\right) v\alpha(x^t)x = v$$

We see that:

$$\mathrm{Ad}^{\times}_{(\alpha(x^t)x)^{-1}}(v) = \alpha\left((\alpha(x^t)x)^{-1}\right) v\alpha(x^t)x$$

so $(\alpha(x^t)x)^{-1} \in \ker \mathrm{Ad}^{\times}$. It follows that $\alpha(x^t)x \in \ker \mathrm{Ad}^{\times}$, and hence in $\mathbb{R}^*$. Since t is the identity on $\mathbb{R}^{t,s}$, we also have that $x^t\alpha(x) \in \ker \mathrm{Ad}^{\times}$, implying that $N(x^t) \in \ker \mathrm{Ad}^{\times}$. However, we see that for any $y \in \mathrm{Cl}^*(t,s)$, then $y^{-1} \in \mathrm{Cl}^*(t,s)$, hence:

$$\alpha(y^{-1})vy = v'$$

for some $v' \in \mathbb{R}^{t,s}$. Apply α and t to both sides, we obtain:

$$\alpha(y^t)v(y^t)^{-1} = v'$$

so $y^t \in \mathrm{Cl}^*(t)$. Since t is anti automorphism of $\mathrm{Cl}(t,s)$, and is in particular one to one and onto, it follows that t restricted $\mathrm{Cl}^*(t,s)$ is an anti automorphism[26] of $\mathrm{Cl}^*(t,s)$. Therefore, $N(x^t) \in \ker \mathrm{Ad}^\times$ implies that $N(x) \in \ker \mathrm{Ad}^\times$, so N restricted to $\mathrm{Cl}^*(t,s)$ has image in $\mathbb{R}^*$.

We see that N is a group homomorphism as for $x, y \in \mathrm{Cl}^*(t,s)$:

$$N(xy) = (xy)\alpha((xy)^t) = xy\alpha(y^t)\alpha(x^t) = xN(y)\alpha(x^t) = x\alpha(x^t)N(y) = N(x)\cdot N(y)$$

Furthermore:

$$N(\alpha(x)) = \alpha(x)\alpha(\alpha(x^t)) = \alpha(x)x^t = \alpha(x\alpha(x^t)) = \alpha(N(x)) = N(x)$$

so $N(\alpha(x)) = N(x)$.

Finally since t and α are continuous maps $\mathrm{Cl}^*(t,s) \to \mathrm{Cl}^*(t,s)$, and since multiplication in $\mathrm{Cl}^*(t,s)$ is continuous, it follows that N is continuous as the composition of continuous maps. □

Definition 2.2.17. We define the following subsets of $\mathbb{R}^{t,s}$:

$$\begin{aligned} S_-^{t,s} =& \{v \in \mathbb{R}^{t,s} : \eta(v,v) = -1\} \\ S_+^{t,s} =& \{v \in \mathbb{R}^{t,s} : \eta(v,v) = +1\} \\ S_\pm^{t,s} =& S_-^{t,s} \cup S_+^{t,s} \end{aligned}$$

With this definition, we can now construct the groups Pin, Spin and Spin^+:

Theorem 2.2.6. *The following subsets form Lie subgroups of the group $Cl^\times(t,s)$:*

$$Pin(t,s) = \{v_1 \cdots v_r : v_i \in S_\pm^{t,s}, r \geq 0\}$$

where $r = 0$ is the empty product equal to ± 1. We also set:

$$Pin(n) = Pin(0,n)$$

*We call $Pin(t,s)$ the **pin group**.*

Proof. We first define the following subgroup of $\mathrm{Cl}^*(t,s)$:

$$\tilde{\mathrm{Cl}}^*(t,s) = \{x \in \mathrm{Cl}^*(t,s) : \exists v_i \in \mathbb{R}^{t,s}, \eta(v_i,v_i) \neq 0, \text{ and } x = v_1 \cdots v_r\}$$

i.e. the subgroup generated by elements of $\mathbb{R}^{t,s}$. It is clear that this a group, and endowed with the subspace topology it is a topological group such that the restriction of N is a continuous group homomorphism.

We now show that $\mathrm{Pin}(t,s)$ is a group. By construction, $\mathrm{Pin}(t,s)$ contains the identity, as $1 \in \mathrm{Pin}(t,s)$. Furthermore, for any $x, y \in \mathrm{Pin}(t,s)$ we have that:

$$x \cdot y = v_1 \cdots v_{r_x} \cdot u_1 \cdots u_{r_y} \in \mathrm{Pin}(t,s)$$

for some $u_i, v_i \in S_\pm^{t,s}$ and $r_x, r_y \geq 0$. Finally, if $x = v_1 \cdots v_r$, and then x has an inverse given by:

$$x = (-1)^k\alpha((v_1 \cdots v_r)^t) = (-1)^{k+r}v_r \cdots v_1$$

where k is the number of $v_i \in S_-^{t,s}$. We check that this is an inverse:

$$\begin{aligned} xx^{-1} =& (-1)^{k+r}v_1 \cdots v_r \cdot v_r \cdots v_1 \\ =& (-1)^{k+r}(-1)^r\eta(v_1,v_1)\cdots\eta(v_r,v_r) \\ =& (-1)^{k+r}(-1)^{k+r} \\ =& 1 \end{aligned}$$

[26]In the sense of groups, not algebras.

Note that this implies that:

$$N(x) = \pm 1$$

so:

$$\mathrm{Pin}(t, s) \subset N^{-1}(1) \cup N^{-1}(-1)$$

where here N denotes the restriction of N to $\tilde{\mathrm{Cl}}^*(t, s)$. We want to show that $N^{-1}(1) \cup N^{-1}(-1) \subset \mathrm{Pin}(t, s)$, which, since N is continuous, would imply that $\mathrm{Pin}(t, s)$ is a closed subgroup of $\tilde{CL}^*(t, s)$. In particular, $\mathrm{Pin}(t, s)$ would be a closed subgroup of $\mathrm{Cl}^\times(t, s)$, so by **Theorem 1.2.1**, $\mathrm{Pin}(t, s)$ would be an embedded Lie subgroup of $\mathrm{Cl}^\times(t, s)$.

Let $x \in N^{-1}(1) \cup N^{-1}(-1)$, we have that for some $r \geq 0$ and $v_i \in \mathbb{R}^{t,s}$ such that $\eta(v_i, v_i) \neq 0$:

$$x = v_1 \cdots v_r$$

We want to show x can be rewritten as:

$$x = u_1 \cdots u_r$$

where each u_i satisfies $\eta(u_i, u_i) = \pm 1$. We prove this by induction on r, the base case $r = 1$ is trivial, as for all $v \in \mathbb{R}^{t,s}$:

$$N(v) = \eta(v, v) = \pm 1$$

Suppose the rth case, and let $x = v_1 \cdots v_r \cdot v_{r+1} \in N^{-1}(1) \cup N^{-1}(-1)$, then:

$$N(x) = N(v_1 \cdots v_r)N(v_{r+1}) = N(v_1 \cdots v_r)\eta(v_{r+1}, v_{r+1}) = \pm 1$$

Assume that $\eta(v_{r+1}, v_{r+1}) > 0$, then

$$\begin{aligned} N(\sqrt{\eta(v_{r+1}, v_{r+1})}v_1 \cdots v_r) =& \eta(v_{r+1}, v_{r+1})\eta(v_1, v_1)N(v_2 \cdots v_r) \\ =& \eta(v_{r+1}, \eta(v_{r+1}))N(v_1 \cdots v_r) \\ =& \pm 1 \end{aligned} \tag{2.2.11}$$

The inductive hypothesis then implies that for some $u_i \in \mathbb{R}^{t,s}$ such that $\eta(u_i, u_i) = \pm 1$:

$$\sqrt{\eta(v_{r+1}, v_{r+1})}v_1 \cdots v_r = u_1 \cdots u_r$$

Let:

$$u_{r+1} = \frac{1}{\sqrt{\eta(v_{r+1}, v_{r+1})}} v_{r+1}$$

then:

$$\eta(u_{r+1}, u_{r+1}) = 1$$

while:

$$u_1 \cdots u_r \cdot u_{r+1} = v_1 \cdots v_r$$

If $\eta(v_{r+1}, v_{r+1}) < 0$ then replace $\sqrt{\eta(v_{r+1}, v_{r+1})}$ with $\sqrt{-\eta(v_{r+1}, v_{r+1})}$, and ± 1 with ∓ 1 in (2.2.11); the same argument shows that x can be written into a product of vectors of norm ± 1. We thus have that $x \in \mathrm{Pin}(t, s)$, so:

$$\mathrm{Pin}(t, s) = N^{-1}(1) \cup N^{-1}(-1)$$

and is therefore an embedded Lie subgroup of $\mathrm{Cl}^\times(t, s)$ by our earlier discussion. $\square$

Corollary 2.2.6. *The following subset is a Lie subgroup of Pin(t,s):*

$$\begin{aligned}\mathit{Spin}(t,s) =&\mathit{Pin}(t,s)\cap \mathit{Cl}^0(t,s)\\ =&\{v_1\cdots v_{2m} : v_i \in S^{t,s}_{\pm}, m\geq 0\}\end{aligned}$$

We also set:

$$\mathit{Spin}(n) = \mathit{Spin}(0,n)$$

*and call Spin(t,s) the **spin group**.*

Proof. It is clear that Spin(t,s) is a subgroup of Pin(t,s). Note that $\mathrm{Cl}^0(t,s)$ is a closed subset of $\mathrm{Cl}(t,s)$, and that:

$$\mathrm{Pin}(t,s) = \mathrm{Pin}(t,s)\cap \mathrm{Cl}^{\times}(t,s)$$

It follows that:

$$\mathrm{Spin}(t,s) = \mathrm{Pin}(t,s)\cap (\mathrm{Cl}^{\times}(t,s)\cap \mathrm{Cl}^0(t,s))$$

Note that $\mathrm{Cl}^{\times}(t,s)\cap\mathrm{Cl}^0(t,s)$ is closed in $\mathrm{Cl}^{\times}(t,s)$ by definition of the subspace topology. Therefore Spin(t,s) is closed in Pin(t,s), and thus by **Theorem 1.2.1** is an embedded Lie subgroup of Pin(t,s). □

Corollary 2.2.7. *The following subset is a Lie subgroup of Spin(t,s):*

$$\mathit{Spin}^+(t,s) =\{v_1\cdots v_{2m} : m = p+q\geq 0, 2p \text{ of the } v_i\in S^{t,s}_{-} \text{ and } 2q \text{ of the } v_i \in S^{t,s}_{+}\}$$

*We call Spin$^+(t,s)$, the **orthochronus spin group**.*

Proof. It is clear that $\mathrm{Spin}^+(t,s)\subset \mathrm{Spin}(t,s)$ is a subgroup. In particular, we note that if $x\in \mathrm{Spin}^+(t,s)$, then $x\in\mathrm{Spin}(t,s)\cap N^{-1}(1)$, as if:

$$x = v_1\cdots v_{2m}$$

and $2p$ of the v_i satisfy $\eta(v_i,v_i) = -1$, then:

$$N(x) = \eta(v_1,v_1)\cdots\eta(v_{2m},v_{2m}) = (-1)^{2p} = 1$$

While if $x\in\mathrm{Spin}(t,s)\cap N^{-1}(1)$ then we have:

$$N(x) = \eta(v_1,v_1)\cdots\eta(v_{2m},v_{2m}) = (-1)^{k} = 1$$

where k is the number of v_i such that $\eta(v_i,v_i) = -1$. It follows that k is even, hence $x\in\mathrm{Spin}^+(t,s)$. This implies that:

$$\mathrm{Spin}^+(t,s) = \mathrm{Spin}(t,s)\cap N^{-1}(1)$$

We see that $N^{-1}(1)\subset\mathrm{Pin}(t,s)$ is closed in Pin(t,s), so Spin$^+(t,s)$ is closed in Spin(t,s), and is thus an embedded Lie subgroup of Spin(t,s) by **Theorem 1.2.1**. □

We want to realize these groups as double covers of the pseudo orthogonal groups. We have not touched on covering spaces in this paper, and we do not intend to, so for our purposes it is perfectly fine to think of the double cover of a Lie group G as an other Lie group H such that there exists a surjective Lie group homomorphism ϕ satisfying $\ker\phi = \mathbb{Z}_2 = \{\pm e\}$. Note that this implies that for any $g\in G$, there exist precisely two elements $h,-h\in H$ such that $\phi(h) = \phi(-h) = g$. Incidentally, we have already constructed this homomorphism. Indeed, let $v,u\in\mathbb{R}^{t,s}$, with $\eta(v,v) = \pm1$, then $v^{-1} = \mp v$, so:

$$\begin{aligned}\mathrm{Ad}^{\times}_v(u) =&\pm v\cdot u\cdot v\\ =&\pm v(-v\cdot u - 2\eta(v,u))\\ =&u\mp 2\eta(v,u)v\end{aligned}$$

If v is perpendicular to u then:

$$\mathrm{Ad}^{\times}_v(u) = u$$

while if v is parallel to u, then we have that for some $a \in \mathbb{R}$, $u = av$, so:

$$\begin{aligned}\mathrm{Ad}^{\times}_v(u) =& av \mp 2a\eta(v,v)v \\ =& av - 2av \\ =& -u\end{aligned}$$

This implies that $\mathrm{Ad}^{\times}_v$ is a reflection in the hyperplane $v^{\perp}$, so $\mathrm{Ad}^{\times}_v \in O(t,s)$. The composition of any amount of reflections is still an element of $O(t,s)$, so it follows that $\mathrm{Ad}^{\times}$ restricted to $\mathrm{Pin}(t,s)$ is a Lie group homomorphism with image in $O(t,s)$, as:

$$\mathrm{Ad}^{\times}_{v_1\cdots v_r} = \mathrm{Ad}^{\times}_{v_1} \circ \cdots \circ \mathrm{Ad}^{\times}_{v_r}$$

To demonstrate that this map is a surjection with $\ker \mathrm{Ad}^{\times} = \mathbb{Z}_2$, we first cite the following theorem by Cartan. The proof can be found in Hamilton's *Mathematical Gauge Theory.*

Theorem 2.2.7. *Every element of $O(t,s)$ can be written as a composition of at most $2(t+s)$ reflections in hyperplanes $v_i^{\perp}$ with vectors $v_i \in S^{t,s}_{\pm}$.*

We also need the following result:

Theorem 2.2.8. *Let R in $O(t,s)$ be a composition of reflections, in hyperplanes $v_i^{\perp}$ with vectors $v_i \in S^{t,s}_{\pm}$. Then:*

a) R is an element of $SO(t,s)$ if an only if the number of reflections is even.

b) R is an element of $SO^{+}(t,s)$ if and only both the number of vectors $v_i \in S^{t,s}_{-}$ and the number of vectors $v_i \in S^{t,s}_{+}$ are even.

However, we require this lemma:

Lemma 2.2.17. *Let $A, B \in O(t,s)$, if A and B both have time orientability -1 then $AB = A \circ B$ has time orientability $+1$. Furthermore, if A has time orientability -1 and B has time orientability $+1$ then $AB = A \circ B$ has time orientability -1.*

Proof. Let V be an arbitrary maximally negative definite subspace of $\mathbb{R}^{t,s}$. If A and B both have time orientability -1, we see that the composition:

$$(A \circ B)|_V = A|_{B(V)} \circ B|_V$$

is an isomorphism $V \to A \circ B(V)$. Since it is a composition of orientation reversing isomorphism's, the composition $(A \circ B)|_V$ is orientation preserving, hence $AB = A \circ B$ has time orientability $+1$ by **Lemma 1.2.2**. If B has time orientability $+1$, then this is an orientation reversing isomorphism, so AB has time orientability -1. □

We can now prove **Theorem 2.2.8**:

Proof. Let $R = R_{v_1} \circ \cdots \circ R_{v_r}$, where R_{v_r} denotes the reflection in the hyper plane $v_i^{\perp}$. For every vector $v_i \in S^{t,s}_{-}$, we can decompose $\mathbb{R}^{t,s}$ into maximally negative definite and maximally positive definite subspaces $W_- \oplus W_+$ such that for a suitable change of basis matrix $A_i : W_- \oplus W_+ \to \mathbb{R}^{t,s}$:

$$A_i^{-1} R_{v_i} A_i = \begin{pmatrix} -1 & 0 & 0 \\ 0 & I_{t-1} & 0 \\ 0 & 0 & I_s \end{pmatrix} \tag{2.2.12}$$

where the 0's represent row and column vectors of length $t-1$ and s. Similarly, if $v_i \in S^{t,s}_{+}$, we can write:

$$A_i^{-1} R_{v_i} A_i = \begin{pmatrix} I_t & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & I_{s-1} \end{pmatrix} \tag{2.2.13}$$

The determinant of both of these are equal to -1, implying that the determinant of any reflection is -1. Suppose $R \in SO(t,s)$, then:

$$\det(R) = 1$$

However this implies that:

$$\det(R_{v_1}) \cdots \det(R_{v_r}) = (-1)^r = 1$$

implying that r is even, hence then number of reflections is even. If the number of reflections is even then:

$$\det(R_{v_1} \cdots R_{v_r}) = (-1)^r = 1$$

so $R \in SO(t,s)$, implying a).

For brevity denote $A_i^{-1} \circ R_{v_i} \circ A_i$ by R'_{v_i}. We see that the restriction $R'_{v_i}|_{W_-} : W_- \to W_-$ is given by the matrix:

$$R'_{v_i}|_{W_-} = \begin{pmatrix} -1 & 0 \\ 0 & I_{t-1} \end{pmatrix}$$

as by (2.2.12), R'_{v_i} maps W_- to W_-, so we can think of this restriction as a $t \times t$ matrix. Similarly if $v_i \in S^{t,s}_+$ we have that:

$$R'_{v_i} = I_s$$

It follows by **Lemma 1.2.2** that if $v_i \in S^{t,s}_\pm$ then R_{v_i} has time orientability ± 1. Let R be a composition of reflections, it then follows by **Lemma 2.2.17**, that if p is the number of reflections in R which correspond to vectors in $S^{t,s}_-$, then we can rewrite R as a composition of p time orientation reversing transformations in $SO(t,s)$. Suppose R in $SO^+(t,s)$, and that $p = 2m+1$ is odd, then

$$\begin{aligned} R =& A_1 \circ \cdots \circ A_{2m+1} \\ =& (A_1 \circ A_2) \circ \cdots \circ (A_{2m-1} \circ A_{2m}) \circ A_{2m+1} \end{aligned}$$

where each $A_i \in SO(t,s)$ has time orientability -1. Each pair of compositions has time orientability $+1$, so this becomes:

$$R = B \circ A_{2m+1}$$

where $B \in SO^+(t,s)$. By **Lemma 2.2.17**, it follows that R has time orientability -1, a contradiction, thus p must be even. Since $R \in SO^+(t,s)$ it follows that $R \in SO(t,s)$, so there must be an even amount of reflections in R corresponding to vectors in $S^{t,s}_+$. Conversely, suppose that R is a composition of reflections, and that $p = 2m$ and $q = 2n$ are the number of reflections corresponding to vectors in $S^{t,s}_-$ and $S^{t,s}_+$ respectively. It follows that $R \in SO(t,s)$, and furthermore, that:

$$R = A_1 \circ \cdots \circ A_{2m}$$

where each $A_i \in SO(t,s)$ has time orientability -1. We can rewrite this as:

$$\begin{aligned} R =& (A_1 \circ A_2) \circ \cdots (A_{2m-1} \circ A_{2m}) \\ =& B_1 \circ \cdots \circ B_m \end{aligned}$$

where $B_i \in SO^+(t,s)$, hence $R \in SO^+(t,s)$ which implies b). □

Corollary 2.2.8. *Denote the restriction of $Ad^\times$ to $Pin(t,s)$, by λ. The following then hold:*

a) The Lie group homomorphism:

$$\lambda : Pin(t,s) \longrightarrow O(t,s)$$

is surjective and has kernel $\mathbb{Z}_2 = \{\pm 1\}$

b) *The restriction of λ to Spin(t,s) and Spin$^+(t,s)$ defines surjective Lie groups homomorphisms:*

$$\lambda : Spin(t,s) \longrightarrow SO(t,s)$$
$$\lambda : Spin^+(t,s) \longrightarrow SO^+(t,s)$$

with kernel $\mathbb{Z}_2 = \{\pm 1\}$

Proof. It is clear that the map:

$$\lambda : \mathrm{Pin}(t,s) \longrightarrow O(t,s)$$

is surjective, by **Theorem 2.2.7**. We want to show that the kernel of $\lambda = \pm 1$. Note that $\mathrm{Pin}(t,s) \subset \mathrm{Cl}^*(t,s)$, and thus if:

$$\mathrm{Ad}^{\times}_{v_1 \cdots v_r} = \mathrm{Id}$$

by **Proposition 2.2.5**, we have that $v_1 \cdots v_r \in \mathbb{R}^*$. However, the only scalars in $\mathrm{Pin}(t,s)$ are ± 1 by construction, hence $\ker \lambda = \mathbb{Z}_2 = \{\pm 1\}$, implying *a*).

Note that λ restricted to $\mathrm{Spin}(t,s)$ and $\mathrm{Spin}^+(t,s)$ has image in $SO(t,s)$ and $SO^+(t,s)$, and is surjective by **Theorem 2.2.8**. Furthermore, the kernel of λ restricted to $\mathrm{Spin}(t,s)$ and $\mathrm{Spin}^+(t,s)$ is clearly $\{\pm 1\}$, by the same argument for $\mathrm{Pin}(t,s)$, implying *b*). □

A simple fact from algebra demonstrates that for any $g \in O(t,s)$, we have that:

$$\lambda^{-1}(g) = \{\pm x\} = x \cdot \mathbb{Z}_2$$

where x is any arbitrary element of $\lambda^{-1}(g)$. This shows that for each element $g \in O(t,s)$, there exist precisely two elements in $\mathrm{Pin}(t,s)$ which map to g. It can also be shown that $\mathrm{Spin}^+(t,s)$ is connected.

The following lemma will allow us to prove a variety of important properties the groups spin groups:

Lemma 2.2.18. *Let $f : G \to H$ be a surjective Lie group homomorphism such that $\ker f$ is a discrete subgroup of G. Then the induced Lie algebra homomorphism $f_* : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism.*

Proof. Note that the kernel of f is closed in the G, and thus by **Theorem 1.2.1** is an embedded Lie subgroup of G. Since $\ker f$ is discrete it follows that this Lie subgroup is 0 dimensional. With this in mind, we prove that f_* is injective as follows; let $X \in \mathfrak{g}$ and suppose that $f_*(X) = 0$. We then see by **Proposition 1.2.9** that:

$$f(\exp(tX)) = \exp(tf_*(X)) = e_H$$

However, this implies that the one dimensional embedded subgroup generated by X under the exponential map lies in the kernel of f, a contradiction. Therefore, f_* has trivial kernel, and is thus injective.

We now show that f_* is surjective. Let $Y \in \mathfrak{h}$, $t \neq 0$, then $\exp(tY) \in H$, and since f is surjective, it follows that there exists a $g \in G$ such that:

$$f(g) = \exp(tY)$$

Let t be small enough such that tY lies in a small enough open neighborhood U of 0 such that the exponential map is a local diffeomorphism. It follows that $\exp(U)$ is an open neighborhood of e_H, and that $f^{-1}(\exp(U))$ is an open neighborhood of e_G. If necessary, make t smaller, so that $f^{-1}(\exp(U))$ is locally diffeomorphic to an open neighborhood of 0 in $\mathfrak{g}$. It then follows that for some $\epsilon > 0$ there exists an $s \in (-\epsilon, \epsilon)$, and $X \in \mathfrak{g}$ such that:

$$\exp(sX) = g$$

hence:

$$f(\exp(sX)) = \exp(sf_*X) = \exp(tY)$$

Since the exponential map is a local diffeomorphism in both neighborhoods we have that:

$$sf_*X = tY$$

Therefore:

$$f_*\left(\frac{s}{t}X\right) = Y$$

so f_* is surjective, implying the claim. $\square$

We then immediately obtain the following result regarding the Lie algebras of the spin groups.

Corollary 2.2.9. *The following Lie algebra homomorphisms are isomorphisms:*

$$\lambda_* : \mathfrak{pin}(t,s) \longrightarrow \mathfrak{o}(t,s)$$
$$\lambda_* : \mathfrak{spin}(t,s) \longrightarrow \mathfrak{so}(t,s)$$
$$\lambda_* : \mathfrak{spin}^+(t,s) \longrightarrow \mathfrak{so}^+(t,s)$$

In particular,

$$\mathfrak{spin}^+(t,s) \cong \mathfrak{so}(t,s)$$

and the dimension of each Pin and Spin group is:

$$\frac{n^2-n}{2}$$

where $n = s + t$.

Though the preceding isomorphism is certainly convenient, it is at times important to have a description of $\mathfrak{spin}^+(t,s)$ as a subset of the Lie algebra $\mathfrak{cl}^\times(t,s) \cong \mathrm{Cl}(t,s)$. We need the following definition:

Definition 2.2.18. Let $\{e_i\}$ be an orthonormal basis for $\mathbb{R}^{t,s}$. We denote by $M(t,s)$ the subspace of $\mathrm{Cl}(t,s)$ given by the span of the set:

$$\{e_ie_j \in \mathrm{Cl}(t,s) : 1 \leq i < j \leq n\}$$

Lemma 2.2.19. *The vector space $M(t,s)$ is a Lie subalgebra of $\mathfrak{cl}^\times(t,s)$ of dimension:*

$$\dim_{\mathbb{R}} M(t,s) = \frac{n^2-n}{2}$$

where $n = t + s$.

Proof. It follows that are n choose 2 basis vectors of $M(t,s)$, as the set:

$$\{e_ie_j \in \mathrm{Cl}(t,s) : 1 \leq i < j \leq n\}$$

spans the space by construction, and is linearly independent by **Corollary 2.2.3**.

We need to check that that $M(t,s)$ is closed under the bracket operation on $\mathfrak{cl}^\times(t,s)$ given by:

$$[x,y] = x \cdot y - y \cdot x$$

for all $x, y \in \mathrm{Cl}(t,s)$. Let $x = A^{ij}e_ie_j$, and $y = B^{lk}e_le_k$, then:

$$[x,y] = A^{ij}B^{kl}[e_ie_j, e_le_k]$$

We see that:

$$[e_ie_j, e_le_k] = e_ie_je_le_k - e_le_ke_ie_j$$

Examine the second term:

$$\begin{aligned} e_le_ke_ie_j &= -e_l(e_ie_k + 2\eta_{ik})e_j \\ &= -e_le_ie_ke_j - 2\eta_{ik}e_le_j \end{aligned}$$

Repeating this process we find that:

$$e_l e_k e_i e_j = e_i e_j e_l e_k + 2\eta_{lj} e_i e_k - 2\eta_{jk} e_i e_l + 2\eta_{il} e_k e_j - 2\eta_{ik} e_l e_j$$

hence:

$$[e_i e_j, e_l e_k] = -2\eta_{lj} e_i e_k + 2\eta_{jk} e_i e_l - 2\eta_{il} e_k e_j + 2\eta_{ik} e_l e_j$$

When $i = l$ and $j = k$ this is expression is zero as expected. Furthermore, if $i \neq l, k$ and $j \neq l, k$, then this expression is also zero. If $i = l$, and $j \neq k$, then $l = i < k \neq j$:

$$[e_i e_j, e_l e_k] = -\, 2\eta_{ii} e_k e_j$$

and if $k > j$, then we can reorder using the Clifford relation. Similarly if $i = k$ then $l < k = i < j$, so :

$$[e_i e_j, e_l e_k] = 2\eta_{ii} e_l e_j$$

If $j = l$ then $i < j = l < k$ so:

$$[e_i e_j, e_l e_k] = -2\eta_{jj} e_i e_k$$

Finally, if $j = k$, and $i \neq l$ then $l \neq i < j = k$ so:

$$[e_i e_j, e_l e_k] = 2\eta_{jj} e_i e_l$$

and if $i > l$ then we can reorder as before. It follows that $M(t, s)$ is closed under the Lie bracket, and thus a Lie subalgebra of $\mathfrak{cl}^{\times}(t, s)$. $\square$

We now show that $M(t, s)$ is actually to $\mathfrak{spin}^{+}(t, s)$:

Proposition 2.2.7. *For all $s, t \geq 0$, the following identity holds:*

$$\mathfrak{spin}(t, s) = M(t, s)$$

Proof. If we can show that $M(t, s)$ is a subspace of $\mathfrak{spin}^{+}(t, s)$ then since $\dim_{\mathbb{R}} M(t, s) \cong \dim_{\mathbb{R}} \mathfrak{spin}^{+}(t, s)$ we are done.

Suppose that $e_i e_j$ satisfy $\eta_{ii} = \eta_{jj}$. Then the curve:

$$\gamma(t) = e_i(-\eta_{ii} \cos(t) e_i + \sin(t) e_j)$$

is a smooth curve in $\mathrm{Spin}^{+}(t, s)$ as:

$$\begin{aligned} \eta(-\eta_{ii} \cos(t) e_i + \sin(t) e_j, -\eta_{ii} \cos(t) e_i + \sin(t) e_j) =& \eta_{ii}^3 \cos^2(t) + \sin^2(t) \eta_{jj} \\ =& \eta_{ii} \cos^2(t) + \sin^2(t) \eta_{jj} \\ =& \eta_{ii} \end{aligned}$$

Similarly, if $\eta_{ii} \neq \eta_{jj}$, then $\eta_{ii} = -\eta_{jj}$, so the curve

$$\gamma(t) = e_i(-\eta_{ii} \cosh(t) e_i + \sinh(t) e_j)$$

lies in $\mathrm{Spin}^{+}(t, s)$ as:

$$\begin{aligned} \eta(-\eta_{ii} \cosh(t) e_i + \sinh(t) e_j, -\eta_{ii} \cosh(t) e_i + \sinh(t) e_j) =& \eta_{ii} \cosh^2(t) + \sinh^2(t) \eta_{jj} \\ =& \eta_{ii} \cosh^2(t) - \sinh^2(t) \eta_{ii} \\ =& \eta_{ii} \end{aligned}$$

At $t = 0$, it s clear that both curves go through the identity, as $\sinh(0) = \sin(0) = 0$, and $-\eta_{ii} e_i = e_i^{-1}$. Taking the derivative at $t = 0$, we find that in both cases:

$$\dot{\gamma}(0) = e_i e_j \in \mathfrak{spin}^{+}(t, s)$$

for all $1 \leq i < j \leq s + t$. Therefore, $M(t, s) \subset \mathfrak{spin}^{+}(t, s)$, implying the claim. $\square$

In particular, the above proposition implies that the spanning set of $M(t,s)$ given in **Definition 2.2.18** is a basis for $\mathfrak{spin}^+(t,s)$. We wish to find the the image of this set under the isomorphism λ_*

Proposition 2.2.8. *Let $\{e_i\}$ be an orthonormal basis for $\mathbb{R}^{t,s}$, then for all $1 \le i < j \le s+t$, the Lie algebra isomorphism:*

$$\lambda_* : \mathfrak{spin}^+(t,s) \longrightarrow \mathfrak{so}^+(t,s)$$

satisfies:

$$\lambda_*(e_i e_j) = 2(\eta_{ii} E^i_j - \eta_{jj} E^j_i)$$

where there is no summation over i and j, and E^i_j is the $(s+t) \times (s+t)$ matrix with a 1 in the ith column of the jth row, and zeroes elsewhere. In other words:

$$E^i_j = e^i \otimes e_j$$

where $\{e^{\}}$ is the basis dual to $\{e_i\}$.

Proof. Note that the map in the case of $\mathrm{Spin}^+(t,s)$ we have that:

$$\lambda(x)(y) = \mathrm{Ad}^\times_x(y) = \alpha(x)yx^{-1} = xyx^{-1} = \mathrm{Ad}_x(y)$$

for all $x \in \mathrm{Spin}^+(t,s)$ and all $y \in \mathfrak{spin}^+(t,s)$. It follows from **Theorem 1.2.7**, that:

$$\lambda_* : \mathfrak{spin}^+(t,s) \longrightarrow \mathfrak{so}^+(t,s)$$

satisfies:

$$\lambda_*(x)(y) = [x,y]$$

for all $x \in \mathfrak{spin}^+(t,s)$ and $y \in \mathrm{Cl}(t,s)$. Suppose that $y \in \mathbb{R}^{t,s} \subset \mathrm{Cl}(t,s)$, and let $y = y^m e_m$, then since $\eta_{im} = 0$ unless $m = i$ we have:

$$\begin{aligned}\lambda_*(e_i e_j)(e_m) =& [e_i e_j, y^m e_m] \\ =& y^m e_i e_j e_m - y^m e_m e_i e_j \\ =& 2(-y^m \eta_{jm} e_i + y^m \eta_{im} e_j) \\ =& 2(-y^j \eta_{jj} e_i + y^i \eta_{ii} e_j)\end{aligned}$$

where there is no implied summation over j and i. Let $\{e^i\}$ denote the dual basis to $\{e_i\}$, then the matrix:

$$\eta_{ii} E^i_j - \eta_{jj} E^j_i$$

can be written as:

$$\eta_{ii} e^i \otimes e_j - \eta_{jj} e^j \otimes e_i$$

where again there is no implied summation over i and j. It follows that:

$$\begin{aligned}2(\eta_{ii} e^i \otimes e_j - \eta_{jj} e^j \otimes e_i)(y) =& 2(y^m \eta_{ii} e^i(e_m) \otimes e_j - \eta_{jj} e^j(e_m) \otimes e_i) \\ =& 2(y^i \eta_{ii} e_j - y^j \eta_{jj} e_i)\end{aligned}$$

Hence:

$$\lambda_*(e_i e_j) = 2(\eta_{ii} E^i_j - \eta_{jj} E^j_i)$$

$\square$

It should be clear from **Example 1.2.11** that $\lambda_*(e_i e_j)$ does indeed lie in $\mathfrak{so}^+(t,s)$. Furthermore, for brevity, we sometimes write:

$$\epsilon_{ij} = \eta_{ii} E^i_j - \eta_{jj} E^j_i$$

Now that we have sufficiently characterized the Lie algebra $\mathfrak{spin}^+(t,s)$, we turn to our second application of **Lemma 2.2.18**. Specifically, can obtain the following result regarding the quotients of the spin groups by the kernel of λ.

Proposition 2.2.9. *There exist unique Lie group isomorphisms:*

$$\begin{aligned} Pin(t,s)/\mathbb{Z}_2 &\cong O(t,s) \\ Spin(t,s)/\mathbb{Z}_2 &\cong SO(t,s) \\ Spin^+(t,s)/\mathbb{Z}_2 &\cong SO^+(t,s) \end{aligned}$$

Proof. Let G denote any of the groups $\mathrm{Pin}(t,s)$, $\mathrm{Spin}(t,s)$, or $\mathrm{Spin}^+(t,s)$, and H denote the corresponding image.

Note that $\ker \lambda = \mathbb{Z}_2$, so $\mathbb{Z}_2$ is a closed, normal subgroup of G. In particular, $\mathbb{Z}_2$ is a compact embedded Lie subgroup of each group, thus the natural right action of $\mathbb{Z}_2$ on each group is proper. This action is easily seen to be free, as for any $g \in G$, and any $h_1, h_2 \in \mathbb{Z}_2$ we have that:

$$\phi_g(h_1) = g \cdot h_1 = g \cdot h_2 = \phi_g(h_2)$$

implies that:

$$h_1 = h_2$$

as we can apply g^{-1} on the left to both sides. **Corollary 1.2.8** and **Theorem 1.2.4** then tell us that $G/\mathbb{Z}_2$, has the structure of a smooth manifold such that projection π is a smooth submersion.

However, this smooth manifold also has a well defined group structure as $\mathbb{Z}_2$ is a normal subgroup. The group structure is given by:

$$[g] \cdot [h] = [g \cdot h] = \pi(g \cdot h)$$

Let $s_1, s_2 : U \to G$ be two smooth local sections of π, then we have that $s_1 \times s_2 : U \times U \to G \times G$. The multiplication map on $G/\mathbb{Z}_2$ is then given locally by:

$$\mu = \pi(s_1 \cdot s_2)$$

which is a composition of smooth maps and thus smooth. **Lemma 1.2.3** then implies that $G/\mathbb{Z}_2$ is a Lie group. By the universal property of quotient groups,[27] there thus exists a unique group homomorphism:

$$\psi : G/\mathbb{Z}_2 \longrightarrow H$$

such that:

$$\psi \circ \pi = \lambda \tag{2.2.14}$$

By **Lemma 1.2.10**, we have that ψ is a smooth surjective, and thus a surjective Lie group homomorphism. It is also injective since $\ker \pi = \ker \lambda$. The induced Lie algebra homomorphism is thus an isomorphism by **Lemma 2.2.18**, so at all points $g \in G$ the differential:

$$D_g \psi = D_e L_{\psi(g)} \circ D_e \psi$$

is an isomorphism as it is a composition of isomorphisms. It follows that ψ is a Lie group homomorphism which is also a diffeomorphism, and thus a Lie group isomorphism, implying the claim. ∎

[27] Quotient groups have a similar universal property as algebras.

Corollary 2.2.10. *The Lie group homomorphisms:*

$$\begin{aligned}\lambda : Pin(t,s) &\longrightarrow O(t,s)\\ \lambda : Spin(t,s) &\longrightarrow SO(t,s)\\ \lambda : Spin^+(t,s) &\longrightarrow SO^+(t,s)\end{aligned}$$

are open surjective submersions.

Proof. This follows from (2.2.14), as the composition of open submersions is an open submersion. □

With this corollary, we can determine when $\mathrm{Spin}^+(t,s)$ is a connected Lie group.

Proposition 2.2.10. *If $t \geq 2$ or $s \geq 2$, then $Spin^+(t,s)$ is connected.*

Proof. Recall that if $f : X \to Y$ is a continuous open surjective map between topological spaces, and Y is connected, then X is connected if any two points $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$ lie in a connected subset of X. Since every connected manifold is path connected, it suffices to show that for any $x \in \mathrm{Spin}^+(t,s)$ we can find a path connecting x to $-x$.

Suppose that $t \geq 2$, and let $x \in \mathrm{Spin}^+(t,s)$, and consider the curve:

$$\gamma(\theta) = x(e_1\cos(\theta) + e_2\sin(\theta))(e_1\cos(\theta) - e_2\sin(\theta))$$

We see that this is indeed in $\mathrm{Spin}^+(t,s)$ as:

$$\begin{aligned}\eta(e_1\cos(\theta) + e_2\sin(\theta), e_1\cos(\theta) + e_2\sin(\theta)) =& -\cos^2(t) - \sin^2(t)\\ =& -1\end{aligned}$$

and similarly for the second term. This curve is smooth, and connects x to $-x$ as:

$$\begin{aligned}\gamma(0) =& x\\ \gamma(\pi/2) =& -x\end{aligned}$$

Meanwhile, if $s \geq 2$, the curve:

$$\gamma(\theta) = x(e_{t+1}\cos(\theta) + e_{t+2}\sin(\theta))(e_{t+1}\cos(\theta) - e_{t+2}\sin(\theta))$$

lies in $\mathrm{Spin}^+(t,s)$ by a similar argument, and satisfies:

$$\begin{aligned}\gamma(0) =& -x\\ \gamma(\pi/2) =& x\end{aligned}$$

so $\mathrm{Spin}^+(t,s)$ is connected. □

Before ending with an explicit description of $\mathrm{Spin}^+(1,3)$, we wish to develop the spinor representation of $\mathrm{Spin}^+(t,s)$. Note that $\mathrm{Spin}^+(t,s) \subset \mathrm{Cl}^0(t,s) \subset \mathbb{C}\mathrm{l}^0(t,s)$, and thus restriction of the representation discussed in **Corollary 2.2.5** and **Proposition 2.2.3** to $\mathrm{Spin}^+(t,s)$ yields an induced representation of $\mathrm{Spin}^+(t,s)$ on Δ_{s+t}.

Definition 2.2.19. Let $s + t = n$, we denote by:

$$\kappa : \mathrm{Spin}^+(t,s) \longrightarrow GL(\Delta_n)$$

the **spinor representation** induced by the restriction of the even Clifford algebra $\mathrm{Cl}^0(t,s)$

Proposition 2.2.11. *The spinor representation of $Spin^+(t,s)$ is compatible with Clifford multiplication in the following way:*

$$\kappa(g)(x \cdot \psi) = (\lambda(g) \cdot x)(\kappa(g) \cdot \psi)$$

for all $g \in Spin^+(t,s)$, $x \in \mathbb{R}^{t,s}$ and $\psi \in \Delta_n$. Here $\lambda(g) \cdot x$ is shorthand for $\rho_{SO^+}(\lambda(g)) \cdot x$, where ρ_{SO^+} is the standard representation of $SO^+(t,s)$ on $\mathbb{R}^{t,s}$.

Proof. Let ρ denote the representation:

$$\rho : \mathrm{Cl}(t,s) \longrightarrow \Delta_n$$

where $n = s + t$. Then:

$$\begin{aligned}\kappa(g)(x \cdot \psi) =& \rho(g)\rho(x) \cdot \psi \\ =& \rho(gx)\rho(g^{-1}g) \cdot \psi \\ =& \rho(gxg^{-1})\rho(g) \cdot \psi \\ =& (\lambda(g) \cdot x)(\kappa(g) \cdot \psi)\end{aligned}$$

as desired. □

As mentioned earlier, we now end with an explicit description of $\mathrm{Spin}^+(1,3)$. Specifically, as this next example shows, $\mathrm{Spin}^+(1,3)$ can be identified with the Lie group $SL_2(\mathbb{C})$.

Example 2.2.6. Let $SL_2(\mathbb{C})$ denote the subset:

$$SL_2(\mathbb{C}) = \{A \in GL_2(\mathbb{C}) : \det(A) = 1\}$$

Via similar methods in **Example 1.2.2**, one can show that this is a Lie group of real dimension $2 \cdot 4 - 2 = 6$, as the determinant maps to $\mathbb{C}^2$ which has real dimension 2, and $\mathrm{Mat}_{2\times 2}(\mathbb{C})$ has real dimension $2 \cdot 4 = 8$. We wish to show that $SL_2(\mathbb{C})$ is the double cover of $SO^+(1,3)$ and thus satisfies[28]:

$$SL_2(\mathbb{C}) \cong \mathrm{Spin}^+(1,3)$$

We write a vector $v \in \mathbb{R}^{1,3}$ as:

$$(t,x,y,z) \longmapsto X = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix}$$

Note that this matrix lies in the four dimensional real vector space $H_2(\mathbb{C})$ of Hermitian 2×2 matrices[29], which, as a real vector spaces, is spanned by the Pauli spin matrices, and the identity. Indeed, this assignment $f : \mathbb{R}^{t,s} \to H_2(\mathbb{C})$ is equivalent to:

$$(x^0, x^1, x^2, x^3) \longrightarrow x^\mu \sigma_\mu = X$$

where $\sigma_0 = I_2$. This assignment is clearly injective, and linear, so it is an isomorphism of real vector spaces by rank-nullity. We also note that:

$$-\det(X) = -(t^2 - z^2 - x^2 - y^2) = -t^2 + x^2 + y^2 + z^2$$

so $-\det(X) = \eta(x,x)$, where η is the standard Minkowski inner product of signature $(- + ++)$.

Examine the following map:

$$\begin{aligned}F : SL_2(\mathbb{C}) \times \mathbb{R}^{1,3} &\longrightarrow \mathbb{R}^{1,3} \\ (M, X) &\longmapsto MXM^\dagger\end{aligned}$$

where $\dagger$ denotes the conjugate transpose. Let:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

for $a, b, c, d \in \mathbb{C}$. We want to show that F truly has image in $\mathbb{R}^{1,3}$. Note that:

$$MXM^\dagger = Mx^\mu \sigma_\mu M^\dagger = x^\mu M \sigma_\mu M^\dagger$$

[28] In particular this follows from the fact that $\mathrm{Spin}^+(1,3)$ is the *universal cover* of $SO^+(1,3)$, and that $SL_2(\mathbb{C})$ is simply connected, but working out these details would take us too far afield. We do note however, that $\mathrm{Spin}^+(t,s)$ is not the universal cover for most combinations of t, and s.

[29] Those which satisfy $H = H^\dagger$

so it suffices to check that each Pauli matrix gets mapped to a Hermitian matrix. We do this as follows

$$M\sigma_0 M^\dagger = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} |a|^2+|b|^2 & a\bar{c}+b\bar{d} \\ \bar{a}c+\bar{b}d & |c|^2+|d|^2 \end{pmatrix} \in H_2(\mathbb{C})$$
$$M\sigma_1 M^\dagger = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} \bar{a}b+a\bar{b} & b\bar{c}a\bar{d} \\ d\bar{a}+c\bar{b} & d\bar{c}+\bar{d}c \end{pmatrix} \in H_2(\mathbb{C})$$
$$M\sigma_2 M^\dagger = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} i(\bar{a}b-a\bar{b}) & i(b\bar{c}-a\bar{d}) \\ i(d\bar{a}-c\bar{b}) & i(d\bar{c}-\bar{d}c) \end{pmatrix} \in H_2(\mathbb{C})$$
$$M\sigma_3 M^\dagger = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} |a|^2-|b|^2 & a\bar{c}-b\bar{d} \\ \bar{a}c-\bar{b}d & |c|^2-|d|^2 \end{pmatrix} \in H_2(\mathbb{C})$$

It is easily seen that each matrix lies in $H_2(\mathbb{C})$ by taking the conjugate transpose, hence F is a well defined map. Furthermore, we see that if $M \in SL_2(\mathbb{C})$ then:

$$-\det\bigl(MXM^\dagger\bigr) = -\det(M)\det(X)\det\bigl(M^\dagger\bigr) = -\det(X)$$

so F preserves $-\det(X) = \eta(x,x)$. It then follows by a similar argument to **Theorem 1.2.6**, that F induces a Lie group homomorphism:

$$\begin{aligned} \phi : SL_2(\mathbb{C}) &\longmapsto O(1,3) \\ M &\longmapsto \rho(M) = F(M,\cdot) \end{aligned}$$

We want to show that the homomorphism actually has image in $SO^+(t,s)$, and to do this, we need an inverse map $H_2(\mathbb{C}) \to \mathbb{R}^{1,3}$. We define an inner product on $H_2(\mathbb{C})$ by:

$$\langle M, N\rangle = \frac{1}{2}\operatorname{Tr}(MN)$$

It is then easily verifiable that the Pauli spin matrices are orthonormal:

$$\langle \sigma_\mu, \sigma_\nu\rangle = \delta_{\mu\nu}$$

hence the map:

$$X \longrightarrow (\langle X, \sigma_0\rangle, \ldots \langle X, \sigma_3\rangle)$$

is easily seen to be the inverse of $\mathbb{R}^{1,3} \to H_2(\mathbb{C})$ as:

$$(x^0, x^1, x^2, x^3) \longmapsto x^\mu \sigma_\mu \longmapsto (x^0, x^1, x^2, x^3)$$

is the identity. We then see that

$$\begin{aligned} f^{-1}(\phi(M)(x^\mu\sigma_\mu)) =& x^\mu f^{-1}(M\sigma_\mu M^\dagger) \\ =& x^\mu(\langle M\sigma_\mu M^\dagger, \sigma_0\rangle, \ldots \langle M\sigma_\mu M^\dagger, \sigma_3\rangle) \end{aligned}$$

Since we prefer to think about column vectors when talking of linear transformation, this implies that the matrix corresponding to $\phi(M)$ is given by:

$$\phi(M) = \begin{pmatrix} \langle M\sigma_0M^\dagger, \sigma_0\rangle & \langle M\sigma_1M^\dagger, \sigma_0\rangle & \langle M\sigma_2M^\dagger, \sigma_0\rangle & \langle M\sigma_3M^\dagger, \sigma_0\rangle \\ \langle M\sigma_0M^\dagger, \sigma_1\rangle & \langle M\sigma_2M^\dagger, \sigma_1\rangle & \langle M\sigma_1M^\dagger, \sigma_0\rangle & \langle M\sigma_3M^\dagger, \sigma_1\rangle \\ \langle M\sigma_0M^\dagger, \sigma_2\rangle & \langle M\sigma_2M^\dagger, \sigma_2\rangle & \langle M\sigma_2M^\dagger, \sigma_2\rangle & \langle M\sigma_3M^\dagger, \sigma_2\rangle \\ \langle M\sigma_0M^\dagger, \sigma_3\rangle & \langle M\sigma_2M^\dagger, \sigma_3\rangle & \langle M\sigma_2M^\dagger, \sigma_3\rangle & \langle M\sigma_3M^\dagger, \sigma_3\rangle \end{pmatrix}$$

We easily see that:

$$\langle M\sigma_0M^\dagger, \sigma_0\rangle = \frac{1}{2}(|a|^2 + |b|^2 + |c^2| + |d|^2)$$

so $\det(\phi(M)_{11}) > 0$, hence $\phi(M) \in O^+(t,s)$. Furthermore, under the assumption that $M \in SL_2(\mathbb{C})$, we obtain that if $a \neq 0$:

$$\det(M) = 1 \Longrightarrow a = \frac{1+bc}{d}$$

A lengthy calculation, if done by hand, then demonstrates that:

$$\det(\phi(M)) = 1$$

We can repeat this process for each entry in M, and obtain the same result, so since we can't have $a = b = c = d = 0$, ϕ has image in $SO^+(1,3)$. One could also argue that since $SL_2(\mathbb{C})$ is simply connected, that it's image must be connected, hence ϕ takes image in $SO^+(1,3)$.

Let $M \in \ker\phi$, then for all $x \in \mathbb{R}^{1,3}$ we have that:

$$MXM^\dagger = X$$

If $x = (t,0,0,0)$, this implies that:

$$MM^\dagger = I_2$$

so $M \in \ker\phi$ implies that M is unitary. In particular, we have that:

$$MX = XM$$

for all $X \in \mathbb{H}_2(\mathbb{C})$. Examining this relationship on the Pauli spin matrices σ_1 and σ_3 gives us:

$$\begin{pmatrix} b & a \\ d & c \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$
$$\begin{pmatrix} a & -b \\ c & -d \end{pmatrix} = \begin{pmatrix} a & b \\ -c & -d \end{pmatrix}$$

The second condition implies that $c = d = 0$, while the first condition implies that $a = d$. Since $\det(M) = 1$, we conclude that $M = \pm I_2$, thus $\ker\phi = \{\pm I_2\} \cong \mathbb{Z}_2$.

It follows from our work in **Lemma 2.2.18** that ϕ_* is an injective Lie algebra homomorphism, so since $\dim_{\mathbb{R}} SL_2(\mathbb{C}) = \dim_{\mathbb{R}} SO^+(1,3)$ we have that ϕ_* is an isomorphism. In particular, ϕ_* is surjective, so by **Proposition 1.2.9**, ϕ is a surjection onto an open neighborhood of the identity of $SO^+(1,3)$. However, $SO^+(1,3)$ is connected, and an open neighborhood of the identity generates the connected component of the identity[30], so ϕ is surjective as it is a group homomorphism. It follows that:

$$\mathrm{Spin}^+(1,3) \cong SL_2(\mathbb{C})$$

In particular, we have that:

$$SL_2(\mathbb{C})/\mathbb{Z}_2 \cong SO^+(1,3)$$

by **Proposition 2.2.7**

2.2.6 The Dirac Form

Definition 2.2.20. Let Δ_n be the complex vector space with the spinor representation of $\mathrm{Cl}(t,s)$, where $n = s+t$. We fix a constant $\delta = \pm 1$, and call a nondegenerate $\mathbb{R}$-bilinear form:

$$\langle\cdot,\cdot\rangle : \Delta_n \times \Delta_n \to \mathbb{C}$$

a **Dirac form** if it satisfies the following conditions:

a) $\langle v\cdot\psi, \phi\rangle = \delta\langle\psi, v\cdot\phi\rangle$ for all $v \in \mathbb{R}^{t,s}$, and all $\psi,\phi \in \Delta_n$.

b) $\langle\psi,\phi\rangle = \overline{\langle\phi,\psi\rangle}$ for all $\psi,\phi \in \Delta_n$.

c) $\langle\psi, c\phi\rangle = c\langle\psi,\phi\rangle = \langle\bar{c}\psi,\phi\rangle$ for all $\psi,\phi \in \Delta_n$ and all $c \in \mathbb{C}$.

[30]This is not difficult to see, just show that the subgroup generated by U is both open and closed.

Note that we do not assume that $\langle\cdot,\cdot\rangle$ is positive definite.

Lemma 2.2.20. *Let $\{\chi_i\}$ be a complex basis for Δ_n, and define the matrix A by:*

$$A_{ij} = \langle\chi_i, \chi_j\rangle$$

If we write the column vectors as $\psi, \phi \in \Delta_n$, as $\phi = \phi^i\chi_i$, and $\psi^i = \chi_i$, then:

$$\langle\psi, \phi\rangle = \psi^\dagger A\phi$$

Furthermore, if γ_a are the mathematical gamma matrices for the representation of $Cl(t, s)$ on Δ_n, then $a)$ and $b)$ are equivalent to:

i) $\gamma_a^\dagger = \delta A\gamma_a A^{-1}$ *for all* $a = 1, \ldots, s+t$.

ii) $A^\dagger = A$.

Proof. We see that by property $c)$ in **Definition 2.2.20**:

$$\begin{aligned}\langle\psi^i\chi_i, \phi^j\chi_j\rangle =& \bar{\psi}^i\phi^j\langle\chi_i, \chi_j\rangle \\ =& \bar{\psi}^i\phi^j A_{ij}\end{aligned} \tag{2.2.15}$$

We see that:

$$A = \begin{pmatrix}\langle\chi_1, \chi_1\rangle & \cdots & \langle\chi_1, \chi_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\chi_n, \chi_1\rangle & \cdots & \langle\chi_n, \chi_n\rangle\end{pmatrix}$$

It follows that sum in (2.2.15) is equivalent to:

$$\langle\psi, \phi\rangle = \psi^\dagger A\phi$$

To prove $i)$ note that:

$$\begin{aligned}\langle e_a \cdot \psi, \phi\rangle =& \langle\gamma_a \cdot \psi, \phi\rangle \\ =& (\gamma_a \cdot \psi)^\dagger A\phi \\ =& \psi^\dagger \cdot \gamma_a^\dagger A\phi\end{aligned}$$

Property $a)$ in **Definition 2.2.20** then implies that for all $\psi, \phi \in \Delta_n$:

$$\psi^\dagger(\gamma_a^\dagger A)\phi = \delta\psi^\dagger(A\gamma_a)\phi$$

Since this holds for all ψ, ϕ, we can choose basis elements for ψ, ϕ and pick out the components of the matrices $(\gamma_a^\dagger A)$ and $\delta(A\gamma_a)$, implying that:

$$\gamma_a^\dagger A = \delta A\gamma_a$$

Since $\langle\cdot,\cdot\rangle$ is nondegenerate, it follows that A is invertible, hence:

$$\gamma_a^\dagger = \delta A\gamma_a A^{-1}$$

Finally, to prove $ii)$ we see that by condition $b)$ of **Definition 2.2.20**:

$$A_{ij} = \bar{A}_{ji}$$

implying that:

$$A = A^\dagger$$

as the ijth component of $A^\dagger$ is precisely $\bar{A}_{ji}$. $\square$

Lemma 2.2.21. *Every Dirac form is invariant under the action of $Spin^+(t, s)$.*

Proof. Let $\langle \cdot, \cdot \rangle$ be a Dirac from. Then condition $a)$ of **Definition 2.2.20** implies that:

$$\langle v \cdot \psi, v \cdot \psi \rangle = \delta \langle \psi, (v \cdot v) \psi \rangle$$

It follows that for an element $v_1 \cdots v_{2r} \in \mathrm{Spin}^+(t,s)$ that:

$$\begin{aligned}\langle v_1 \cdots v_{2r} \cdot \psi, v_1 \cdots v_{2r} \cdot \psi \rangle =& \delta^{2r} \langle \psi, v_{2r} \cdots v_1 \cdot v_1 \cdots v_{2r} \cdot \psi \rangle \\ =& \langle \psi, \phi \rangle\end{aligned}$$

as $(v_1 \cdots v_{2r})^t$ is the inverse of $v_1 \cdots v_{2r}$. □

Definition 2.2.21. A complex representation of $\mathrm{Cl}(t,s)$ is called **basis unitary** if all of the gamma matrices are unitary.

Corollary 2.2.11. *Every complex representation of Cl(t,s) admits a basis unitary representation.*

Proof. Let $\mathrm{Cl}(t,s)$ be the spinor representation on Δ_n equipped with the standard Hermitian inner product

$$\langle \psi, \phi \rangle = \psi^\dagger \phi$$

Let $\{e_i\}$ be an orthonormal basis for $\mathbb{R}^{t,s}$ and examine the subgroup of $\mathrm{Cl}^\times(t,s)$ generated by $\{e_1, \ldots, e_{s+t}\}$. It follows that this subgroup is closed and compact, as it has 2^n elements, and is thus a compact Lie subgroup of $\mathrm{Cl}^\times(t,s)$. The claim then follows from an extension of **Theorem 1.2.5**. In particular, the new Hermitian inner product is given by:

$$\langle \psi, \phi \rangle' = \sum_{g \in G} \langle \rho(g)\psi, \rho(g)\phi \rangle$$

where G is the subgroup generated by $\{e_1, \ldots, e_{t+s}\}$. With this new Hermitian inner product, the gamma matrices are clearly basis unitary. □

Proposition 2.2.12. *For any basis unitary representation of Cl(t,s), there exists a Dirac form given by the matrix:*

$$A = \epsilon \Gamma_1 \cdots \Gamma_t$$

where $\delta = (-1)^{t+1}$, and $\epsilon \in \mathbb{C}$ satisfies:

$$\bar{\epsilon} = (-1)^{t(t+1)/2} \epsilon$$

There also exists a Dirac form:

$$A = \epsilon \Gamma_{t+1} \cdots \Gamma_{t+s}$$

where $\delta = (-1)^s$ and $\epsilon \in \mathbb{C}$ satisfies:

$$\bar{\epsilon} = (-1)^{s(s-1)/2} \epsilon$$

Proof. We prove the first statement, as the second follows from the same argument. First note that since $\Gamma_a = -i\gamma_a$ that the physical gamma matrices must satisfy:

$$\begin{aligned}\Gamma_a^\dagger =& (-i\gamma_a)^\dagger \\ =& i\gamma_a^\dagger \\ =& i\delta A \gamma_a A^{-1} \\ =& \delta A \Gamma_a A^{-1}\end{aligned}$$

We also see that:

$$\Gamma_a^{-1} = \eta_{aa} \Gamma_a$$

Furthermore, since γ_a is unitary, it follows that $-i\gamma_a$ is unitary, so Γ_a is unitary. This implies that

$$\Gamma_a^\dagger = \eta_{aa}\Gamma_a$$

If $t+1 \leq a \leq s+t$, then:

$$\begin{aligned}\Gamma_a A =&(-1)^t A\Gamma_a \\ =&-(-t)^{t+1}A\Gamma_a \\ =&-\delta A\gamma_a\end{aligned}$$

Applying A^{-1} to both sides on the right we obtain:

$$\Gamma_a = -\delta A\Gamma_a A^{-1}$$

Since $\eta_{aa} = 1$, we have that $\Gamma_a = \Gamma_a^\dagger$, so:

$$\Gamma_a^\dagger = -\delta A\Gamma_a A^{-1}$$

IF $1 \leq a \leq t$, then:

$$\begin{aligned}A\Gamma_a =&\epsilon\Gamma_a\Gamma_1\cdots\Gamma_t \\ =&(-1)^{1-a}\epsilon\Gamma_1\cdots\Gamma_a\Gamma_a\cdots\Gamma_t \\ =&(-1)^{1-a+t-a}A\Gamma_a \\ =&(-1)^{t+1}A\Gamma_a \\ =&\delta A\Gamma_a\end{aligned}$$

It follows that $\Gamma_a = -\Gamma^\dagger$, so:

$$\begin{aligned}-A\Gamma_a^\dagger =&\delta A\Gamma_a \\ \Rightarrow \Gamma_a^\dagger =&-\delta A\Gamma_a A^{-1}\end{aligned}$$

Furthermore:

$$A = \epsilon\Gamma_1\cdots\Gamma_t$$

while:

$$\begin{aligned}A^\dagger =&(-1)^{t(t+1)/2}\epsilon\Gamma_t^\dagger\cdots\Gamma_1^\dagger \\ =&(-1)^{t(t+1)/2}(-1)^t\epsilon\Gamma_t\cdots\Gamma_1 \\ =&(-1)^{t(t+1)/2}(-1)^t(-1)^{t(t-1)/2}\epsilon\Gamma_1\cdots\Gamma_t \\ =&(-1)^{t^2+t}\epsilon\Gamma_1\cdots\Gamma_t \\ =&A\end{aligned}$$

so A is unitary. It follows that the nondegenerate inner product defined by:

$$\langle\psi,\phi\rangle = \psi^\dagger A\phi$$

defines a Dirac form as desired. □

Given the Dirac form $\langle\cdot,\cdot\rangle$, a spinor ψ, we want to construct ‘dual spinors’ ω which satisfy:

$$\omega(\phi) = \langle\psi,\phi\rangle$$

for all $\phi \in \Delta_n$. This is however easy, as any Dirac form can be viewed as an element of $\bar{\Delta}^* \otimes \Delta$, so we are essentially performing a contraction in the first entry.

Definition 2.2.22. Let $\langle \cdot, \cdot \rangle$ be a Dirac form, and $\psi \in \Delta_n$ a spinor. The **Dirac Conjugate** of ψ, denoted $\bar{\psi}$ is the complex linear map:

$$\begin{aligned}\bar{\psi} : \Delta &\longrightarrow \mathbb{C} \\ \phi &\longmapsto \langle \psi, \phi \rangle\end{aligned}$$

In other words $\bar{\psi}$ is the contraction of the Dirac form in the first entry::

$$\bar{\psi} = \langle \psi, \cdot \rangle \in \Delta^*$$

We see that given a basis $\{\chi_i\}$ for Δ_n, and the matrix A corresponding to the Dirac form, that $\bar{\psi}$ is given by:

$$\bar{\psi} = \psi^\dagger A$$

We also employ the following notation:

$$\langle \psi, \phi \rangle = \bar{\psi}\phi$$

We end with the following example:

Example 2.2.7. Recall from **Example 2.2.5** that the spinor representation of $\mathrm{Cl}(1,3)$ on $\mathbb{C}^4$ was given by the mathematical gamma matrices:

$$\begin{aligned}\gamma_0 &= \begin{pmatrix} 0 & -I_2 \\ -I_2 & 0 \end{pmatrix} \\ \gamma_i &= \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}\end{aligned}$$

The physical Gamma matrices are then given by:

$$\begin{aligned}\Gamma_0 &= \begin{pmatrix} 0 & iI_2 \\ iI_2 & 0 \end{pmatrix} \\ \Gamma_i &= \begin{pmatrix} 0 & i\sigma_i \\ -i\sigma_i & 0 \end{pmatrix}\end{aligned}$$

We see that this representation is basis unitary as:

$$\begin{aligned}\Gamma_0\Gamma_0^\dagger &= \begin{pmatrix} 0 & iI_2 \\ iI_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & -iI_2 \\ -iI_2 & 0 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix} \\ \Gamma_i\Gamma_i^\dagger &= \begin{pmatrix} 0 & i\sigma_i \\ -i\sigma_i & 0 \end{pmatrix}\begin{pmatrix} 0 & i\sigma_i^\dagger \\ -i\sigma_i^\dagger & 0 \end{pmatrix} = \begin{pmatrix} \sigma_i\sigma_i^\dagger & 0 \\ 0 & \sigma_i\sigma_i^\dagger \end{pmatrix}\end{aligned}$$

Each Pauli spin matrix is Hermitian, so $\sigma_i = \sigma_i^\dagger$, and each satisfies $\sigma_i^2 = I_2$, hence $\Gamma_i\Gamma_i^\dagger = I_4$. The physical chirality element is given by:

$$\begin{aligned}\Gamma_5 &= -\, i^3\Gamma_0\Gamma_1\Gamma_2\Gamma_3 \\ &= \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix}\end{aligned}$$

implying that first two components of any spinor are right handed Weyl spinors, and the second two components are left handed Weyl spinors. We can thus decompose any ψ into:

$$\psi = \psi_R + \psi_L$$

Furthermore, we could define the Dirac form as:

$$A = \epsilon\Gamma_0$$

however, this would give us $\delta = (-1)^{1+1} = 1$, and for reasons which will be clear in chapter 3 we do not want this. Indeed, if $\delta = 1$, it is easy to see that the Dirac equation found in **Theorem 3.1.13** demands trivial spinor fields. Instead, we define A by:

$$\begin{aligned} A &= \pm\, i\Gamma_1\Gamma_2\Gamma_3 \\ &= \begin{pmatrix} 0 & \mp iI \\ \pm iI & 0 \end{pmatrix} \end{aligned}$$

It follows that for $\psi = \psi_R + \psi_L$:

$$\begin{aligned} \bar{\psi} =& (\psi_R^\dagger, \psi_L^\dagger) \begin{pmatrix} 0 & \mp iI \\ \pm iI & 0 \end{pmatrix} \\ =& (\pm i\psi_L^\dagger, \mp i\psi_R^\dagger) \end{aligned}$$

so:

$$\bar{\psi}\psi = \pm i\psi_L^\dagger\psi_R \mp i\psi_R^\dagger\psi_L$$

Note that the Dirac form is not positive definite. Indeed, both right handed and left handed spinors are null.

2.2.7 Spinor Bundles

Recall that given a smooth manifold M an orientation on M was a pointwise orientation for each tangent such that such that for all $p \in M$ there existed a smooth local frame which agreed with the pointwise orientation. Equivalently, M could be covered by coordinate charts such that the Jacobian of each transition function was positive. Note that this implies that the frame bundle of M can be reduced to a principal $GL_n^+(\mathbb{R}) \subset GL_n(\mathbb{R})$ bundle, where $GL_n^+(\mathbb{R})$ is the subgroup consisting invertible linear transformations with positive determinants. Furthermore, recall that if (M, g) is a pseudo-Riemannian manifold, then the frame bundle can be reduced to a principal $O(t, s)$ bundle, as every neighborhood admits an orthonormal frame. With this in mind we turn to the following definition:

Definition 2.2.23. Let (M, g) be a pseudo-Riemannian manifold, then:

a) (M, g) is called **orientable** if the frame bundle can be reduced to a principal $SO(t, s)$ bundle under the embedding $SO(t, s) \hookrightarrow O(t, s)$.

b) (M, g) is called **time orientable** if the frame bundle can be reduced to a principal $O^+(t, s)$ bundle under the embedding $O^+(t, s) \hookrightarrow O(t, s)$.

c) (M, g) is called **orientable and time orientable** if the frame bundle can be reduced to a principal $SO^+(t, s)$ bundle under the embedding $SO^+(t, s) \hookrightarrow O(t, s)$.

We call (M, g) **oriented, time oriented, or oriented and time oriented** once a choice of reduction has been made.

Equivalently, in the Lorentzian signature $(1, n-1)$ case, we have that a Lorentzian manifold is time orientable if there exists a nowhere vanishing time like vector field.

Suppose that M is oriented, and time oriented, we denote by $SO^+(M)$ the the $SO^+(t, s)$ frame bundle:

$$\pi_{SO^+} : SO^+(t, s) \longrightarrow M$$

Definition 2.2.24. A **spin structure** on a pseudo-Riemannian manifold (M, g) is a $\mathrm{Spin}+(t, s)$ principal bundle:

$$\pi_{\mathrm{Spin}^+} : \mathrm{Spin}^+(M) \longrightarrow M$$

with a double covering:

$$\Lambda : \mathrm{Spin}^+(M) \longrightarrow SO^+(M)$$

such that the following diagram commutes:

$$\begin{tikzcd}
\mathrm{Spin}^+(M)\times\mathrm{Spin}^+(t,s) \arrow[r] \arrow[dd, "\Lambda\times\lambda"'] & \mathrm{Spin}^+(M) \arrow[dr, "\pi_{\mathrm{Spin}^+}"] \arrow[dd, "\Lambda"'] & \\
& & M \\
SO^+(M)\times SO^+(t,s) \arrow[r] & SO^+(M) \arrow[ur, "\pi_{SO^+}"'] &
\end{tikzcd}$$

where λ is the Lie group homomorphism $\mathrm{Spin}^+(t,s)$, and the horizontal arrows denote the group action on each principal bundle.

Note that a spin structure on M is a equivalent to λ-equivariant bundle homomorphism $\Lambda : \mathrm{Spin}^+(M) \to SO^+(M)$, and thus a λ-reduction of $SO^+(M)$.

Definition 2.2.25. Two spin structures $\Lambda : \mathrm{Spin}^+(M) \to SO^+(M)$, $\Lambda' : \mathrm{Spin}^+(M)' \to SO^+(M)$, are called **isomorphic** if there exists $\mathrm{Spin}^+(t,s)$ equivariant bundle isomorphism:

$$F : \mathrm{Spin}^+(M) \longrightarrow \mathrm{Spin}^+(M)'$$

such that:

$$\Lambda = \Lambda' \circ F$$

We note that the existence of spin structure on a pseudo Riemannian manifold is not guaranteed. There are obvious obstructions, such as the orientability and time orientability of (M,g), however, there are in fact deeper topological restrictions to the existence of a spin structure. Indeed the existence of a spin structure is intimately related to characteristic classes of vector bundles. In particular the vanishing of the second Stiefel-Whitney class of TM is a necessary and sufficient condition for $SO^+(M)$ to admit a spin structure. One can also show that the first vanishing of the first Stiefel-Whitney class of M is a necessary and sufficient condition for M to be orientable, so a spin structure can be thought of as a generalized orientability condition on M. Moreover, spin structures are in general not unique, but instead are related to the cohomology group of M. To the interested reader, we recommend Milnor's *Characteristic Classes* for an in depth treatment of Stiefel-Whitney classes, and Michelsohn and Lawson's *Spin Geometry* for a treatment of spin structures in the Riemannian case.

Note that every principal bundle over $\mathbb{R}^{t,s}$ is trivial, and that $\mathbb{R}^{t,s}$ is trivially a orientable and time orientable manifold. It follows that $SO^+(\mathbb{R}^{t,s})$ admits a spin structure:

$$\mathrm{Spin}^+(\mathbb{R}^{t,s}) = \mathbb{R}^{t,s} \times \mathrm{Spin}^+(t,s)$$

where the map Λ is given by $\mathrm{Id}_{\mathbb{R}^{t,s}} \times \lambda$, as:

$$SO^+(\mathbb{R}^{t,s}) = \mathbb{R}^{t,s} \times SO^+(t,s)$$

Going forward we assume that (M,g) is oriented, time oriented, and that $SO^+(M)$ admits a spin structure.

Definition 2.2.26. A local section $e = (e_1, \cdots, e_n)$ of $SO^+(M)$ is a **veilbein**, and corresponds to an oriented, and time oriented orthonormal frame of TM

Lemma 2.2.22. *Suppose we have chosen a spin structure on (M,g), then for every veilbein on a contractible open set of $U \subset M$ there exist precisely two local sections $\epsilon_\pm$ of Spin$^+(t,s)$ over U such that $\Lambda \circ \epsilon_\pm = e$.*

Proof. Note that the e is an injective immersion, and that π_{SO^+} restricted to im e is a continuous inverse of $e : U \to \mathrm{im}\ e$, so e is an embedding. It follows that im e is an embedded submanifold of $SO^+(M)$, and thus contractible. The restriction of the double covering:

$$\Lambda|_{\Lambda^{-1}(\mathrm{im}\ e)} : \Lambda^{-1}(\mathrm{im}\ e) \longrightarrow \mathrm{im}\ e$$

is then a trivial two sheeted covering of im e which admits precisely two sections. Denote these sections by $s_\pm$. It follows that:

$$\epsilon_\pm = s_\pm \circ e$$

are sections of $\mathrm{Spin}^+(M)$ satisfying $\Lambda \circ \epsilon_\pm = e$. $\square$

Definition 2.2.27. Let $\text{Spin}^+(M) \to M$ be a spin structure on an n dimensional pseudo Riemannian spin manifold M and:

$$\kappa : \text{Spin}^+(t,s) \longrightarrow GL(\Delta_n)$$

the spinor representation. The **Dirac spinor bundle** is then the associated complex vector bundle:

$$S = \text{Spin}^+(M) \times_\kappa \Delta_n$$

over M. Sections of this bundle are called **spinor fields** or **spinors**.

Proposition 2.2.13. *Let $S \to M$ be a Dirac spinor bundle associated to a spin structure $Spin^+(M) \to M$ of an n dimensional pseudo Riemannian manifold M. Then the following hold:*

a) There exists a well defined bilinear Clifford multiplication:

$$\begin{aligned} TM \times S &\longrightarrow S \\ (X, \Psi) &\longmapsto X \cdot \Psi \end{aligned}$$

on the level bundles, which restricts to a bilinear map $T_xM \times S_x \to S_x$ for all $x \in M$. This map also induces well-defined Clifford multiplication of forms with spinors.

*b) If n is even, then S splits into a direct sum of **Weyl Spinor Bundles** $S = S_+ \oplus S_-$ defined by:*

$$S_\pm = Spin^+(M) \times_\kappa \Delta_n^\pm$$

Proof. We begin with a). Recall that if ρ_{SO^+} is the standard representation of $SO^+(t,s)$ on $\mathbb{R}^{t,s}$ then the tangent bundle of M is given by:

$$TM = SO^+(M) \times_{\rho_{SO^+}} \mathbb{R}^{t,s}$$

We can then define the map:

$$\begin{aligned} (SO^+(M) \times_{\rho_{SO^+}} \mathbb{R}^{t,s}) \times (\text{Spin}^+(M) \times_\kappa \Delta^n) &\longrightarrow \text{Spin}^+(M) \times_\kappa \Delta^n \\ ([\Lambda(p), v], [p, \psi]) &\longmapsto [p, v \cdot \psi] \end{aligned}$$

where $p \in \text{Spin}^+(M)$, $v \in \mathbb{R}^{t,s}$, and $\psi \in \Delta_n$. Note that since $p \in \text{Spin}^+(M)$ we have that $\Lambda(p) \in SO^+(M)$, and both lie in a fibre over $\pi_{\text{Spin}^+}(p)$. To see that this map is well defined, let $g \in \text{Spin}^+(t,s)$, and $q = p \cdot g$, and $\phi = \kappa(g)^{-1} \cdot \psi$, then:

$$[q, \phi] = [p \cdot g, \kappa(g)^{-1} \cdot \psi] = [p, \psi]$$

Furthermore, we have that $\Lambda(p \cdot g) = \Lambda(p) \cdot \lambda(g)$, so if $w = \rho_{SO^+}(\lambda(g))^{-1} \cdot v$, we have that:

$$[\Lambda(q), w] = [\Lambda(p) \cdot \lambda(g), \rho_{SO^+}(\lambda(g)^{-1}) \cdot v] = [\Lambda(p), v]$$

Hence by **Proposition 2.2.11**:

$$\begin{aligned} [\Lambda(q), w] \cdot [q, \phi] =& [p, w \cdot \phi] \\ =& [p \cdot g, (\rho_{SO^+}(\lambda(g)^{-1}) \cdot v)(\kappa(g)^{-1} \cdot \psi)] \\ =& [p \cdot g, \kappa(g)^{-1}(v \cdot \psi)] \\ =& [p, v \cdot \psi] \end{aligned}$$

so the map is well defined. By examining local sections, it is easy to see that this map is smooth, and moreover when restricting the map to the fibre $T_xM \times S_x$ for $x \in M$ it is clear we obtain a bilinear map between vector spaces, implying the claim.

Note that there is an induced representation ρ'_{SO^+} of $SO^+(t,s)$ on $\Lambda^k(\mathbb{R}^{t,s})$ given by:

$$\rho'_{SO^+}(g) \cdot \sum_{i_1 \cdots i_k} (v_{i_1} \wedge \cdots \wedge v_{i_k}) = \sum_{i_1 \cdots i_k} (\rho_{SO^+}(g) \cdot v_{i_1}) \wedge \cdots \wedge (\rho_{SO^+}(g) \cdot v_{i_k})$$

One can then easily show that:

$$\Lambda^k(TM) = SO^+(M) \times_{\rho'_{SO^+}} \Lambda^k(t,s)$$

and the same argument demonstrates that the map:

$$\begin{aligned}(SO^+(M) \times_{\rho'_{SO^+}} \Lambda^k(\mathbb{R}^{t,s})) \times (\mathrm{Spin}^+(M) \times_\kappa \Delta^n) &\longrightarrow \mathrm{Spin}^+(M) \times_\kappa \Delta^n \\ ([\Lambda(p), \omega], [p, \psi]) &\longmapsto [p, \omega \cdot \psi]\end{aligned}$$

is well defined; here $\omega\cdot\psi$ is given by **Definition 2.2.14**, Furthermore, the metric g induces a canonical bundle isomorphism $F : \Lambda^k(T^*M) \to \Lambda^k(TM)$, so we can define the Clifford multiplication of differential k forms, by composing Clifford multiplication with the isomorphism F.

To prove b) recall that M is orientable, in particular M admits a nowhere vanishing volume form given in any orthonormal frame by:

$$\mathrm{dvol}_g = e^1 \wedge \cdots \wedge e^n$$

Under the isomorphism $\Lambda^n(T^*M) \to \Lambda^n(TM)$ induced by g, we obtain a nowhere vanishing element α of $\Lambda^n(TM)$ which in any orthonormal frame is given by:

$$\alpha = e_1 \wedge \cdots \wedge e_n$$

Define $\xi \in \Lambda^n(TM) \otimes \mathbb{C}$ by:

$$\xi = -i^{\frac{n}{2}+t}\alpha$$

It follows that for any spinor $\Psi = [p, \psi] \in S$, and any orthonormal basis $\{e_i\}$ for $\mathbb{R}^{t,s}$:

$$\begin{aligned}\xi_x \cdot \Psi =& [\Lambda(p), -i^{n/2+t} e_1 \wedge \cdots \wedge e_n] \cdot [p, \psi] \\ =& [p, -i^{n/2+t}(e_1 \wedge \cdots \wedge e_n) \cdot \psi] \\ =& [p, \omega \cdot \psi]\end{aligned}$$

where ω is the chirality element. With this in mind, we can think of the action of ξ on S as the action of the chirality element on the Spinor bundle. Since ξ is well defined, i.e. independent of our orthonormal frame, and nowhere vanishing it follows that the map:

$$\begin{aligned}\omega : S &\longrightarrow S \\ \Psi &\longmapsto \xi \cdot \Psi\end{aligned}$$

is a well defined global bundle automorphism, that satisfies $\omega^2 = \mathrm{Id}_S$. The following bundle map:

$$\begin{aligned}F : S &\longrightarrow S^+ \oplus S^- \\ [p, \psi] &\longmapsto ([p, \psi_R], [p, \psi_L])\end{aligned}$$

is then a well defined global bundle isomorphism, since every $\Psi \in S$ can be decomposed into the $\pm\mathrm{Id}$ eigenspaces of $\omega : S \to S$. In particular, we have that:

$$S^\pm = \ker(\omega \mp \mathrm{Id}_s)$$

which are smooth subbundles of S as the kernel of $\omega \mp \mathrm{Id}_s$ has constant rank. □

As Dirac spinor bundles are in essence vector bundles, we can construct bundle metrics on them. In particular we are interested in a specific class of bundle metrics defined below:

Definition 2.2.28. A **Dirac bundle metric** is a bundle metric $\langle\cdot,\cdot\rangle_S$ on the associated complex vector bundle $\mathrm{Spin}^+(M) \times_\kappa \Delta_n$ satisfying the following properties:

a) $\langle X \cdot \Psi, \Phi\rangle_{S_x} = \delta\langle \Psi, X \cdot \Phi\rangle_{S_x}$ for all $X \in T_xM$, and all $\Psi, \Phi \in S_x$.

b) $\langle \Psi, \Phi\rangle_{Sx} = \overline{\langle \phi, \psi\rangle}_{S_x}$ for all $\Psi, \Phi \in S_x$.

c) $\langle \Psi, c\Phi\rangle_{S_x} = c\langle \Psi, \Phi\rangle_{S_x} = \langle \bar{c}\Psi, \Phi\rangle_{S_x}$ for all $\Psi, \Phi \in S_x$ and all $c \in \mathbb{C}$.

Note that since bundle metrics are $C^\infty(M)$ bilinear, similar conditions follow for vector fields, spinor fields, and complex valued functions on M

From **Lemma 2.2.21** and **Proposition 2.1.13** we have the following corollary:

Corollary 2.2.12. *Let M be an oriented and time oriented pseudo Riemannian manifold with spin structure $Spin^+(M)$. Then, $S = Spin^+(M) \times_\kappa \Delta_n$ admits a Dirac bundle metric.*

Proof. Choose a basis unitary representation of $\mathrm{Cl}(t,s)$ on Δ_n, then there exists a Dirac form by **Proposition 2.2.12**. By **Lemma 2.2.21** the Dirac form is invariant under the action of $\mathrm{Spin}(t,s)$, so the claim follows by **Proposition 2.2.13**. □

As mentioned earlier, given two vector bundles E and F over M, we can define a new vector bundle $E \otimes F$, where the fibres are given by the tensor product $E_x \otimes F_x$. We are interested in the following specific case of this construction:

Definition 2.2.29. Let M an oriented and time oriented pseudo Riemannian manifold with spin structure $\mathrm{Spin}^+(M)$, $\pi : P \to M$ be a principal bundle with structure group G, and $E = P \times_\rho V$ be a complex vector bundle associated to the complex representation $\rho : G \to GL(V)$. The vector bundle:

$$S \otimes E$$

is called the **twisted spinor bundle** or **gauge multiplet spinor bundle**.

We want to show that the twisted spinor bundle is a vector bundle associated to a principal $\mathrm{Spin}^+(t,s) \times G$ bundle. However, it should be clear that $\mathrm{Spin}^+(M) \times P$ does not fit the bill. Instead we need the following construction:

Definition 2.2.30. The **fibre product** of $\mathrm{Spin}^+(M)$ and $\pi : P \to M$ is the disjoint union:

$$\mathrm{Spin}^+(M) \times_M P = \coprod_{x \in M} \pi^{-1}_{\mathrm{Spin}^+}(x) \times \pi^{-1}(x)$$

Equivalently the fibre product is the set:

$$\mathrm{Spin}^+(M) \times_M P = \{(p,q) \in \mathrm{Spin}^+(M) \times P : \pi_{\mathrm{Spin}^+}(p) = \pi(q)\}$$

Proposition 2.2.14. *The fibre product of $Spin^+(M)$ and P has the structure of a principal $Spin^+(t,s) \times G$ principal bundle.*

Proof. The disjoint union:

$$\mathrm{Spin}^+(M) \times_M P = \coprod_{x \in M} \pi^{-1}_{\mathrm{Spin}^+}(x) \times \pi^{-1}(x)$$

comes equipped with a natural projection map $\pi_\times : \mathrm{Spin}^+(M) \times_M P \to M$. We want to construct a smoothly compatible fibre bundle atlas. Let $\{U_i\}_i$ be and open cover of M such that each U_i is contained in a coordinate chart. It follows that each U_i is contractible and diffeomorphic to an open set of $\mathbb{R}^n$, and thus $\mathrm{Spin}^+(M)_{U_i}$ and P_{U_i} are trivial, hence we have principal bundle atlases $\{U_i, \phi_i\}$ and $\{U_i, \psi_i\}$ for $\mathrm{Spin}^+(M)$ and P respectively. Define charts by:

$$\begin{aligned} \phi_i \times_M \psi_i : \pi_\times^{-1}(U_i) &\longrightarrow U_i \times (\mathrm{Spin}^+(t,s) \times G) \\ (p,q) &\longmapsto (\pi(q), \pi_{\mathrm{Spin}^+(t,s)} \circ \phi_i(p), \pi_G \circ \psi_i(q)) \end{aligned}$$

where $\pi_{\mathrm{Spin}^+(t,s)}$ and π_G are the projections onto $\mathrm{Spin}^+(t,s)$ and G in their respective trivializations. This map has an inverse given by:

$$\begin{aligned} (\phi_i \times_M \psi_i)^{-1} : U_i \times (\mathrm{Spin}^+(M) \times G) &\longrightarrow \pi_\times^{-1}(U_i) \\ (x,s,g) &\longmapsto (\phi_i^{-1}(x,s), \psi_i^{-1}(x,g)) \end{aligned}$$

It should be clear that these are indeed inverses of another from the fact that $\pi(q) = \pi_{\mathrm{Spin}^+}(p)$, and that ϕ_i^{-1} and ψ_i^{-1} are smooth inverses of ϕ_i and ψ_i. It is also easy to see that the transition functions are smooth, as they are smooth in each component. From **Theorem 2.1.1** it follows that $\mathrm{Spin}^+(M) \times_M P$ is a fibre bundle over M. In particular, we obtain a smooth right action of $\mathrm{Spin}^+(t,s) \times G$ on $\mathrm{Spin}^+(M) \times_M P$ given by:

$$\begin{aligned} \Phi : (\mathrm{Spin}^+(M) \times_M P) \times (\mathrm{Spin}^+(t,s) \times G) &\longmapsto \mathrm{Spin}^+(M) \times_M P \\ ((p,q),(s,g)) &\longmapsto (p,q)\cdot(s,g) = (p\cdot s, q\cdot g) \end{aligned}$$

Since the separate actions of $\mathrm{Spin}^+(t,s)$ and G on $\mathrm{Spin}^+(M)$ and P are free and transitive on the respective fibres, it follows that this action is free and transitive on the fibres of $\mathrm{Spin}^+(M) \times_M P$ as:

$$(\mathrm{Spin}^+(M) \times_M P)_x = \pi^{-1}_{\mathrm{Spin}^+}(x) \times \pi^{-1}(x)$$

by construction. Moreover, since the separate actions preserve the fibres, we have that this action preserves the fibres of $\mathrm{Spin}^+(M) \times_M P$, again by construction. By **Proposition 2.1.5**, we have that $\mathrm{Spin}^+(M) \times_M P$ is a principal bundle with structure group $\mathrm{Spin}^+(t,s) \times G$. $\square$

Corollary 2.2.13. *Let $\pi_{Spin^+}(p) = \pi(q)$, then $T_{(p,q)}(Spin^+(M) \times_M P) \subset T_pSpin^+(M) \times T_qP$ is the subspace defined by:*

$$W = \{(X_p, Y_q) \in T_pSpin^+(M) \times T_qP : \pi_{Spin^+*}X_p = \pi_* Y_q\}$$

Proof. Let $Z \in T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P)$, be the tangent vector to a curve $\gamma : I \to \mathrm{Spin}^+(M) \times_M P$, such that $\gamma(0) = (p,q)$ and $\dot\gamma(0) = Z$. It follows that γ can be written as:

$$\gamma(t) = (\gamma_{\mathrm{Spin}^+}(t), \gamma_P(t))$$

where $\gamma_{\mathrm{Spin}^+}(0) = p$, $\gamma_P(0) = q$ and for all $t \in I$:

$$\pi_{\mathrm{Spin}^+}(\gamma_{\mathrm{Spin}^+}(t)) = \pi(\gamma_P(t))$$

If $\dot\gamma_{\mathrm{Spin}^+}(0) = X_p$ and $\dot\gamma_P(0) = Y_q$, then $Z = (X_p, Y_q) \in T_p\mathrm{Spin}^+(M) \times T_qP$, such that:

$$\pi_{\mathrm{Spin}^+*}X_p = \pi_* Y_q$$

therefore:

$$T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P) \subset W$$

We note that the dimension of $T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P)$ is $\dim M + \dim \mathrm{Spin}(t,s) + \dim G$, while dimension of $T_p\mathrm{Spin}^+(M) \times T_qP$ is $\dim M + \dim T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P)$. Via a principal bundle chart, we can construct an isomorphism:

$$T_p\mathrm{Spin}^+(M) \times T_qP \cong T_xM \oplus T_p\mathrm{Spin}^+(M)_x \oplus T_xM \oplus T_qP_x$$

Let $W' \cong W$ be the image of W under this isomorphism. The condition that $\pi_{\mathrm{Spin}^+(M)}*X_p = \pi_*Y_q$ will force that W' is the vector subspace consisting of vectors where the T_xM components are the same. We can thus construct an isomorphism:

$$W' \cong T_xM \oplus T_p\mathrm{Spin}^+(M)_x \oplus T_qP_x$$

implying that W', and thus W, has the same dimension as $T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P)$. It follows that $W = T_{(p,q)}(\mathrm{Spin}^+(M) \times_M P)$. $\square$

We can now show the following

Proposition 2.2.15. *The twisted spinor bundle $S \otimes E$ is a vector bundle associated to the principal bundle $Spin^+(M) \times_M P$.*

Proof. Consider the representation $\kappa \otimes \rho$ of $\mathrm{Spin}^+(t,s) \times G$ on $\Delta_n \otimes V$ given on simple tensors by:

$$\kappa \otimes \rho(s,g)(\psi \otimes v) = \kappa(s)\psi \otimes \rho(g)v$$

It should be clear that this a representation, as it is a smooth map, and and homomorphism, and thus a Lie group homomorphism:

$$\mathrm{Spin}^+(t,s) \times G \longrightarrow GL(\Delta_n \otimes V)$$

Let $\mathscr{E}$ be the associated vector bundle:

$$\mathscr{E} = (\mathrm{Spin}^+(M) \times_M P) \times_{\kappa\otimes\rho} (\Delta_n \otimes V)$$

and define the smooth map given on simple tensors by:

$$\begin{aligned} F : \mathscr{E} &\longrightarrow S \otimes E \\ [(p,q), \psi \otimes v] &\longmapsto [p,\psi] \otimes [q,\phi] \end{aligned}$$

We check that this map is well defined, first note that for any $(s,g) \in \mathrm{Spin}^+(t,s) \times G$, we have that:

$$\begin{aligned} [(p,q)\cdot(s,g), \kappa\otimes\rho(s,g)^{-1}\psi\otimes v] =& [(p\cdot s, q\cdot g), \kappa\otimes\rho(s^{-1},g^{-1})\psi\otimes v] \\ =& [(p\cdot s, q\cdot g), \kappa(s)^{-1}\psi\otimes\rho(g)^{-1}v] \end{aligned}$$

Hence:

$$\begin{aligned} F([(p,q)\cdot(s,g), \kappa\otimes\rho(s,g)^{-1}\psi\otimes v]) =& [p\cdot s, \kappa(s)^{-1}\psi] \otimes [q\cdot g, \rho(g)^{-1}v] \\ =& [p,\psi]\otimes[q,\phi] \end{aligned}$$

It is also clear that this satisfies:

$$\pi_{S\otimes E} \circ F = \pi_{\times}$$

as p and q both project down to the same $x \in M$. The restriction of F to the fibre $\mathscr{E}_x$ is clearly linear and surjective as for any:

$$\sum_i [p,\psi_i] \otimes [q, v_i] \in (S\otimes E)_x$$

we have that:

$$a = \left[(p,q), \sum_i \psi_i \otimes v_i\right]$$

satisfies :

$$F_x(a) = \sum_i [p,\psi_i]\otimes[q,v_i]$$

Since $\dim_{\mathbb{C}} \mathscr{E}_x = \dim_{\mathbb{C}}(S\oplus E)_x$, we have that F_x is a linear isomorphism for all $x \in M$, and thus a vector bundle isomorphism. $\square$

Proposition 2.2.16. *Let $\dim_{\mathbb{C}} V = r$, and $\{\tau_i\}$ be a local frame for E_U. Then any $\Psi \in \Gamma(S\otimes E)$ can be written locally as:*

$$\Psi|_U = \sum_{i=1}^r \Psi_i \otimes \tau_i$$

where ψ_i are local sections of S. Equivalently:

$$\Psi = [\epsilon \times s, \psi]$$

where ϵ and s are local sections of Spin$^+(M)$ and P, and ψ is the multiplet:

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_r \end{pmatrix} : U \to \Delta_n \otimes \mathbb{C}^r$$

Here each ψ_i is a smooth map $U \to \Delta_n$.

Proof. Let Ψ be a spinor field, then for all $x \in U$, we have that $\Psi_x \in S_x \otimes E_x$. Any tensor product can be decomposed into a sum of simple tensors, so:

$$\Psi_x = \sum_j \Psi_j \otimes v_j$$

for some $\Psi_j \in S_x$ and $v_j \in E_x$. Since $\{\tau_i\}$ is a local frame, it follows that τ_{ix} forms a basis for E_x, hence we can write this as:

$$\Psi_x = \sum_{i=1}^{r} \Psi_i \otimes \tau_{ix} \tag{2.2.16}$$

where we have absorbed the coefficients of each v_j into new elements $\psi_i \in S_x$. Doing this for all $x \in U$ we obtain maps $\Psi_i : U \to S_U$, which must be smooth as each τ_{ix} is smooth and linearly independent, and $\Psi|_U$ is smooth. It follows that:

$$\Psi|_U = \sum_{i=1}^{r} \Psi_i \otimes \tau_i$$

To proceed with the second part of the proof we note that for local sections $\epsilon : U \to \mathrm{Spin}^+(M)_U$, and $s : U \to P_U$, the smooth map:

$$\begin{aligned} \epsilon \times_M s : U &\longrightarrow \mathrm{Spin}^+(M) \times_M P \\ x &\longmapsto (\epsilon(x), s(x)) \end{aligned}$$

is a local section of $(\mathrm{Spin}^+(M) \times_M P)_U$. It then follows that for smooth maps $\psi : U \to \Delta_n$ and $\phi : U \to V$ the smooth map:

$$\begin{aligned} \Psi|_U : U &\longmapsto (S \otimes E)_U \\ x &\longmapsto [\epsilon \times_M s(x), \psi(x) \otimes \phi(x)] \end{aligned}$$

Let the local frame τ_i be given by:

$$\tau_i = [s, v_i]$$

where $\{v_i\}$ forms a basis for V. Representing each v_i as the column vector with a 1 in the ith column, and 0's in the rest, we can represent any smooth map:

$$\psi : U \to \Delta_n \otimes E$$

can be written as:

$$\psi = \sum_i \psi_i \otimes v_i = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_r \end{pmatrix}$$

where each ψ_i is a smooth map $U \to \Delta_n$. Suppose that each ψ_i satisfies:

$$[\epsilon, \psi_i] = \Psi_i$$

where Ψ_i are the maps defined in (2.2.16). We then have that:

$$[\epsilon \times s, \psi] = \sum_i^r [\epsilon \times s, \psi_i \otimes v_i]$$

Under the isomorphism $F : \mathscr{E} \to S \otimes E$ we obtain:

$$\begin{aligned} F([\epsilon \times s, \psi]) &= \sum_i [\epsilon, \psi_i] \otimes [s, v_i] \\ &= \sum_i^r \psi_i \otimes \tau_i \end{aligned}$$

so the two constructions are equivalent. $\square$

2.2.8 The Spin Covariant Derivative and the Dirac Operator

Definition 2.2.31. Let (M,g) be a pseudo Riemannian manifold, then the **Levi-Civita connection** is the unique covariant derivative on the tangent bundle TM which is torsion free, i.e.

$$\nabla_X Y - \nabla_Y X = [X,Y]$$

for all $X,Y,Z \in \mathfrak{X}(M)$, and metric compatible:

$$\nabla g = 0$$

We will prove the existence and uniqueness of the Levi-Civita connection in a later chapter, with the use of a soldering form, and a connection on $SO^+(M)$. We will also prove the more standard method found in most textbooks on Riemannian geometry.Furthermore, note that the covariant derivative of a $(0,2)$ tensor field ξ is defined implicitly by:

$$(\nabla_X \xi)(Y,Z) = \nabla_X(\xi(Y,Z)) - \xi(\nabla_X Y, Z) - \xi(Y, \nabla_X Z)$$

Since $\nabla_X(\xi(X,Y)) = \mathscr{L}_X(\xi(Y,Z))$, we see that the metric compatible condition is given by:

$$\mathscr{L}_X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

which is exactly equivalent to **Proposition 2.1.23**. To give away the plot a bit, one can use the metric compatible condition, and the torsion free condition to demonstrate that the Levi-Civita connection is uniquely determined by the *Koszul formula*:

$$\begin{aligned} 2g(\nabla_X Y, Z) =& \mathscr{L}_X g(Y,Z) + \mathscr{L}_Y g(Z,X) - \mathscr{L}_Z g(X,Y) \\ &+ g([X,Y],Z) - g([Y,Z],X) + g([Z,X],Y) \end{aligned} \tag{2.2.17}$$

Using this, one can also show that ∇ satisfies the properties of a covariant derivative, a result we will explicitly demonstrate in chapter 3.2.

Proposition 2.2.17. *In any local orthonormal frame $\{e_i\}$ for TM_U the Levi Civita connection satisfies:*

$$\nabla e_a = \xi_{ab}\eta^{bc} \otimes e_c$$

where ξ_{ab} are one forms that satisfy:

$$\xi_{ab} = -\xi_{ba}$$

and the components of η are the components of the standard metric tensor on $\mathbb{R}^{t,s}$.

Proof. Note that ∇ is a map $\mathfrak{X}(M) \to \Omega^1(M,TM)$. It follows that locally, we can write:

$$\nabla e_a = \xi_a^c \otimes e_c$$

where ξ_a^c are one forms on U. Note that:

$$g(e_a, e_b) = \eta_{ab}$$

is constant on U since e_a and e_b are orthonormal. The metric compatible condition then states that for any arbitrary $X \in \mathfrak{X}(M)$:

$$g(\nabla_X e_a, e_b) = -g(e_a, \nabla_X e_b)$$

We see that:

$$g(\nabla_X e_a, e_b) = g(\xi_a^c(X) e_c, e_b) = \eta_{cb}\xi_a^c(X)$$

while:

$$g(e_a, \nabla_X e_b) = \eta_{ad}\xi_b^d(X)$$

It follows that:

$$\eta_{cb}\xi_a^c(X) = -\eta_{ad}\xi_b^d(X)$$

Define one forms by:

$$\xi_{ab} = \xi_a^c\eta_{cb} \qquad \text{and} \qquad \xi_{ba} = \xi_b^d\eta_{da}$$

we then obtain that:

$$\xi_{ab} = -\xi_{ba}$$

Furthermore, we see that:

$$\begin{aligned} \xi_{ab}\eta^{bc} \otimes e_c =& \eta_{db}\xi_a^d\eta^{bc} \otimes e_c \\ =& \delta_d^c\xi_a^d \otimes e_c \\ =& \xi_a^c \otimes e_c \end{aligned}$$

so the Levi-Civita connection can be written as:

$$\nabla e_a = \xi_a^c \otimes e_c = \xi_{ab}\eta^{bc} \otimes e_c$$

where ξ_{ab} is antisymmetric in the indices a and b, as desired. □

We now want to show that these one forms are uniquely determined in any local orthonormal frame.

Definition 2.2.32. The **anaholonomy coefficients** of a local orthonormal frame are defined by:

$$[e_a, e_b] = \Omega_{ab}^c e_c$$

Proposition 2.2.18. *In a local orthonormal frame the one forms ξ_{ab} are uniquely determined by the formula:*

$$\xi_{ab}(e_c) = \frac{1}{2}\left(\Omega_{cab} - \Omega_{abc} + \Omega_{bca}\right)$$

where $\Omega_{abc} = \Omega_{ab}^d\eta_{dc}$.

Proof. This is follows from the Koszul equation; we have that:

$$\begin{aligned} g(\nabla_{e_c}e_a, e_b) =& \xi_{ad}(e_c)\eta^{df}g(e_f, e_b) \\ =& \xi_{ad}(e_c)\eta^{df}\eta_{fb} \\ =& \xi_{ad}(e_c)\delta_b^d \\ =& \xi_{ab}(e_c) \end{aligned}$$

Furthermore, since $\{e_i\}$ is a local orthonormal frame we have that:

$$\mathscr{L}_{e_c}g(e_a, e_b) = \mathscr{L}_{e_a}g(e_b, e_c) = \mathscr{L}_{e_b}(e_c, e_a) = 0$$

as the inner product is constant. Now note that:

$$g([e_c, e_a], e_b) = g(\Omega_{ca}^d e_d, e_b) = \Omega_{ca}^d\eta_{db} = \Omega_{cab}$$

It those follows from (2.2.17) that:

$$2\xi_{ab}(e_c) = \Omega_{cab} - \Omega_{abc} + \Omega_{bca}$$

implying the claim. □

As mentioned earlier, in chapter 3 we will demonstrate the existence and uniqueness of a connection one form on $SO^+(M)$ which corresponds to the Levi-Civita connection. For now, we show the following converse:

Proposition 2.2.19. *The Levi-Civita connection induces a unique connection one form $A_{SO^+} \in \Omega^1(SO^+(M), \mathfrak{so}^+(t,s))$.*

Proof. Note that the tangent bundle TM is the associated bundle:

$$TM = SO^+(M) \times_\rho \mathbb{R}^n$$

where ρ is the standard representation of $SO^+(t,s)$ on $\mathbb{R}^{t,s}$. Since ρ is the standard representation of $SO^+(t,s)$, it follows that the induced representation ρ_* of $\mathfrak{so}^+(t,s)$ on $\mathbb{R}^{t,s}$ is given by:

$$\rho_*(X)v = Xv$$

where X is any matrix satisfying:

$$X^T\eta + \eta X = 0 \tag{2.2.18}$$

Let $\{e_i\}$ be an local oriented and time oriented orthonormal frame for an open neighborhood U of $x \in M$, then **Proposition 2.2.17** implies that that:

$$\eta_{cb}\xi^c_a + \eta_{ac}\xi^c_b = 0$$

so the one forms ξ^c_a are the components of a one form valued in the Lie algebra $\mathfrak{so}^+(t,s)$, i.e. a matrix valued one form which satisfies (2.2.18). Let $\{f_i\}$ be the standard orthonormal basis of $\mathbb{R}^{t,s}$, then if $e : U \to SO^+(M)$ is the frame $(e_1, \ldots, e_n)$, we have that by **Example 2.1.4**:

$$e_i = [e, f_i]$$

as f_i is the standard column vector with a 1 in the ith entry, and a 0 in all others. Let:

$$T^j_i = f_i \otimes f^j$$

where $\{f^j\}$ is the basis dual to $\{f_i\}$, then we see that in the coordinates $\{x^i\}$ the matrix valued one form

$$\xi = (\xi^i_{j\mu} dx^\mu) \otimes T^j_i$$

satisfies:

$$\begin{aligned}
[e, \rho_*(\xi(X))f_a] =& [e, \xi^c_j(X)T^j_c(f_a)] \\
=& [e, \xi^c_j(X)f_c \otimes f^j(f_a)] \\
=& [e, \xi^c_a(X)f_c] \\
=& \xi^c_a(X)[e, f_c] \\
=& \xi^c_a(X)e_c
\end{aligned}$$

which is precisely equal to $\nabla_X e_c$. We thus define the local connection one form on U by:

$$e^*(A_{SO^+}) = \xi$$

Let ϕ be the trivialization corresponding to e, then we define a one form on $U \times SO^+(t,s)$ by:

$$\omega^e((X,Y)_{(x,g)}) = \mathrm{Ad}_{g^{-1}} \circ \xi(X_x) + \mu_{SO^+}(Y_g) \tag{2.2.19}$$

for all $(X,Y)_{x,g} \in T_xU \oplus T_gSO^+(t,S)$. Note that this is smooth as the above definition of ω^e is equivalent to setting:

$$\omega^e_{(x,g)} = \mathrm{Ad}_{g^{-1}} \circ (\pi^*_U \xi_x) + \pi^*_{SO^+(t,s)}\mu_{SO^+}$$

We define A locally by:

$$A_{SO^+}|_{\pi^{-1}(U)} = \phi^*\omega^e \tag{2.2.20}$$

As a sanity check, let us pull this local form back by e. Note that $\phi \circ e(x) = (x, \mathrm{Id}_n)$, so for all $X \in T_xU$:

$$(\phi \circ e)_*(X) = (X, 0)$$

It follows that:

$$e^*(A_{SO^+}|_{\pi^{-1}(U)})_x(X) = \omega^e_{(x,\mathrm{Id}_n)}(X, 0) = \xi_x(X)$$

hence:

$$e^*(A_{SO^+}|_{\pi^{-1}(U)}) = \xi$$

as desired. We now show that this $A_{SO^+}|_{\pi^{-1}(U)}$ satisfies the properties of a connection one form. First, let $X \in\in \mathfrak{g}$, and let $\tilde{X}$ be the associated vertical vector field, then:

$$\phi_*(\tilde{X})_{(x,g)} = (0_x, X_g)$$

where 0_x is the zero vector in T_xM, and X_g is evaluated at g. Therefore:

$$A_{SO^+}|_{\pi^{-1}(U)}(\tilde{X}_p) = \mu_{SO^+}(X_g) = X$$

Finally, for $p \in P$ let $\phi(p) = (x, g)$, and set $\phi_*(Z_p) = (X_x, Y_g)$ for $Z_p \in T_pP$, then, for all $h \in H$ we have that by the $SO^+(t, s)$ equivariance of ϕ:

$$\begin{aligned}
(R_h^* A_{SO^+}|_{\pi^-(U)})_p(Z_p) =&((\phi \circ R_h)^*\omega^e)_p(Z_p)\\
=&((R_h \circ \phi)^*\omega^e)_p(Z_p)\\
=&(R_h^*\omega^e_{(x,g)})(X_x, Y_g)\\
=&R_h^*(\mathrm{Ad}_{g^{-1}} \circ \xi)(X_x) + (R_h^*\mu_{SO^+})(Y_g)\\
=&\mathrm{Ad}_{h^{-1}g^{-1}} \circ \xi(X_x) + \mathrm{Ad}_{h^{-1}} \circ \mu_{SO^+}(Y_g)\\
=&\mathrm{Ad}_{h^{-1}} \circ (\mathrm{Ad}_{g^{-1}} \circ \xi(X_x) + \mu_{SO^+}(Y_g))\\
=&\mathrm{Ad}_{h^{-1}} \circ (A_{SO^+}|_{\pi^-(U)})_p(Z_p)
\end{aligned}$$

Hence $A|_{\pi^{-1}(U)}$ is Ad invariant, and it follows that (2.2.20) defines a unique connection one form on $\pi^{-1}(U)$.

Now let $\tilde{e} : V \to SO^+(M)$ be another local section such that $U \cap V \neq 0$, and ψ it's corresponding trivialization. Then $\tilde{e}$ defines a a $\mathfrak{so}^+(t, s)$ valued one form on $\tilde{\xi}$ satisfying:

$$\nabla(\tilde{e}_a) = \tilde{\xi}^c_a \otimes \tilde{e}_c$$

With $\tilde{\xi}$ instead of ξ, and ψ instead of ϕ, we define $\omega^{\tilde{e}}$, and $A_{SO^+}|_{\pi^{-1}(V)}$, as in 2.2.19 and 2.2.20 respectively. In order for A_{SO^+} to be globally defined, we thus need to show that on $\pi^{-1}(U \cap V)$ both definition of A_{SO^+} agree, i.e.

$$\phi^*\omega^e = \psi^*\omega^{\tilde{e}}$$

which is equivalent to showing that:

$$\omega^{\tilde{e}} = (\phi \circ \psi^{-1})^*\omega^e$$

Let $h : U \cap V \to SO^+(t, s)$ be the physical gauge transformation satisfying:

$$\tilde{e} = e \cdot h$$

then since h is a matrix of functions on $U \cap V$ we have that:

$$e \cdot h = (h^i_1 e_i, \dots, h^i_n e_i)$$

hence:

$$\tilde{e}_i = h^j_i e_j$$

Therefore, by the Leibniz rule for covariant derivatives:

$$\begin{aligned}\nabla(\tilde{e}_a) =& h^i_a \nabla e_i + dh^j_a \otimes e_j \\ =& h^i_a \xi^j_i \otimes e_j + dh^j_a \otimes e_j\end{aligned}$$

Note that:

$$e_j = (h^{-1})^k_j \tilde{e}_k$$

so:

$$\begin{aligned}\nabla(\tilde{e}_a) =& h^i_a \xi^j_i (h^{-1})^k_j \otimes \tilde{e}_k + (h^{-1})^k_j dh^j_a \otimes \tilde{e_k} \\ =& (h^i_a \xi^j_i (h^{-1})^k_j + (h^{-1})^k_j dh^j_a) \otimes \tilde{e}_k\end{aligned}$$

It follows that the kth component is given by:

$$\nabla(\tilde{e}_a)^k = (h^{-1})^k_j \xi^j_i h^i_a + (h^{-1})^k_j dh^j_a$$

Therefore:

$$\tilde{\xi}^k_a = (h^{-1})^k_j \xi^j_i h^i_a + (h^{-1})^k_j dh^j_a$$

hence we obtain that:

$$\tilde{\xi} = (h^{-1})\xi h + h^{-1} dh$$

In our usual notation, this is equivalent to:

$$\tilde{\xi} = \mathrm{Ad}_{h^{-1}} \circ \xi + h^* \mu_{SO^+} \tag{2.2.21}$$

Moreover, since:

$$\psi^{-1}(x, g) = \tilde{e}(x) \cdot g = e(x) \cdot (h(x) \cdot g)$$

we have that:

$$(\phi \circ \psi^{-1})(x, g) = (x, h(x) \cdot g)$$

Since $h(x) \cdot g$ is the composition of maps:

$$F : (U \cap V) \times G \longrightarrow G \times G \longrightarrow G$$

where the first map is (h, Id_G), and the second is multiplication in G, we have that:

$$(\phi \circ \psi^{-1})_*(X, Y) = (X, R_{g*}(D_x h(X)) + L_{h*} Y)$$

for all $(X, Y) \in T_x(U \cap V) \oplus T_g G$. We thus have that:

$$\begin{aligned}((\phi \circ \psi^{-1})^* \omega^e)_{(x,g)}(X, Y) =& \omega^e_{(x,h(x)g)}(X, R_{g*}(D_x h(X)) + L_{h(x)*} Y) \\ =& (\mathrm{Ad}_{g^{-1}} \circ \mathrm{Ad}_{h(x)^{-1}}) \circ (\xi(X)) + \mu_{SO^+}(R_{g*} D_x h(X) + L_{h(x)*} Y) \\ =& \mathrm{Ad}_{g^{-1}} \circ \left(\mathrm{Ad}_{h(x)^{-1}} \xi(X) + h^* \mu_{SO^+}(X)\right) + \mu_G(Y) \\ =& \mathrm{Ad}_{g^{-1}} \circ \tilde{\xi}(X) + \mu_G(Y) \\ =& \omega^{\tilde{e}}_{(x,g)}(X, Y)\end{aligned}$$

implying the claim. $\square$

Note that in the preceding proof, we never used the fact that the Levi-Civita connection is torsion free, so this result applies to any metric compatible covariant derivative on an oriented and time oriented pseudo-Riemannian manifold. If instead (M, g) is just a pseudo-Riemannian manifold, and not assumed to be orientable and time orientable, we could perform the same process to obtain the more general result that any metric compatible covariant derivative induces a principal connection on $O(M)$. This is clearly due to the fact that $\mathfrak{o}(t, s) \cong \mathfrak{so}^+(t, s)$. In fact, in greater generality, one can take any covariant derivative on an arbitrary vector bundle E, and find a unique connection one form in the bundle of linear frames of E.

Note that by **Proposition 2.2.17**, we have in any local oriented and time oriented orthonormal frame e:

$$(A^e_{SO^+})^c_a = \xi^c_a = \xi_{ab}\eta^{bc}$$

Including the inverse matrix η^{bc} will prove convenient for our purposes.

Proposition 2.2.20. *Let:*

$$\Lambda : Spin^+(M) \longrightarrow SO^+(M)$$

be the covering map given by the spin structure. Then:

$$A_{Spin^+} = (\lambda_*)^{-1} \circ \Lambda^*(A_{SO^+})$$

*is a connection one form on $Spin^+(M)$, called the **spin connection**.*

Proof. We need to show that this define a connection one form $\mathrm{Spin}^+(M)$. Let $g \in \mathrm{Spin}^+(t,s)$, then using the λ equivariance of the spin structure:

$$\Lambda \circ R_g = R_{\lambda(g)} \circ \Lambda$$

we see that since λ_* is a Lie algebra isomorphism $\mathfrak{spin}^+(t,s) \to \mathfrak{so}^+(t,s)$:

$$\begin{aligned} R^*_g A_{\mathrm{Spin}^+} =& (\lambda_*)^{-1} \circ R^*_g \Lambda^*(A_{SO^+}) \\ =& (\lambda_*)^{-1} \circ (\Lambda \circ R_g)^*(A_{SO^+}) \\ =& (\lambda_*)^{-1} \circ (R_{\lambda(g)} \circ \Lambda)^*(A_{SO^+}) \\ =& (\lambda_*)^{-1} \circ \Lambda^* R^*_{\lambda(g)}(A_{SO^+}) \\ =& (\lambda_*)^{-1} \circ \mathrm{Ad}_{\lambda(g)^{-1}} \circ \Lambda^*(A_{SO^+}) \end{aligned}$$

We introduce the identity transformation $\lambda_* \circ (\lambda_*)^{-1}$ and obtain that:

$$\begin{aligned} R^*_g A_{\mathrm{Spin}^+} =& (\lambda_*)^{-1} \circ \mathrm{Ad}_{\lambda(g)^{-1}} \circ \lambda_* \circ (\lambda_*)^{-1} \circ \Lambda^*(A_{SO^+}) \\ =& (\lambda_*)^{-1} \circ \mathrm{Ad}_{\lambda(g)^{-1}} \circ \lambda_* \circ A_{\mathrm{Spin}^+} \end{aligned}$$

Note that since λ is a group homomorphism:

$$\begin{aligned} \mathrm{Ad}_{\lambda(g)^{-1}} \circ \lambda_* =& (c_{\lambda(g)^{-1}} \circ \lambda)_* \\ =& (R_{\lambda(g)} \circ L_{\lambda(g)^{-1}} \circ \lambda)_* \\ =& (\lambda \circ R_g \circ L_{g^{-1}})_* \\ =& \lambda_* \circ (c^{-1}_g)_* \\ =& \lambda_* \circ \mathrm{Ad}_{g^{-1}} \end{aligned}$$

hence:

$$\begin{aligned} R^*_g A_{\mathrm{Spin}^+} =& (\lambda_*)^{-1} \circ \lambda_* \circ \mathrm{Ad}_{g^{-1}} \circ A_{\mathrm{Spin}^+} \\ =& \mathrm{Ad}_{g^{-1}} \circ A_{\mathrm{Spin}^+} \end{aligned}$$

as desired. Now let $X \in \mathfrak{spin}^+(t,s)$, and recall that for any $p \in \mathrm{Spin}^+(M)$, we have that the vertical vector field $\tilde{X}$ at p is given by:

$$\tilde{X}_p = \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tY)$$

With:

$$A_{\mathrm{Spin}^+}(\tilde{X}_p) = (\lambda_*)^{-1} \circ A_{SO^+}(\Lambda_* \tilde{X}_p)$$

we see that:

$$\begin{aligned}\Lambda_*\tilde{X}_p =& \frac{d}{dt}\Big|_{t=0}\Lambda\left(p\cdot\exp(tX)\right)\\ =& \frac{d}{dt}\Big|_{t=0}\Lambda(p)\cdot\exp(t\lambda_*(X))\\ =& \widetilde{\lambda_*(X)}_{\Lambda(p)}\end{aligned}$$

hence:

$$\begin{aligned}A_{\text{Spin}^+}(\tilde{X}_p) =& (\lambda_*)^{-1}\circ A_{SO^+}\left(\widetilde{\lambda_*(X)}_{\Lambda(p)}\right)\\ =& (\lambda_*)^{-1}\circ\lambda_*(X)\\ =& X\end{aligned}$$

implying the claim. □

Note that this implies that the Levi-Civita connection induces a connection A_{Spin^+} on $\text{Spin}^+(M)$. We will use this connection to define the spin covariant derivative.

Definition 2.2.33. We call the covariant derivative on the spinor bundle S induced by the spin connection the **spin covariant derivative**, which we also denote by ∇.

Given a section $\epsilon : U \to \text{Spin}^+(M)_U$, and a section $\psi : U \to \Delta_n$ we write sections of S locally as:

$$\Psi = [\epsilon, \psi]$$

The spin covariant derivative then acts as usual:

$$\nabla_X\Psi = [\epsilon, \nabla_X\psi]$$

where:

$$\nabla_X\psi = d\psi(X) + A_{\text{Spin}^+}(X)\cdot\psi$$

and $A_{\text{Spin}^+}(X)$ acts on ψ through the induced spinor representation $\kappa_* : \mathfrak{spin}^+(t,s) \to \text{End}(\Delta_n)$.

Lemma 2.2.23. *Write the components of a matrix $A \in \mathfrak{so}^+(t,s)$ as:*

$$A_a^c = w_{ab}\eta^{bc}$$

then the map:

$$\kappa_*\circ(\lambda_*)^{-1} : \mathfrak{so}^+(t,s) \longrightarrow End(\Delta_n)$$

is given by:

$$\kappa_*(\lambda_*)^{-1}(A) = \frac{1}{4}w_{ab}\gamma^{ab}$$

Proof. Recall from **Proposition 2.2.7** that for any orthonormal basis $\{e_i\}$ of $\mathbb{R}^{t,s}$ a basis for the Lie algebra is given by:

$$\mathfrak{spin}^+(t,s) = \{e_ie_j \in \text{Cl}(t,s) : 1 \le i < j \le n\}$$

and from **Proposition 2.2.8** that[31]:

$$\lambda_*(e_ie_j) = 2(\eta_{ii}T_j^i - \eta_{jj}T_i^j)$$

where if $\{e^i\}$ is the basis dual to $\{e_i\}$:

$$T_j^i = e^i\otimes e_j$$

[31] We used E_j^i in the **Proposition 2.2.8**, but to stay consistent with the notation of this section, we write T_j^i.

and there is no implied summation. Let $z \in \mathfrak{spin}^+(t,s)$, then:

$$z = \sum_{a<c} z^{ac} e_a e_c$$

for some real coefficients z^{ac}, so:

$$\begin{aligned}\lambda_*(z) =& \sum_{a<c} z^{ac}\lambda_*(e_a e_c)\\ =& 2\sum_{a<c} z^{ac}\left(\eta_{aa}T^a_c - \eta_{cc}T^c_a\right)\end{aligned}$$

Note that for any k we have that:

$$(\eta_{aa}T^a_c)e_k = \eta_{aa}e^a(e_k)\otimes e_c = \delta^a_k\eta_{aa}e_c$$

which is only non zero when $k = a$. This implies that:

$$\eta(\lambda_*(e_a e_c)e_a, e_c) = 2\eta_{aa}\eta(e_c, e_c) = 2\eta_{aa}\eta_{cc}$$

We deduce that:

$$\eta(\lambda_*(e_a e_c)e_a, e_c)\eta_{aa}\eta_{cc}e_a e_c = 2e_a e_c$$

It follows that:

$$z = \frac{1}{2}\sum_{k<l}\eta(\lambda_*(z)e_k, e_l)\eta_{kk}\eta_{ll}e_k e_l$$

We now write A as:

$$A = w_{ab}\eta^{bc}T^a_c$$

Since A is in the image of λ_*, we can find $(\lambda_*)^{-1}(A)$ by:

$$(\lambda_*)^{-1}(A) =\frac{1}{2}\sum_{k<l}\eta(Ae_k, e_l)\eta_{kk}\eta_{ll}e_k e_l$$

We see that:

$$\begin{aligned}Ae_k =& w_{ab}\eta^{bc}e^a(e_k)\otimes e_c\\ =& w_{kb}\eta^{bc}e_c\end{aligned}$$

hence:

$$\begin{aligned}(\lambda_*)^{-1}(A) =& \frac{1}{2}\sum_{k<l} w_{kb}\eta^{bc}\eta(e_c, e_l)\eta_{kk}\eta_{ll}e_k e_l\\ =& \frac{1}{2}\sum_{k<l} w_{kb}\eta^{bc}\eta_{cl}\eta_{kk}\eta_{ll}e_k e_l\\ =& \frac{1}{2}\sum_{k<l} w_{kl}\eta_{kk}\eta_{ll}e_k e_l\end{aligned}$$

If we sum over all possible k and l we over count by a factor of 2 so:

$$(\lambda_*)^{-1}(A) =\frac{1}{4}\sum_{k,l} w_{kl}\eta_{kk}\eta_{ll}e_k e_l$$

Furthermore, we see that $e_k e_l = -e_l e_k$, so we can obtain an other factor of 2 and write:

$$(\lambda_*)^{-1}(A) =\frac{1}{8}\sum_{k,l} w_{kl}[\eta_{kk}e_k, \eta_{kk}e_l]$$

Note that κ is just the restriction of the spinor representation of $\mathrm{Cl}(t,s)$ to $\mathrm{Spin}^+(t,s)$, so if we have mathematical gamma matrices γ_a, then:

$$\kappa_*([e_k, e_l]) = [\gamma_k, \gamma_l]$$

hence:

$$\kappa_* \circ (\lambda_*)^{-1}(A) = \frac{1}{8}\sum_{k,l} w_{kl}[\eta_{kk}\gamma_k, \eta_l\gamma_l]$$

Since η is diagonal it follows that $\eta_{kk} = \eta^{ll}$, and that the sum:

$$\gamma^a = \eta^{ab}\gamma_b$$

only nonzero $a = b$, hence:

$$\begin{aligned}\kappa_* \circ (\lambda_*)^{-1}(A) =& \frac{1}{8}\sum_{k,l} w_{kl}[\eta^{kk}\gamma_k, \eta^{ll}\gamma_l] \\ =& \frac{1}{8}\sum_{k,l} w_{kl}[\gamma^k, \gamma^l] \\ =& \frac{1}{4}w_{kl}\gamma^{kl}\end{aligned}$$

as desired. □

From this lemma we obtain the following result:

Proposition 2.2.21. *The spin covariant derivative is locally given by:*

$$\begin{aligned}\nabla_X\psi =& d\psi(X) + \frac{1}{4}(X)\xi_{ab}\gamma^{ab}\psi \\ =& d\psi(X) - \frac{1}{4}(X)\xi_{ab}\Gamma^{ab}\psi\end{aligned}$$

Proof. Let ϵ be a section which under Λ maps to the oriented and time oriented orthonormal frame e. We have that in this frame:

$$(A^e_{SO^+})^c_a = \xi_{ab}\eta^{bc}$$

Then it follows that:

$$\begin{aligned}A^\epsilon_{\mathrm{Spin}^+} =& \epsilon^* A_{\mathrm{Spin}^+} \\ =& \epsilon^*\left((\lambda_*)^{-1} \circ \Lambda^*(A_{SO^+})\right) \\ =& (\lambda_*)^{-1} \circ (\Lambda \circ \epsilon)^* A_{SO^+} \\ =& (\lambda_*)^{-1} \circ e^* A_{SO^+}\end{aligned}$$

In any local orthonormal frame $A^e_{SO^+}$ is given by the one forms:

$$(A^e_{SO^+}(X))^c_a = \xi_{ab}(X)\eta^{bc}$$

or any $X \in \mathfrak{X}(M)$. It follows from **Lemma 2.2.23** that:

$$\kappa_* \circ (\lambda_*)^{-1} \circ (A^e_{SO^+})(X) = \frac{1}{4}\xi_{ab}(X)\gamma^{ab}$$

Therefore, for some smooth map $\psi : U \to \Delta_n$, since A_{Spin^+} acts on ψ through κ_*, we have:

$$\begin{aligned}\nabla_X\psi =& d\psi(X) + A^\epsilon_{\mathrm{Spin}^+}(X)\psi \\ =& d\psi(X) + \frac{1}{4}\xi_{ab}(X)\gamma^{ab}\psi\end{aligned}$$

Furthermore, since:

$$\gamma^{ab} = -\Gamma^{ab}$$

we obtain:

$$\nabla_X \psi = d\psi(X) - \frac{1}{4}\xi_{ab}(X)\Gamma^{ab}\psi$$

□

From this proposition we obtain the following result:

Corollary 2.2.14. *If the dimension, n, of M is even, then the spin covariant derivative preserves the splitting of the spinor bundle S into the Weyl spinor bundles $S^{\pm}$. This means that if $\Psi \in S^{\pm}$ then:*

$$\nabla_X \Psi \in \Gamma(S^{\pm})$$

for all $X \in \mathfrak{X}(M)$

Proof. Recall from **Proposition 2.2.13**that there exists a global bundle automorphism ω on S which preserves the subbundles $S^{\pm}$. It follows that locally, given a section $\epsilon : U \to \mathrm{Spin}^+(M)$, this global bundle automorphism is given by:

$$\omega \cdot \Psi = [\epsilon, \Gamma^{n+1} \cdot \psi]$$

for some smooth map $\psi : U \to \Delta_n$, and where:

$$\Gamma^{n+1} = -\, i^{n/2+1}\Gamma^1 \cdots \Gamma^{n+1}$$

From **Lemma 2.2.10** we obtain that:

$$[\Gamma^{n+1}, \Gamma^{ab}] = [\Gamma^{n+1}, \Gamma^a\Gamma^b] - [\Gamma^{n+1}, \Gamma^b\Gamma^a] = 0$$

Therefore, if $\Psi \in \Gamma(S^{\pm})$:

$$\omega \cdot \nabla_X \Psi = [\epsilon, \Gamma^{n+1} d\psi(X) - \frac{1}{4}\xi_{ab}(X)\Gamma^{n+1}\Gamma^{ab}\Psi$$

We see that clearly $[\epsilon, d\psi(X)] \in S^{\pm}$, so since Γ^{n+1} preserves the subspaces $\Delta_n^{\pm}$, we have that $[\epsilon, \Gamma^{n+1} d\psi(X)] \in S^{\pm}$. Furthermore, since $[\Gamma^{n+1}, \Gamma^{ab}] = 0$, we have that:

$$\Gamma^{n+1}\Gamma^{ab}\psi = \pm\, \Gamma^{ab}\psi$$

hence:

$$[\epsilon, \Gamma^{n+1}\Gamma^{ab}\psi] = [\epsilon, \pm\Gamma^{ab}\psi]$$

implying the claim. □

We also have that Clifford multiplication on the level of bundles is compatible with the spin covariant derivative and Levi-Civita connection in the following sense:

Proposition 2.2.22. *For all vector fields $X, Y \in \mathfrak{X}(M)$, and all $\Psi \in \Gamma(S)$ we have that:*

$$\nabla_X (Y \cdot \Psi) = (\nabla_X Y) \cdot \Psi + Y \cdot (\nabla_X \Psi)$$

where $\nabla_X Y$ denotes the Levi-Civita connection, and $\nabla_X \Psi$ denotes the spin covariant derivative.

Proof. We need only prove this locally. We have that for some smooth maps $\phi : U \to \mathbb{R}^{t,s}$, and $\psi : U \to \Delta_n$, and a section $\epsilon : U \to \mathrm{Spin}^+(M)_U$ that:

$$Y = [\Lambda(\epsilon), \phi] \qquad \text{and} \qquad \Psi = [\epsilon, \phi]$$

It follows from **Proposition 2.2.13** that:

$$Y \cdot \Psi = [\epsilon, \phi \cdot \psi]$$

It follows that:

$$\nabla_X(Y \cdot \Psi) =[\epsilon, (d\phi(X)) \cdot \psi + \phi \cdot (d\psi(X)) + A^{\epsilon}_{\mathrm{Spin}^+}(X)(\phi \cdot \psi)]$$

We have that $A^{\epsilon}_{\mathrm{Spin}^+}(X)$ acts on Δ_n through the representation κ_*. We see that by :

$$\kappa(g)(\phi \cdot \psi) = (\lambda(g) \cdot \phi)(\kappa(g) \cdot \psi)$$

Taking the derivative at $e \in \mathrm{Spin}^+(t, s)$ we obtain that for $z \in \mathfrak{spin}^+(t, s)$:

$$\kappa_*(z)(\phi \cdot \psi) = (\lambda_*(z) \cdot \phi) \cdot \psi + \phi \cdot \kappa_*(z) \cdot \psi$$

Since:

$$\lambda_*(A^{\epsilon}_{\mathrm{Spin}^+}) = A^{e}_{SO^+}$$

where $e = \Lambda \circ \epsilon$, we thus have that:

$$\begin{aligned}
\nabla_X(Y \cdot \Psi) =&[\epsilon, (d\phi(X)) \cdot \psi + \phi \cdot d\psi(X) + (A^{e}_{SO^+}(X)\phi) \cdot \psi + \phi \cdot A^{\epsilon}_{\mathrm{Spin}^+}(X)\psi]\\
=&[\epsilon, d\phi(X)) \cdot \psi + (A^{e}_{SO^+}(X)\phi) \cdot \psi] + [\epsilon, \phi \cdot d\psi(X) + \phi \cdot A^{\epsilon}_{\mathrm{Spin}^+}(X)\psi]\\
=&(\nabla_X Y) \cdot \Psi + Y \cdot (\nabla_X \Psi)
\end{aligned}$$

implying the claim. □

By **Proposition 2.1.23** we clearly have the following result:

Corollary 2.2.15. *For any spinor bundle equipped with a Dirac bundle metric, a connection in $A \in \Omega^1(Spin^+(M), \mathfrak{spin}^+(t, s))$ induces a metric compatible covariant derivative. In particular, the spin connection, induced by the Levi-Civita connection, induces a metric compatible covariant derivative on the spinor bundle.*

We now define the Dirac operator in terms of the spin covariant derivative.

Definition 2.2.34. The **Dirac operator**, denoted $D : \Gamma(S) \to \Gamma(S)$ is given in a local oriented orthonormal frame by:

$$D\Psi = \eta^{ab} e_a \cdot \nabla_{e_b} \Psi$$

If we let $\Psi = [\epsilon, \psi]$, for some $\epsilon : U \to \mathrm{Spin}^+(M)_U$ satisfying $\Lambda \circ \epsilon = e$, we have that:

$$\begin{aligned}
D\Psi =&[\epsilon, D\psi]\\
=&[\epsilon, \gamma^a \nabla_{e_a} \psi]\\
=&\left[\epsilon, \gamma^a \left(d\psi(e_a) + \frac{1}{4}\xi_{bc}(e_a)\gamma^{bc}\psi\right)\right]\\
=&\left[\epsilon, i\Gamma^a \left(d\psi(e_a) - \frac{1}{4}\xi_{bc}(e_a)\Gamma^{bc}\psi\right)\right]
\end{aligned}$$

Proposition 2.2.23. *The Dirac operator is independent of the local oriented and time oriented orthonormal frame $\{e_i\}$.*

Proof. Let $\{f_i\}$ be another local oriented and time oriented orthonormal frame, then:

$$f_i = B^j_i e_j$$

for some matrix of functions B^j_i valued in $SO^+(t, s)$. It follows that:

$$\begin{aligned}
D\Psi =&\eta^{ab} f_a \cdot \nabla_{f_b} \Psi\\
=&\eta^{ab}(B^j_a e_j) \cdot \nabla_{B^k_b e_k} \Psi\\
=&\eta^{ab}(B^j_a e_j) B^k_b \cdot \nabla_{e_k} \Psi
\end{aligned}$$

Note that since $\eta = \eta^{-1}$:

$$B^T \eta^{-1} B = \eta^{-1}$$

we have:

$$\begin{aligned} D\Psi =& (B^j_a \eta^{ab} B^k_b) e_j \cdot \nabla_{e_k} \Psi \\ =& \eta^{jk} e_j \cdot \nabla_{e_k} \Psi \end{aligned}$$

implying the claim. □

This implies that the Dirac operator D is indeed well defined. Furthermore, the Dirac operator is a first order differential operator on the sections of S. Clearly, if the dimension of M is even, since Clifford multiplication of a vector maps $\Delta_n^\pm$ to $\Delta_n^\mp$, and since ∇ preserves the subbundles $S^\pm$, we obtain the following corollary:

Corollary 2.2.16. *If the dimension of M is even, then the Dirac operator is a map:*

$$D : \Gamma(S^\pm) \longrightarrow \Gamma(S^\mp)$$

Example 2.2.8. Let $(M, g) = (\mathbb{R}^{t,s}, \eta)$, then in global coordinates x^i, Levi-Civita connection is given by

$$\nabla_X Y = dY(X) = X^i \partial_i Y^j \partial_j$$

as the metric is constant. One can directly verify this by use of **Proposition 2.2.18**, and the fact the coordinate vector fields are orthonormal, and commute. It follows that on $S = \mathbb{R}^{t,s} \times \Delta_{t+s}$, that we can define a Dirac operator in any global orthonormal frame $\{e_i\}$ by:

$$D = \gamma^a \cdot \nabla_{e_a}$$

Note that if we just the standard basis, then the coordinate frame is orthonormal normal, so setting $e_i = \partial_i$ gives the following for any spinor field $\Psi : \mathbb{R}^{t,s} \to \mathbb{R}^{t,s} \times \Delta_{t+s}$:

$$\begin{aligned} D\Psi =& \gamma^a \cdot \nabla_{\partial_a} \Psi \\ =& \gamma^a \cdot \frac{\partial \Psi}{\partial x^a} \end{aligned}$$

It follows that:

$$\begin{aligned} D^2\Psi =& \gamma^b \cdot \nabla_{e_b} \left(\gamma^a \cdot \frac{\partial \Psi}{\partial x^a} \right) \\ =& \gamma^b \gamma^a \frac{\partial^2 \Psi}{\partial x^b x^a} \\ =& \frac{1}{2} \sum_{a,b} \{\gamma^b, \gamma^a\} \frac{\partial^2 \Psi}{\partial x^b x^a} \end{aligned}$$

Note that:

$$\begin{aligned} \{\gamma^b, \gamma^a\} =& \eta^{ca} \eta^{db} \{\gamma_c, \gamma_d\} \\ =& -2\eta^{ca} \eta^{db} \eta cd I_N \\ =& -2\eta^{ca} \delta^b_c \\ =& -2\eta^{ab} \end{aligned}$$

hence, with the fact $\eta^{ab} = \eta_{ab} = 0$ unless $a = b$ obtain:

$$\begin{aligned} D^2\Psi =& -\sum_{a,b} \eta^{ba} \frac{\partial^2 \Psi}{\partial x^b x^a} \\ =& -\sum_a \eta_{bb} \frac{\partial^2 \Psi}{\partial (x^b)^2} \end{aligned}$$

implying that D^2 is the Laplacian, so D is the ”square root” of the Laplacian as desired.

We now wish to extend the spin covariant derivative, and Dirac operator to sections of twisted spinor bundles, i.e. $S \otimes E$, where E is a vector bundle associated to some principal G bundle $\pi : P \to M$. We first need the following lemmas:

Lemma 2.2.24. *Let $\kappa \otimes \rho$ be the representation of Spin$^+(t,s) \times G$ on the vector space $\Delta_n \otimes V$, given on simple tensors by:*

$$\kappa \otimes \rho(s,g)(\psi \otimes v) = \kappa(s)\psi \otimes \rho(g)v$$

for all $s, g \in$ Spin$^+(t,s) \times G$, and $v \in V$, $\psi \in \Delta_n$. Then the induced Lie algebra representation is given on simple tensors by:

$$(\kappa \otimes \rho)_*(X,Y)(\psi \otimes v) = \kappa_*(X)\psi \otimes v + \psi \otimes \rho_*(Y)v$$

for all $(X,Y) \in \mathfrak{spin}^+(t,s) \oplus \mathfrak{g}$.

Proof. This follows trivially by taking the derivative at $t = 0$ as follows:

$$\begin{aligned}(\kappa \otimes \rho)_*(X,Y)(\psi \otimes v) =& \frac{d}{dt}\Big|_{t=0} \kappa \otimes \rho(\exp(tX), \exp(tY))(\psi \otimes v) \\ =& \frac{d}{dt}\Big|_{t=0} \kappa(\exp(tX))\psi \otimes \rho(\exp(tY))(v) \\ =& \kappa_*(X)\psi \otimes v + \psi \otimes \rho_*(Y)v\end{aligned}$$

$\square$

Lemma 2.2.25. *If A is a connection one form P, and A_{Spin^+}[32] is a connection one form on Spin$^+(M)$, then the following defines a connection one form on Spin$^+(M) \times_M P$:*

$$(A_{Spin^+(M)} \oplus A)(X_p, Y_q) = (A_{Spin^+}(X_p), A(Y_q)) \in \mathfrak{spin}^+(t,s) \oplus \mathfrak{g}$$

for all $(X_p, Y_q) \in T_p$Spin$^+(M) \times T_qP$, such that $\pi_{Spin^+}(p) = \pi(q)$, and $\pi_{Spin^+*}(X_p) = \pi_*(Y_q)$.

Proof. The fact that $A_{\mathrm{Spin}^+(M)} \oplus A$ is an element of $\Omega^1(\mathrm{Spin}^+(M) \times_M P, \mathfrak{spin}^+(t,s) \oplus \mathfrak{g})$ is clear. We thus need to check that this indeed a connection one form. Let $(s,g) \in \mathrm{Spin}^+(t,s) \times G$, and recall that:

$$R_{(s,g)} \circ (p,q) = (p \cdot s, q \cdot g)$$

It follows that for $(X,Y)_{(p,q)} \in T_p\mathrm{Spin}^+(M) \times T_qP$, where $(p,q) \in \mathrm{Spin}^+(M) \times_M P$:

$$R_{(s,g)*}(X_p, Y_q) = (R_{s*}X_p, R_{g*}Y_q)$$

Therefore:

$$\begin{aligned}(R^*_{(s,g)}(A_{\mathrm{Spin}^+(M)} \oplus A))_{(p,q)}(X_p, Y_q) =& (A_{\mathrm{Spin}^+(M)} \oplus A)_{(p\cdot s, q\cdot g)}(R_{s*}X_p, R_{g*}Y_q) \\ =& (A_{\mathrm{Spin}^+ p\cdot s}(R_{s*}X_p), A_{q\cdot s}(R_{g*}Y_q)) \\ =& (\mathrm{Ad}_{s^{-1}} \circ A_{\mathrm{Spin}^+ p}(X_p), \mathrm{Ad}_{g^{-1}} \circ A_q(Y_q)) \\ =& c_{(s,g)^{-1}*} \circ (A_{\mathrm{Spin}^+ p}(X_p), A_q(Y_q)) \\ =& \mathrm{Ad}_{(s,g)^{-1}} \circ (A_{\mathrm{Spin}^+(M)} \oplus A)(X_p, Y_q)\end{aligned}$$

so $(A_{\mathrm{Spin}^+(M)} \oplus A)$ is Ad invariant. Furthermore, we see that for any $(X,Y) \in \mathfrak{spin}^+(t,s) \oplus \mathfrak{g}$, the fundamental vector filed $\tilde{Z}$ associated to this Lie algebra element is defined by:

$$\begin{aligned}\tilde{Z}_{(p,q)} =& \frac{d}{dt}\Big|_{t=0} (p,q) \cdot (\exp(tX), \exp(tY)) \\ =& (\tilde{X}_p, \tilde{Y}_q)\end{aligned}$$

hence:

$$\begin{aligned}(A_{\mathrm{Spin}^+(M)} \oplus A)_{(p,q)}(\tilde{Z}) =& (A_{\mathrm{Spin}^+ p}(\tilde{X}_p), A_q(\tilde{Y}_q)) \\ =& (X,Y)\end{aligned}$$

implying that $(A_{\mathrm{Spin}^+(M)} \oplus A)$ is a connection one form on $\mathrm{Spin}^+(M) \times_M P$ $\square$

[32]Not necessarily the spin connection induced by the Levi-Civita connection

We can show the following:

Proposition 2.2.24. *The connection one form $(A_{Spin^+(M)} \oplus A)$ induces a covariant derivative on $S \otimes E$ defined in the usual manner:*

$$\nabla^A_X \Psi = [\epsilon \times_M s, \nabla^A_X \psi] \tag{2.2.22}$$

*where $\epsilon \times_M s$ is a local section $U \to Spin^+(M) \times_M P_U$, and $\psi : U \to \Delta_n \otimes V$ is a smooth map as given in **Proposition 2.2.16**. Furthermore,*

$$\nabla_X \psi = d\psi(X) - \frac{1}{4}\xi_{ab}(X)\Gamma^{ab}\psi + \rho_*(A_s(X))\psi \tag{2.2.23}$$

Here ξ_{ab} are the one forms defining the Levi-Civita connection, and Γ^{ab} acts on the spinor components of ψ, and $\rho_*(A_s)(X)$ acts on the vector part of ψ, i.e mixes the multiplet ψ.

Proof. The fact that (2.2.22) is a covariant derivative, and independent of the sections ϵ and s is evident from **Lemma 2.2.25**. In particular, we could make a gauge transformation on either component ϵ, or s, and obtain a gauge invariant quantity, as the group action is a product group action.

To pull back $(A_{\mathfrak{spin}^+} \oplus A)$ by $\epsilon \times_M s$, we first note that:

$$\epsilon \times_M s(x) = (\epsilon(x), s(x))$$

hence for any vector $X \in T_xM$:

$$\begin{aligned}(\epsilon \times_M s)^*(A_{\mathfrak{spin}^+} \oplus A)_x(X) =& (A_{\text{Spin}^+}(\epsilon_* X), A(s_* X)) \\ =& (A^{\epsilon}_{\text{Spin}^+}(X), A_s(X))\end{aligned}$$

Now let $\{v_i\}$ be any basis for V, and decompose ψ as:

$$\psi = \psi^i \otimes v_i$$

We see that (2.2.24) then follows from **Lemma 2.2.24** as follows:

$$\begin{aligned}\nabla^A_X \psi =& d\psi(X) + (\kappa \otimes \rho)_*(A^{\epsilon}_{\text{Spin}^+}(X), A_s(X))(\psi^i \otimes v_i) \\ =& d\psi(X) + \kappa_*(A^{\epsilon}_{\text{Spin}^+}(X))\psi^i \otimes v_i + \psi^i \otimes \rho_*(A_s(X))v_i \\ =& d\psi(X) - \xi_{ab}(X)\Gamma^{ab}\psi^i \otimes v_i + \psi^i \otimes \rho_*(A_s(X))v_i\end{aligned}$$

implying the claim. □

Definition 2.2.35. The covariant derivative on $S \otimes E$ induced by the connection $A_{\text{Spin}^+} \otimes A$, where A_{Spin^+} is the spin connection, and A is any connection on P is called the **twisted spin covariant derivative.**

Note that for a fixed vector field X, ∇^A_X is still a map $\Gamma(S \otimes E) \to \Gamma(S \otimes E)$, hence if we can perform Clifford multiplication on $S \otimes E$, then we can define a Dirac operator on $\Gamma(S \otimes E)$ in a similar manner to **Definition 2.2.34**.

Lemma 2.2.26. *Let $S \otimes E = (Spin^+(M) \times_M P) \times_{\kappa \otimes \rho} (\Delta_n \otimes V)$, then there exists a well defined bilinear Clifford multiplication:*

$$\begin{aligned}TM \times (S \otimes E) &\longrightarrow S \otimes E \\ (X, \Psi) &\longmapsto X \cdot \Psi\end{aligned}$$

on the level of bundles, which restricts to a bilinear map $T_xM \times (S \otimes E)_x \to (S \otimes E)_x$ for all $x \in M$. This map also induces a well-defined Clifford multiplication of forms with twisted spinors.

Proof. We define the map on simple tensors:

$$\begin{aligned}(SO^+(M) \times_{\rho_{SO^+}} \mathbb{R}^{t,s}) \times (S \otimes E) &\longrightarrow S \otimes E \\ ([\Lambda(p), x], [(p, q), \psi \otimes v]) &\longmapsto [(p, q), (x \cdot \psi) \otimes v]\end{aligned}$$

and extend linearly. It follows from the same argument in **Proposition 2.2.13** that this map is well defined, as it is only acting on the spinor part of an element in $S \otimes E$. Furthermore, the a similar construction as in **Proposition 2.2.13** yields a well defined Clifford multiplication with twisted forms. □

With this we obtain the following corollary:

Corollary 2.2.17. *The map* $D_A : \Gamma(S \otimes E) \to \Gamma(S \otimes E)$ *given in any orthonormal frame by:*

$$D_A \Psi = \eta^{ab} e_a \cdot \nabla^A_{e_b} \Psi$$

where $\Psi \in \Gamma(S \otimes E)$ *is well defined, and independent of the local frame* e_a.

Proof. This follows from **Lemma 2.2.26** and the same argument as applied in **Proposition 2.2.23**. ☐

Due to the importance of this map, we give the following definition:

Definition 2.2.36. The map $D_A : \Gamma(S \otimes E) \to \Gamma(S \otimes E)$ is called the **twisted Dirac operator**, and is locally given by:

$$D_A \Psi = [\epsilon \times_M s, D_A \psi]$$

where:

$$\begin{aligned} D_A \psi =& \gamma^a \left(d\psi(e_a) + \frac{1}{4} \xi_{bc}(e_a) \gamma^{bc} \psi + \rho_*(A_s(e_a)) \psi \right) \\ =& i\Gamma^a \left(d\psi(e_a) - \frac{1}{4} \xi_{bc}(e_a) \Gamma^{bc} \psi + \rho_*(A_s(e_a)) \psi \right) \end{aligned}$$

As we shall see in the section on QED, the twisted Dirac operator will play an important role in defining the Yang-Mills-Dirac Lagrangian, by giving us a way to couple the gauge field to fermion fields. Physically, our connection will play the role of the electromagnetic potential, and we want the dynamics of our fermionic fields to be determined by this potential. Without the coupling given by the twisted Dirac operator, the dynamics of the fermionic fields would be like that of a free particle, hence the importance. Mathematically, the Dirac operator gives rise to a large number of deep results in geometric analysis. To the interested reader, we recommend the following texts: Friedrich's *Dirac Operators in Riemannian Geometry*, Jost's *Riemannian Geometry and Geometric Analysis*, and Michelsohn and Lawson's *Spin Geometry*.

Yang-Mills and Applications to Physics

3.1 Yang-Mills and Electromagnetism

In **Example 2.1.15** we demonstrated how a connection one form on $\mathbb{R}^4 \times U(1)$ satisfies the gauge transformation rules for electromagnetism. However, how does one choose a connection such that the dynamics of electromagnetism are obtained? The purpose of this chapter is to explore this question in detail. As we will see, at least in the case of source free electromagnetism, the answer is given by finding a connection which leaves the Yang-Mills action *stationary.* In other words, we will want the connection to be a critical point of the Yang-Mills action.

That being said, there is still some legwork which must be done in order to get to this point. We first need to be able to write the Yang-Mills action down; this action involves the L^2 inner product on $\Omega^2(M, \mathrm{Ad}(P))$ induced by a (pseudo)-Riemannian metric on the base manifold, and thus unavoidably relies on the chosen geometry of M. Furthermore, in order to find the aforementioned stationary points, we will need to develop a formal adjoint to the covariant derivative with respect to this L^2 inner product, called the *covariant codifferential.* We spend the first part of this section constructing these two operations.

Once, we are able to write down the Yang-Mills action, we will spend the remainder of the section deriving the Yang-Mills equation, and modifying the action so that the stationary points yield a complete theory of classical electromagnetism which incorporates sources. We will also see how to obtain the classical field equations for quantum electrodynamics, which will involve the fermionic matter fields for electrons and positrons, as opposed to a classical charge density.

We continue to draw inspiration from Hamilton's Mathematical Gauge Theory, though some of our conventions may differ.

3.1.1 The Hodge Star Operator and the Codifferential

Let (M, g) be a (pseudo)-Riemannian manifold. Recall from **Proposition 1.1.17** that the (pseudo)-Riemannian metric g induces a bundle isomorphism $\alpha : TM \to T^*M$. We define an inner product[33] on T_p^*M for all $p \in M$ by:

$$\langle \omega, \eta \rangle_p = g_p \left(\alpha^{-1}(\omega), \alpha^{-1}(\eta) \right) \tag{3.1.1}$$

Let x^i be a coordinate system, and let $\omega = \omega_i dx^i$, and $\eta = \eta_j dx^j$, then:

$$\begin{aligned} \langle \omega, \eta \rangle =& g_{ij}(g^{ik}\omega_k, g^{jl}\eta_l) \\ =& \delta^k_j \omega_k g^{jl} \eta_l \\ =& \omega_j \eta^j \end{aligned}$$

so the inner product on T_p^*M is just the contraction of of ω with $\alpha^{-1}(\eta)$. Furthermore, it is clear from (3.1.1) that $\langle \cdot, \cdot \rangle$ is symmetric, smooth and non degenerate. If g is positive definite, then $\langle \cdot, \cdot \rangle$ is positive definite, and thus defines a Euclidean bundle metric on T^*M. If g is pseudo Riemannian, and M is connected, then $\langle \cdot, \cdot \rangle$ defines a pseudo Euclidean bundle metric on T^*M of the same signature as g.

Proposition 3.1.1. *Let (M, g) be a (pseudo)-Riemannian manifold. Then there exists a natural (pseudo)-Euclidean bundle metric on the vector bundle $T^{(0,k)}M$. If M is connected, and g is (pseudo)-Riemannian, then the signature of the bundle metric is well defined.*

Proof. For all $p \in M$, we have that any element of $T_p^{(0,k)}M$ can be written as a finite sum of simple tensors. That is, for $\omega \in T_p^{(0,k)}M$ we have that[34]:

$$\omega = \omega_I \alpha^{i_1} \otimes \cdots \otimes \alpha^{i_k}$$

where each $\alpha_{i_l} \in T_p^*M$. Hence, for all $p \in M$ we define an inner product on simple tensors by:

$$\langle \alpha^1 \otimes \cdots \otimes \alpha^k, \beta^1 \otimes \cdots \otimes \beta^k \rangle_p^k = \langle \alpha^1, \beta^1 \rangle_p \cdots \langle \alpha^1, \beta^1 \rangle_p$$

[33] We are being mildly loose with our definition of an inner product. In this section, an inner product is just a symmetric, bilinear, non- degenerate, scalar valued map.

[34] Since these are general not alternating tensors, the sum $\omega_I \alpha^I$ is over all possible multi indexes, not just ordered ones.

which we then extend to $T_p^{(0,k)}M$ bi-linearly by

$$\langle \omega_I \alpha^I, \eta_J \beta^J \rangle_p^k = \omega_I \eta_J \langle \alpha^I, \beta^J \rangle_p^k$$

where α^I, β^I are indexed sets of simple tensors in $T_p^{(0,k)}M$. It then follows that for all $p \in M$, $\langle \cdot, \cdot \rangle_p^k$ is symmetric, and bilinear. It suffices to check that the inner product is non degenerate on simple tensors since they generate the space. Fix a set of k covectors α^i, then if:

$$\langle \alpha^1 \otimes \cdots \otimes \alpha^k, \beta^1 \otimes \cdots \otimes \beta^k \rangle_p^k = 0$$

for all simple tensors $\beta^1 \otimes \cdots \otimes \beta^k \in T_p^{(0,k)}M$ it follows that:

$$\langle \alpha^1, \beta^1 \rangle_p \cdots \langle \alpha^k, \beta^k \rangle_p = 0 \tag{3.1.2}$$

for all $\beta^i \in T_p^*M$. Since the induced inner product on T_p^*M is non degenerate, fix β^i for $i < k$ such that $\langle \alpha^i, \beta^i \rangle \neq 0$, then we see that (3.1.2) is equivalent to:

$$\langle \alpha^k, \beta^k \rangle_p = 0$$

for all $\beta^k \in T_p^*M$, a contradiction so $\langle \cdot, \cdot \rangle_p^k$ is nondegenerate. If g is Riemannian, it is clear that $\langle \cdot, \cdot \rangle_p^k$ is positive definite.

We have constructed an inner product on each fibre of $T^{(0,k)}M$; to prove that $\langle \cdot, \cdot \rangle^k$ is a bundle metric, we need to show it is smooth. Let x^i be a coordinate system for an open set $U \subset M$, then for the tensor fields ω and η we see that:

$$\begin{aligned}\langle \omega, \eta \rangle^k =& \omega_{i_1 \cdots i_k} \eta_{j_1 \cdots j_k} \langle dx^{i_1}, dx^{j_1} \rangle \cdots \langle dx^{i_k}, dx^{j_k} \rangle \\ =& \omega_{i_1 \cdots i_k} \eta_{j_1 \cdots j_k} g^{i_1 j_1} \cdots g^{i_k j_k}\end{aligned} \tag{3.1.3}$$

which is smooth as it is the product of smooth functions. Importantly, in the line above, we are summing over all possible multi indices $i_1, \ldots i_k$, not just ordered ones, as these are general covariant tensors so there are no restrictions on the coefficients $\omega_{i_1 \cdots i_k}$.

If (M, g) is connected and pseudo Riemannian it follows that for all $p \in M$ $\langle \cdot, \cdot \rangle_p^k$ has a well defined signature. The signature of $\langle \cdot, \cdot \rangle^k$ is then the same as the signature of $\langle \cdot, \cdot \rangle_p^k$, and is independent of the choice of p by a similar argument to **Theorem 1.1.11** □

Note that this inner product also induces a bundle isomorphism $T^{(0,k)}M \to T^{(k,0)}M$. Explicitly, in a local co-frame dx^i, the tensor field:

$$\omega = \omega_{i_1 \cdots i_k} dx^{i_1} \otimes \cdots \otimes dx^{i_k}$$

is mapped to the tensor field:

$$\omega^\sharp = \omega_{i_1 \cdots i_k} g^{i_1 j_1} \cdots g^{i_k j_k} \partial_{j_1} \otimes \cdots \otimes \partial_{j_k}$$

which for brevity we denote by:

$$\omega^\sharp = \omega^{j_1 \cdots j_k} \partial_{j_1} \otimes \cdots \otimes \partial_{j_k}$$

We see that with this notation (3.1.3) can be rewritten as:

$$\langle \omega, \eta \rangle^k = \omega_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k}$$

As we shall see shortly, the restriction of $\langle \cdot, \cdot \rangle^k$ to the vector subbundle $\Lambda^k(TM)$ defines a bundle metric on $\Lambda^k(TM)$. In particular, one should check that such a restriction is nondegenerate, as $\langle \cdot, \cdot \rangle^k$ is pseudo Euclidean. Indeed, the restriction of $\langle \cdot, \cdot \rangle^k$ induces a symmetric bilinear map:

$$\begin{aligned}\Omega^k(M) \times \Omega^k(M) &\longrightarrow C^\infty(M) \\ (\omega, \eta) &\longmapsto \langle \omega, \eta \rangle^k\end{aligned}$$

In coordinates:

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) dx^{\sigma(i_1)} \otimes \cdots \otimes dx^{\sigma(i_k)}$$

hence the inner product is given by:

$$\langle dx^{i_1} \wedge \cdots \wedge dx^{i_k}, dx^{j_1} \wedge \cdots \wedge dx^{j_k} \rangle^k = \sum_{\sigma \in S_k} \sum_{\alpha \in S_k} \operatorname{sgn}(\sigma)\operatorname{sgn}(\alpha) g^{\sigma(i_1)\alpha(j_k)} \cdots g^{\sigma(i_k)\alpha(j_k)}$$

So, for arbitrary k forms:

$$\omega = \omega_I dx^I \qquad \text{and} \qquad \eta = \eta_J dx^J$$

we have that:

$$\langle \omega, \eta \rangle^k = \sum_{\sigma, \alpha \in S_k} \omega_{\sigma(i_1) \cdots \sigma(i_k)} \eta_{\alpha(j_1) \cdots \alpha(j_k)} g^{\sigma(i_1)\alpha(j_1)} \cdots g^{\sigma(i_1) \cdots \alpha(j_k)} \tag{3.1.4}$$

It is important to note that there are technically four sums in the line above. We are first summing over the ordered multi indices $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_k$, then for each multi index, we are summing over all possible permutations of each multi index. Note that:

$$\omega_{\sigma(i_1) \cdots \sigma(i_k)} = \operatorname{sgn}(\sigma) \omega_{i_1 \cdots i_k} \qquad \text{and} \qquad \eta_{\alpha(j_1) \cdots \alpha(j_k)} = \operatorname{sgn}(\alpha) \eta_{j_1 \cdots j_k}$$

so using our notation for raising and lowering indices, we can rewrite (3.1.4) as a sum over α and the multi indices I:

$$\begin{aligned} \langle \omega, \eta \rangle^k &= \sum_{\sigma \in S_k} \omega_{\sigma(i_1) \cdots \sigma(i_k)} \eta^{\sigma(i_1) \cdots \sigma(i_k)} \\ &= \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)^2 \omega_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k} \\ &= k! \cdot \omega_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k} \end{aligned}$$

Hence an orthonormal frame α^i of T_p^*M, $\langle \cdot, \cdot \rangle^k$ has the unfortunate property that:

$$\langle \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}, \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} \rangle_p^k = k! g_p^{i_1 i_1} \cdots g_p^{i_k i_k}$$

where each $g_p^{i_l i_l} = \pm \delta^{i_l i_l}$. The above demonstrates that $\langle \cdot, \cdot \rangle^k$ is indeed a bundle metric on $\Lambda^k(TM)$, as the set:

$$\{\alpha_p^{i_1} \wedge \cdots \wedge \alpha_p^{i_k} : i_1 < \cdots < i_k\}$$

is then a local orthogonal frame for $\Lambda^k(T^*M)$. However, our lives would be easier if a local orthonormal basis for TM induced a local orthonormal frame of $\Lambda^k(TM)$. To fix this, we define a new bundle metric, $\langle \cdot, \cdot \rangle$, by:

$$\langle \cdot, \cdot \rangle = \frac{1}{k!} \langle \cdot, \cdot \rangle^k$$

It will be clear from context which rank of form we are taking the inner product of so we neglect to include the k in the notation. In particular, this inner product satisfies:

$$\begin{aligned} \langle \omega, \eta \rangle &= \omega_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k} \\ &= \frac{1}{k!} \sum_{i_1 \cdots i_k} \omega_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k} \end{aligned}$$

where the second line is a sum over all unordered multi indices. Furthermore, for complex valued k-forms, $\omega, \eta \in \Omega^k(M, \mathbb{C}) \cong \Omega^k(M) \otimes \mathbb{C}$ we set:

$$\langle \omega, \eta \rangle = \bar{\omega}_{i_1 \cdots i_k} \eta^{i_1 \cdots i_k}$$

Definition 3.1.1. Let (M, g) be an orientable (pseudo) Riemannian manifold, where g has signature (t, s). The **Hodge Star Operator**:

$$\star : \Omega^k(M) \longrightarrow \Omega^{n-k}(M)$$

is the linear map defined by:

$$\langle \omega, \eta \rangle \mathrm{dvol}_g = \omega \wedge \star \eta$$

for all $\omega, \eta \in \Omega^k(M)$. For complex valued k-forms we have that:

$$\langle \omega, \eta \rangle \mathrm{dvol}_g = \bar{\omega} \wedge \star \eta$$

Lemma 3.1.1. *In a local oriented orthornormal coframe $\{\alpha_i\}$, the Hodge star operator is given by:*

$$\star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) = g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{i_1 \cdots i_k i_{k+1} \cdots i_n} \alpha^{i_{k+1}} \wedge \cdots \wedge \alpha^{i_n}$$

where there is no summation over the indices, $\{i_{k+1} \cdots, i_n\}$ is a complimentary set to $\{i_1, \cdots, i_k\}$, and ϵ is totally antisymmetric with:

$$\epsilon_{123 \cdots n} = +1$$

In particular:

$$\star dvol_g = (-1)^t \qquad \text{and} \qquad \star(1) = dvol_g$$

Proof. By construction, the set:

$$\{\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} : i_1 < \cdots < i_k\}$$

is an orthornormal frame for $\Omega^k(U)$, hence for any arbitrary k form ω we have that:

$$\begin{aligned} \langle \omega, \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} \rangle =& \omega_{j_1 \cdots j_k} \langle \alpha^{j_1} \wedge \cdots \wedge \alpha^{j_k}, \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} \rangle \\ =& \omega_{i_1 \cdots i_k} g^{i_1 i_1} \cdots g^{i_k i_k} \end{aligned}$$

where the is no implied summation over the indices. According to **Definition 3.1.1** and **Theorem 1.1.12**:

$$\begin{aligned} \omega \wedge \star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) =& \omega_{i_1 \cdots i_k} g^{i_1 i_1} \cdots g^{i_k i_k} \alpha^1 \wedge \cdots \wedge \alpha^n \\ =& \omega_{i_1 \cdots i_k} g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{1 \cdots n} \alpha^1 \wedge \cdots \wedge \alpha^n \end{aligned}$$

There exists a permutation $\sigma \in S_n$ such that for all $j \leq k$, $\sigma(j) = i_j$, hence:

$$\omega \wedge \star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) = \omega_{i_1 \cdots i_k} g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{1 \cdots n} \mathrm{sgn}(\sigma) \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} \wedge \alpha^{\sigma(k_1)} \wedge \cdots \wedge \alpha^{\sigma(n)}$$

Since ϵ is totally antisymmetric:

$$\epsilon_{\sigma(1) \cdots \sigma(n)} = \mathrm{sgn}(\sigma) \epsilon_{1 \cdots n}$$

thus:

$$\omega \wedge \star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) = \omega_{i_1 \cdots i_k} \alpha^{i_1 \cdots i_k} \wedge \left(g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{i_1 \cdots i_k \sigma(k+1) \cdots \sigma(n)} \alpha^{\sigma(k+1)} \wedge \cdots \wedge \alpha^{\sigma(n)} \right)$$

$\sigma(k+1), \ldots, \sigma(n)$ is a set complimentary to $i_1, \ldots, i_k$, and since ω is a k form, and $\alpha^{\sigma(k+1)} \wedge \cdots \wedge \alpha^{\sigma(n)}$ is an $n-k$ form, the only component of ω that appears in the wedge product is the $\omega_{i_1 \cdots i_k}$ component. Relabeling $\alpha(k+j)$ by i_{k+j} we find:

$$\omega \wedge \star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) = \omega \wedge \left(g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{i_1 \cdots i_k i_{k+1} \cdots i_n} \alpha^{i_{k+1}} \wedge \cdots \wedge \alpha^{i_n} \right)$$

Since this holds for arbitrary ω, we have:

$$\star(\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}) = g^{i_1 i_1} \cdots g^{i_k i_k} \epsilon_{i_1 \cdots i_k i_{k+1} \cdots i_n} \alpha^{i_{k+1}} \wedge \cdots \wedge \alpha^{i_n}$$

as desired. Furthermore:

$$\star(1) = \epsilon_{1 \cdots n} \alpha^1 \wedge \cdots \wedge \alpha^n = \mathrm{dvol}_g$$

and:

$$\star(\mathrm{dvol}_g) = g^{11} \cdots g^{nn} \epsilon_{1 \cdots n} = (-1)^t$$

$\square$

With the lemma above we can prove the following:

Proposition 3.1.2. *Let (M, g) be an n-dimensional, orientable, (pseudo)-Riemannian manifold of signature (t, s). The map :*

$$\star\star : \Omega^k(M) \longrightarrow \Omega^k(M)$$

is a linear isomorphism given by:

$$\star\star = (-1)^{t+k(n-k)}$$

Proof. Since $\star$ is linear, we need only check this for an arbitrary k form ω written in an orthonormal dual basis $\{\alpha^i\}$:

$$\omega = \alpha^{m_1} \wedge \cdots \wedge \alpha^{m_k}$$

From **Lemma 3.1.1**, we see that:

$$\star\omega = g^{m_1 m_1} \cdots g^{m_k m_k} \epsilon_{m_1 \cdots m_k m_{k+1} \cdots m_n} \alpha^{m_{k+1}} \wedge \cdots \wedge \alpha^{m_n}$$

Applying the $\star$ again, we see that:

$$\star \star \omega = g^{m_1 m_1} \cdots g^{m_k m_k} \epsilon_{m_1 \cdots m_k m_{k+1} \cdots m_n} \star (\alpha^{m_{k+1}} \wedge \cdots \wedge \alpha^{m_n})$$

where:

$$\star (\alpha^{m_{k+1}} \cdots \alpha^{m_n}) = g^{m_{k+1} m_{k+1}} \cdots g^{m_n m_n} \epsilon_{m_{k+1} \cdots m_n m_1 \cdots m_k} \alpha^{m_1} \wedge \cdots \wedge \alpha^{m_k}$$

If g has signature (s, t), then:

$$g^{m_1 m_1} \cdots g^{m_n m_n} = (1)^s (-1)^t = (-1)^t$$

Furthermore, since $\epsilon_{1 \cdots n}$ is totally antisymmetric, we have:

$$\begin{aligned} \epsilon_{m_{k+1} \cdots m_n m_1 \cdots m_{k+1}} &= (-1)^{n-k} \epsilon_{m_1 m_{k+1} \cdots m_n m_2 \cdots m_k} \\ \Rightarrow \epsilon_{m_{k+1} \cdots m_n m_1 \cdots m_k} &= (-1)^{k(n-k)} \epsilon_{m_1 \cdots m_k m_{k+1} \cdots m_n} \end{aligned}$$

and since:

$$\epsilon_{m_1 \cdots m_k m_{k+1} \cdots m_n} = \pm 1$$

we see that:

$$\begin{aligned} \star \star \omega &= (-1)^t (-1)^{k(n-k)} \alpha^{m_1} \wedge \cdots \wedge \alpha^{m_k} \\ &= (-1)^{t+k(n-k)} \omega \end{aligned}$$

Therefore:

$$\star\star = (-1)^{t+k(n-k)}$$

as desired.

□

Definition 3.1.2. Let (M, g) be an orientable (pseudo)-Riemannian manifold. We define the $\mathbf{L^2}$ **Inner Product on differential** k **forms with compact support** by:

$$\begin{aligned} \langle \omega, \eta \rangle_{L^2} &= \int_M \langle \omega, \eta \rangle \mathrm{dvol}_g \\ &= \int_M \omega \wedge \star\eta \end{aligned}$$

If M is compact, then this defines an inner product on all of $\Omega^k(M)$.

We want to check that this is indeed an inner product. It is clear that the L^2 inner product is symmetric and bilinear; we prove that it is nondegenerate below.

Proposition 3.1.3. *Let (M, g) be an orientable (pseudo)-Riemannian manifold of signature (t, s). The L^2 inner product on k forms with compact support is nondegenerate.*

Proof. We proceed by contradiction; suppose the L^2 inner product is degenerate, then there exists a k form ω with compact support on M such that:

$$\langle \omega, \eta \rangle_{L^2} = 0$$

for all $\eta \in \Omega^k(M)$ with compact support. Let $\operatorname{supp} \omega = K \subset M$ for some compact set K; let $x \in \operatorname{int} K$, and U be an open neighborhood of x. Since K is the closure of an open set in M, it follows that $U \subset K$. The open neighborhood U then admits an orthornormal frame of k forms:

$$\{\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} : i_1 < \cdots < i_k\}$$

In this frame let:

$$\omega = \omega_{i_1 \cdots i_k} \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$$

where $\omega_{i_1 \cdots i_k}$ are smooth functions on U.

There exists a positively oriented coordinate chart ϕ such that $\phi(x) = 0$ and $\phi(U)$ is an open ball of radius r in $\mathbb{R}^n$ centered at 0. The closed ball B^{r_0} of radius $r_0 < r$ is then a nonempty compact subset of $\phi(U)$, and by continuity of ϕ^{-1}, $L = \phi^{-1}(B^{r_0}) \subset K$ is then a compact set in M. We construct a smooth bump function on U by first defining the smooth function f on $\phi(U)$ by:

$$f(x) = \begin{cases} \exp\left(\frac{r_0}{r_0 - (x^1)^2 - \cdots - (x^n)^2}\right) \text{ for} & (x^1)^2 + \cdots + (x^n)^2 < r_0 \\ 0 & \text{otherwise} \end{cases}$$

$\phi^* f$ is then a smooth function on U, satisfying $\operatorname{supp} \phi^* f = L$. This function can be smoothly extended to all of M, by defining:

$$h(p) = \begin{cases} \phi^* f(p) \text{ for} & p \in U \\ 0 & \text{otherwise} \end{cases}$$

Clearly, $\operatorname{supp} h = L$ as well, hence we construct global k forms with compact support equal to L by:

$$\eta^{i_1 \cdots i_k} = h \cdot \omega_{i_1 \cdots i_k} \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$$

where there are is no implied summation in the line above. We then see that for all $i_1 < \cdots < i_k$:

$$\begin{aligned} \langle \omega, \eta^{i_1 \cdots i_k} \rangle_{L^2} &= \int_M \langle \omega, \eta^{i_1 \cdots i_k} \rangle \mathrm{dvol}_g \\ &= \int_{\phi(U)} \phi^{-1*} \left(h \cdot \omega_{j_1 \cdots j_k} \omega_{i_1 \cdots i_k} \langle \alpha^{j_1} \wedge \cdots \wedge \alpha^{j_k}, \alpha^{i_1} \cdots \alpha^{i_k} \rangle \mathrm{dvol}_g\right) \\ &= \pm \int_{\phi(U)} \phi^{-1*} \left(h \cdot \omega^2_{i_1 \cdots i_k} \mathrm{dvol}_g\right) \end{aligned}$$

where the sign depends on $\langle \cdot, \cdot \rangle$. We have that $h > 0$ on L by construction, and clearly $\omega^2_{i_1 \cdots i_k} \geq 0$, for all $i_1 < \cdots < i_k$. Furthermore, $\omega_{i_1 \cdots i_k}$ can't be identically zero for all $i_1 < \cdots < i_k$ on L as $L \subset U \subset \operatorname{int} K$, thus by **Theorem 1.1.5** there exists an ordered multi index $i_1, \ldots, i_k$ such that:

$$\int_{\phi(U)} \phi^{-1*} \left(h \cdot \omega^2_{i_1 \cdots i_k} \mathrm{dvol}_g\right) > 0$$

A contradiction, therefore, $\langle \cdot, \cdot \rangle_{L^2}$ is non degenerate and thus defines an inner product of k forms on M with compact support, as desired. ☐

We wish to extend the work above to k-forms twisted with a sections of a vector bundle. If E is a $\mathbb{K} = \mathbb{C}, \mathbb{R}$ linear vector bundle over M, equipped with a bundle metric $\langle \cdot, \cdot \rangle_E$, then a similar argument to **Proposition 3.1.1** shows that there is a bundle metric on $\Lambda^k(TM) \otimes E$ induced by the (pseudo)-Riemannian metric g, and the bundle metric $\langle \cdot, \cdot \rangle_E$. We commit the mild notational sin of denoting this bundle metric by $\langle \cdot, \cdot \rangle_E$, and see that pointwise it is given by:

$$\langle A_i^j \omega^i \otimes v_j, B_k^l \eta^k \otimes w_l \rangle_{E_p} = A_i^j B_k^l \langle \omega^i, \eta^k \rangle_p \langle v_j, w_k \rangle_{E_p}$$

for all $A_i^j \omega^i \otimes v_j, B_k^l \eta^k \otimes w_l \in (\Lambda^k(TM) \otimes E)_p$.

Definition 3.1.3. The bundle metric $\langle \cdot, \cdot \rangle_E$ on $\Lambda^k(T^*M) \otimes E$ induces an **inner product on twisted k forms**:

$$\langle \cdot, \cdot \rangle_E : \Omega^k(M, E) \times \Omega^k(M, E) \longrightarrow C^\infty(M, E)$$

Given a local frame e_i for E over $U \subset M$, then for $\omega, \eta \in \Omega^k(M, E)$, we have that:

$$\omega = \omega^i \otimes e_i \qquad \text{and} \qquad \eta = \eta^i \otimes e_i$$

where $\omega^i, \eta^i \in \Omega^k(M, \mathbb{K})$. The inner product is then given by:

$$\langle \omega, \eta \rangle_E = \langle \omega^i, \eta^j \rangle \cdot \langle e_i, e_j \rangle_E$$

which is independent of our choice of frame. Furthermore, the **Hodge star operator on twisted differential forms**:

$$\star : \Omega^k(M, E) \longrightarrow \Omega^{n-k}(M, E)$$

is given by:

$$\star\omega = (\star\omega^i) \otimes e_i$$

Finally, we have an $\mathbf{L^2}$ **inner product on twisted differential forms with compact support** given by:

$$\langle \omega, \eta \rangle_{E, L^2} = \int_M \langle \omega, \eta \rangle_E \mathrm{dvol}_g$$

which is nondegenerate by a similar argument to **Proposition 3.1.2**.

Recall that the exterior derivative d is a linear map from k forms to $k+1$ forms, so if $\omega \in \Omega^k(M)$ and $\eta \in \Omega^{k+1}(M)$ both have compact support, it makes sense to take the following L^2 inner product:

$$\langle d\omega, \eta \rangle_{L^2} = \int_M \langle d\omega, \eta \rangle \mathrm{dvol}_g$$

Similarly, if ∇ is a covariant derivative on the vector bundle E, and $\omega \in \Omega^k(M, E)$ and $\eta \in \Omega^{k+1}(M, E)$ both have compact support, we can take[35]:

$$\langle d_\nabla \omega, \eta \rangle_{L^2, E} = \int_M \langle d_\nabla \omega, \eta \rangle \mathrm{dvol}_g$$

As mentioned earlier, our goal is to now develop a *formal adjoint* to d, and d_∇, which will clearly have to be a map a from k forms to $k-1$ forms.

Definition 3.1.4. Let (M, g) be an n-dimensional, orientable, (pseudo)-Riemannian manifold of signature (t, s). We define the **codifferential**:

$$d^\star : \Omega^k(M) \longrightarrow \Omega^{k-1}(M)$$

by:

$$d^\star = (-1)^{t+nk+1} \star d \star$$

[35] d_∇ is the exterior covariant induced by a general covariant derivative. The definition is the same as **Definition 2.1.30**

Example 3.1.1. The above definition allows one extend the notion of the Laplacian, ∇^2, on functions on $\mathbb{R}^n$ to k forms on orientable (pseudo)- Riemannian manifolds. Indeed, the **Laplace-de Rham operator** is defined as:

$$\Delta = dd^\star + d^\star d$$

Then if we set $M = \mathbb{R}^n$ with the standard Euclidean metric on $\mathbb{R}^n$, then for $f \in \mathbb{C}^\infty(\mathbb{R}^n)$ we have that:

$$\begin{aligned}\Delta f =& dd^\star f + d^\star df\\ =&(-1)d \star d(f \mathrm{dvol}_g) + (-1)^{n+1} \star d \star \left(\partial_i f dx^i\right)\end{aligned}$$

The first term goes to 0 as dvol_g is a top form; taking the Hodge star of df then gives:

$$\star(\partial_i f dx^i) = \partial_i f dx^1 \wedge \cdots \wedge d\hat{x}^i \wedge \cdots \wedge dx^n$$

where $d\hat{x}^i$ denotes the deletion of the ith component. Taking the exterior derivative we obtain:

$$d(\star\partial_i f dx^i) = \partial_i^2 f dx^1 \wedge \cdots \wedge dx^i \wedge \cdots \wedge dx^n$$

Finally, taking the Hodge star again we have that:

$$\Delta f = (-1)^{n+1} \sum_i^n \partial_i^2 f = (-1)^{n+1} \nabla^2 f$$

so the general Laplace-de Rham operator agrees with the Laplacian on smooth functions on $\mathbb{R}^n$, up to a sign.

Theorem 3.1.1. *Let M be an orientable, (pseudo)-Riemannian manifold without boundary of signature (t, s). Then the codifferential $d^\star$ is the formal adjoint of d with respect to the L^2 inner product on k forms with compact support, i.e.*

$$\langle d\omega, \eta\rangle_{L^2} = \langle \omega, d^\star \eta\rangle_{L^2}$$

for all $\omega \in \Omega^k(M)$ and $\eta \in \Omega^{k+1}(M)$ with compact support.

Proof. We have that by **Proposition 3.1.2**:

$$\begin{aligned}(\langle d\omega, \eta\rangle - \langle \omega, d^\star\eta\rangle) \mathrm{dvol}_g =&(d\omega) \wedge \star\eta - \omega \wedge \star(d^\star \eta)\\ =&(d\omega) \wedge \star\eta - (-1)^{t+nk+1} \omega \wedge \star(\star d \star \eta)\\ =&(d\omega) \wedge \star\eta + (-1)^{2t+2nk-k^2+2} \omega \wedge (d \star \eta)\\ =&(d\omega) \wedge \star\eta + (-1)^{-k^2} \omega \wedge (d \star \eta)\end{aligned}$$

Note that k is an integer, hence if k is even $-k^2$ is even, and if k is odd $-k^2$ is odd so:

$$\begin{aligned}(\langle d\omega, \eta\rangle - \langle \omega, d^\star\eta\rangle) \mathrm{dvol}_g =&(d\omega) \wedge \star\eta + (-1)^k \omega \wedge (d \star \eta)\\ =&d(\omega \wedge \star\eta)\end{aligned}$$

By **Theorem 1.1.7**, i.e. Stoke's Theorem, we obtain:

$$\begin{aligned}\int_M \langle d\omega, \eta\rangle - \langle \omega, d^\star\eta\rangle \mathrm{dvol}_g =& \int_M d(\omega \wedge \star\eta)\\ =& \int_{\partial M} \omega \wedge \star\eta\\ =&0\end{aligned}$$

as M has empty boundary. □

As before, we can extend these results to forms twisted with E.

Definition 3.1.5. Let E be $\mathbb{K} = \mathbb{C}, \mathbb{R}$ linear vector bundle over a an n dimensional, orientable, (pseudo)-Riemannian manifold signature of (t, s), and ∇ a covariant derivative on E. We define the **covariant codifferential**:

$$d^\star_\nabla : \Omega^k(M, E) \longrightarrow \Omega^{k-1}(M, E)$$

by:

$$d^\star_\nabla = (-1)^{t+nk+1} \star d_\nabla \star$$

We will also need the following definition:

Definition 3.1.6. Let $\pi : E \to M$ be $\mathbb{K} = \mathbb{C}, \mathbb{R}$ linear vector bundle with a bundle metric $\langle \cdot, \cdot \rangle_E$. The **inner product of forms twisted with E and sections of E**:

$$\langle \cdot, \cdot \rangle_E : \Omega^k(M, E) \times \Gamma(E) \longrightarrow \Omega^k(M)$$

by choosing a local frame as did before, and setting:

$$\langle \omega^i \otimes e_i, \phi^j e_j \rangle_E = \omega^i \phi^j \langle e^i, e_j \rangle_E$$

This inner product is independent of the choice of frame.

Lemma 3.1.2. *Let E be a vector bundle over M, with bundle metric $\langle \cdot, \cdot \rangle_E$, and a covariant derivative ∇ which respects the metric, then in any local frame e_i:*

$$d(\langle \omega, \Phi \rangle) = (d\omega^i)\langle e_i, \Phi \rangle_E + (-1)^k \omega^i \wedge (\langle \nabla e_i, \Phi \rangle_E + \langle e_i, \nabla \Phi \rangle_E)$$

for all $\omega \in \Omega^k(M, E)$ and $\Phi \in \Gamma(E)$.

Proof. We see that:

$$\begin{aligned} d(\langle \omega, \Phi \rangle_E) =& d(\omega^i \langle e_i, \Phi \rangle_E) \\ =& (d\omega^i)\langle e_i, \Phi \rangle_E + (-1)^k \omega^i \wedge d\langle e_i, \Phi \rangle_E \end{aligned}$$

Furthermore, for any $X \in \mathfrak{X}(M)$:

$$\begin{aligned} (d\langle e_i, \Phi \rangle_E)(X) =& \mathscr{L}_X \langle e_i, \Phi \rangle_E \\ =& \langle \nabla_X e_i, \Phi \rangle_E + \langle e_i, \nabla_X \Phi \rangle_E \end{aligned}$$

hence:

$$d(\langle \omega, \Phi \rangle) = (d\omega^i)\langle e_i, \Phi \rangle_E + (-1)^k \omega^i \wedge (\langle \nabla e_i, \Phi \rangle_E + \langle e_i, \nabla \Phi \rangle_E)$$

as desired. □

Theorem 3.1.2. *Let E be a $\mathbb{K} = \mathbb{C}, \mathbb{R}$ linear vector bundle over an n dimensional, orientable, (pseudo)-Riemannian manifold without boundary of signature (t, s), and ∇ a covariant derivative on E. Suppose E is equipped with the bundle metric $\langle \cdot, \cdot \rangle_E$, and that ∇ is metric compatible, then:*

$$\langle d_\nabla \omega, \eta \rangle_{E, L^2} = \langle \omega, d^\star_\nabla \eta \rangle_{E, L^2}$$

for all $\omega \in \Omega^k(M, E)$ and $\eta \in \Omega^{k+1}(M)$ with compact support.

Proof. Let e^i be a local frame for E over $U \subset M$, such that:

$$\omega = \omega^i \otimes e_i \qquad \text{and} \qquad \eta = \eta^i \otimes e_i$$

Then:

$$\begin{aligned} d^\star_\nabla \eta =& (-1)^{t+nk+1} \star d_\nabla (\star \eta^i \otimes e_i) \\ =& (-1)^{t+nk+1} \star \left(d(\star \eta^i) \otimes e_i + (-1)^{n-k-1} (\star \eta^i) \wedge \nabla e_i \right) \\ =& (d^\star \eta^i) \otimes e_i + (-1)^{t+nk+n-k} \star \left((\star \eta^i) \wedge \nabla e_i \right) \end{aligned}$$

hence by **Proposition 3.1.3**:

$$\star d^\star_\nabla \eta = (-1)^{k+1}(d\star\eta^i)\otimes e_i + (-1)^n(\star\eta^i)\wedge\nabla e_i$$

If we can show that the integral:

$$\int_U (\langle d_\nabla(\omega^i\otimes e_i), \eta^j\otimes e_j\rangle_E - \langle \omega^i\otimes e_i, d^\star_\nabla(\eta^j\otimes e_j)\rangle_E)\mathrm{dvol}_g = 0$$

in any local frame e_i over U, then a partition of unity argument clearly proves the claim. We see that:

$$\begin{aligned}\langle d_\nabla(\omega^i\otimes e_i), \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g =&\langle d(\omega^i)\otimes e_i + (-1)^k\omega^i\wedge\nabla e_i, \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g\\ =&(d\omega^i\wedge\star\eta^j)\langle e_i, e_j\rangle + (-1)^k\langle\omega^i\wedge\nabla e_i, \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g \qquad (3.1.5)\end{aligned}$$

Note that:

$$\nabla e_i = \sigma_i^k\otimes e_k$$

for some one forms σ_i^k on M, so:

$$\begin{aligned}\langle\omega^i\wedge\nabla e_i, \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g =&\langle\omega^i\wedge\sigma_i^k\otimes e_k, \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g\\ =&\langle\omega^i\wedge\sigma_i^k, \eta^j\rangle\langle e_k, e_j\rangle_E\mathrm{dvol}_g\\ =&(\omega^i\wedge\sigma_i^k)\wedge\star\eta^j\langle e_k, e_j\rangle_E\\ =&(-1)^{n-k-1}(\omega^i\wedge\star\eta^j)\wedge\langle\sigma_k^i\otimes e_k, e_j\rangle_E\\ =&(-1)^{n-k-1}(\omega^i\wedge\star\eta^j)\wedge\langle\nabla e_i, e_j\rangle_E\end{aligned}$$

hence (3.1.5) becomes:

$$\langle d_\nabla(\omega^i\otimes e_i), \eta^j\otimes e_j\rangle_E\mathrm{dvol}_g = (d\omega^i\wedge\star\eta^j)\langle e_i, e_j\rangle_E + (-1)^{n-1}(\omega^i\wedge\star\eta^j)\wedge\langle\nabla e_i, e_j\rangle_E \qquad (3.1.6)$$

Furthermore:

$$\begin{aligned}\langle\omega^i\otimes e_i, d^\star_\nabla(\eta^j\otimes e_j)\rangle_E =&\langle\omega^i\otimes e_i, (d^\star\eta^j)\otimes e_j + (-1)^{t+nk+n-k}\star\big((\star\eta^i)\wedge\nabla e_j\big)\rangle_E\\ =&(-1)^{k+1}\omega^i\wedge(d\star\eta^j)\langle e_i, e_j\rangle_E\\ &+(-1)^{t+nk+n-k}\langle\omega^i\otimes e_i, \star\big((\star\eta^i)\wedge\nabla e_j\big)\rangle_E\mathrm{dvol}_g \qquad (3.1.7)\end{aligned}$$

In a similar manner we see that:

$$\begin{aligned}\langle\omega^i\otimes e_i, \star\big((\star\eta^i)\wedge\nabla e_j\big)\rangle_E\mathrm{dvol}_g =&\langle\omega^i\otimes e_i, \star\big((\star\eta^i)\wedge\sigma_j^k\big)\otimes e_k\rangle_E\mathrm{dvol}_g\\ =&\langle\omega^i, \star\big((\star\eta^i)\wedge\sigma_j^k\big)\rangle\langle e_i, e_k\rangle_E\mathrm{dvol}_g\\ =&(-1)^{t+k(n-k)}\omega^i\wedge(\star\eta^j\wedge\sigma_j^k)\langle e_i, e_k\rangle_E\\ =&(-1)^{t+k(n-k)}(\omega^i\wedge\star\eta^j)\wedge\langle e_i, \nabla e_j\rangle_E\end{aligned}$$

hence (3.1.7) becomes:

$$\begin{aligned}\langle\omega^i\otimes e_i, d^\star_\nabla(\eta^j\otimes e_j)\rangle_E\mathrm{dvol}_g =&(-1)^{k+1}\omega^i\wedge(d\star\eta^j)\langle e_i, e_j\rangle_E\\ &+(-1)^n(\omega^i\wedge\star\eta^j)\wedge\langle e_i, \nabla e_j\rangle_E \qquad (3.1.8)\end{aligned}$$

The difference of (3.1.7) and (3.1.18) then yields:

$$\begin{aligned}(\langle d_\nabla\omega, \eta\rangle_E - \langle\omega, d^\star_\nabla\eta\rangle)\mathrm{dvol}_g =&(d\omega^i\wedge\star\eta^j)\langle e_i, e_j\rangle_E + (-1)^k\omega^i\wedge(d\star\eta^j)\langle e_i, e_j\rangle_E\\ &+(-1)^{n-1}\omega^i\wedge\star\eta^j\wedge(\langle\nabla e_i, e_j\rangle_E + \langle e_i, \nabla e_j\rangle_E)\\ =&d(\omega^i\wedge\star\eta^j)\langle e_i, e_j\rangle_E + (-1)^{n-1}\omega^i\wedge\star\eta^j\wedge(\langle\nabla e_i, e_j\rangle + \langle e_i, \nabla e_j\rangle)\end{aligned}$$

Note that $\omega^i\wedge\star\eta^j$ is an $n-1$ form for all i, j, hence by **Lemma 3.1.2** we see that:

$$(\langle d_\nabla\omega, \eta\rangle_E - \langle\omega, d^\star_\nabla\eta\rangle)\mathrm{dvol}_g = d(\omega^i\wedge\star\eta^j\langle e_i, e_j\rangle_E)$$

Since U is an open submanifold of M without boundary, the claim then follows from Stoke's theorem. □

3.1.2 The Yang-Mills Lagrangian

We are finally in a position to define the Yang-Mill's Lagrangian, and derive the Yang-Mill's equation. We fix the following data:

- An n dimensional, oriented, (pseudo) Riemannian manifold (M, g)
- A principal G bundle over M with compact structure group G
- An Ad-invariant inner product on $\mathfrak{g}$
- An orthonormal basis T^i for $\mathfrak{g}$ with respect to the aforementioned inner product.

Note that the compactness of G guarantees the existence of an Ad-invariant inner product by **Theorem 1.2.5**. Furthermore, by **Proposition 2.1.13**, the Ad-invariant inner product on $\mathfrak{g}$ induces a bundle metric $\langle \cdot, \cdot \rangle_{\mathrm{Ad}(P)}$ on the adjoint bundle.

Recall that the curvature form F^A of any connection A defines a unique twisted two form $F^A_M \in \mathrm{Ad}(P)$.

Definition 3.1.7. The **Yang-Mill's Lagrangian** is defined by:

$$\mathscr{L}_{YM}[A] = \langle F^A_M, F^A_M \rangle_{\mathrm{Ad}(P)}$$

For a fixed connection A, the Yang-Mill's Lagrangian is a global smooth function:

$$\mathscr{L}_{YM}[A] : M \to \mathbb{R}$$

Theorem 3.1.3. *The Yang-Mill's Lagrangian is gauge invariant, i.e. for any global bundle automorphism f:*

$$\mathscr{L}_{YM}[f^*A] = \mathscr{L}_{YM}[A]$$

Proof. We see that by **Theorem 2.1.12**:

$$F^{f^*A} = \mathrm{Ad}_{\sigma_f^{-1}} \circ F^A$$

This then implies that for any $x \in M$, and $X, Y \in T_xM$:

$$\begin{aligned}
\left(F^{f^*A}_M\right)_x (X_1, X_2) &= \left[p, \mathrm{Ad}_{\sigma_f^{-1}(p)} \circ F^A_p(Y_1, Y_2)\right] \\
&= \left[p \cdot \sigma_{f^{-1}}(p), F^A_p(Y_1, Y_2)\right] \\
&= \left[f^{-1}(p), F^A_p(Y_1, Y_2)\right]
\end{aligned}$$

where $\pi(p) = x$ and $\pi_* X_i = Y_i$. Therefore:

$$F^{f^*A}_M = f^{-1} \cdot F^A_M$$

where $f^{-1}\cdot$ denotes the action of $\mathscr{G}(P)$ on $\mathrm{Ad}(P)$ defined in **Theorem 2.1.5**. F^A_M takes values in $\mathrm{Ad}(P)$, so we have that f^{-1} acts on F^A_M via the adjoint action, and, since the bundle metric on $\mathrm{Ad}(P)$ is Ad invariant by construction, it follows that:

$$\left\langle F^A_M, F^A_M \right\rangle_{\mathrm{Ad}(P)} = \left\langle f^{-1} \cdot F^A_M, f^{-1} \cdot F^A_M \right\rangle_{\mathrm{Ad}(P)}$$

Thus:

$$\mathscr{L}_{YM}[A] = \mathscr{L}_{YM}[f^*A]$$

as desired. □

Definition 3.1.8. Let $\mathscr{A}(P)$ denote the set of connections on P. As mentioned earlier, this set is an affine space over the vector space:

$$\Omega^1_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}} \cong \Omega^1(M, \mathrm{Ad}(P))$$

For any $\alpha \in \Omega^1_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$, we denote by $\alpha_M \in \Omega^1(M, \mathrm{Ad}(P))$ the image of α under the isomorphism constructed in **Theorem 2.1.18**.

We now assume M to be a closed manifold, i.e. compact and without boundary.

Definition 3.1.9. The **Yang-Mill's Action** on P is a smooth map:

$$\mathscr{A}(P) \longrightarrow \mathbb{R}$$

given by:

$$\begin{aligned} S_{YM}[A] &= -\frac{1}{2}\int_M \langle F_M^A, F_M^A \rangle_{\mathrm{Ad}(P)} \,\mathrm{dvol}_g \\ &= -\frac{1}{2}\langle F_M^A, F_M^A \rangle_{\mathrm{Ad}(P),L^2} \end{aligned}$$

The integral is well defined because M is compact.

Our goal is to determine the stationary points of the above action.

Definition 3.1.10. We call a connection A a **stationary or critical point** of the Yang-Mill's action if:

$$\frac{d}{dt}\Big|_{t=0} S_{YM}[A + t\alpha] = 0$$

for all $\alpha \in \Omega^1_{\mathrm{hor}}(P, \mathfrak{g})^{\mathrm{Ad}}$.

We denote the covariant codifferential on $\mathrm{Ad}(P)$ associated to A by $d_A^\star$.

Theorem 3.1.4. *The stationary points of the Yang-Mill's actions satisfy:*

$$d_A^\star F_M^A = 0 \tag{3.1.9}$$

Equivalently:

$$d_A \star F_M^A = 0 \tag{3.1.10}$$

Both (3.1.9) *and* (3.1.10) *are called the* ***Yang-Mills Equation****.*

Proof. We first note that:

$$\begin{aligned} F^{A+t\alpha} =& dA + td\alpha + \frac{1}{2}[A, A] + t[A, \alpha] + \frac{1}{2}t^2[\alpha, \alpha] \\ =& F^A + td\alpha + t[A, \alpha] + \frac{1}{2}t^2[\alpha, \alpha] \end{aligned}$$

Note for any local gauge $s : U \to P_U$:

$$\begin{aligned} (d_A\alpha_M)_s =& d\alpha_s + \rho_*(A_s)\alpha_s \\ =& d\alpha_s + [A_s, \alpha_s] \end{aligned}$$

hence under the isomorphism constructed in **Theorem 2.1.18** we obtain:

$$F_M^{A+t\alpha} = F_M^A + td_A\alpha_M + \frac{1}{2}t^2[\alpha_M, \alpha_M]$$

This then implies that:

$$\mathscr{L}_{YM}[A + t\alpha] = -\frac{1}{2}\langle F_M^A, F_M^A \rangle_{\mathrm{Ad}(P)} - t\langle d_A\alpha_M, F_M^A \rangle_{\mathrm{Ad}(P)} + \mathcal{O}(t^2)$$

Hence:

$$\frac{d}{dt}\Big|_{t=0} S_{YM}[A + t\alpha] = -\int_M \langle d_A\alpha_M, F_M^A \rangle_{\mathrm{Ad}(P)} \,\mathrm{dvol}_g$$

which by **Theorem 3.1.2** gives:

$$\frac{d}{dt}\Big|_{t=0} S_{YM}[A + t\alpha] = -\int_M \langle \alpha_M, d_A^\star F_M^A \rangle_{\mathrm{Ad}(P)} \,\mathrm{dvol}_g$$

Since the L^2 inner product is nondegenerate, we conclude that in order for the action to be stationary at A:

$$d_A^\star F_M^A = 0$$

□

Example 3.1.2. Recall from **Example 1.1.26** that the curvature form of the connection:

$$A = \frac{1}{2}\left(\bar{z}_1\alpha_1 - z_1\bar{\alpha}_1 + \bar{z}_2\alpha_2 - z_2\bar{\alpha}_2\right)$$

on the Hopf fibration $\mathbb{S}^3 \to \mathbb{S}^2$ was given by:

$$F_{\mathbb{S}^2} = \frac{i}{2}\mathrm{dvol}_g$$

where dvol_g is the volume associated to the usual round metric on $\mathbb{S}^2$. Furthermore, note that since $\mathbb{S}^1$ is an abelian Lie group, we have that for all local gauge's $s : U \to \mathbb{S}^3_U$, and all form's $\omega \in \Omega^k(\mathbb{S}^2, \mathrm{Ad}(\mathbb{S}^3))$ that:

$$(d_A\omega)_s = d\omega_s + [A_s, \omega_s] = d\omega_s$$

hence:

$$d_A = d$$

We then see that:

$$\begin{aligned} d_A^\star F_{\mathbb{S}^2} &= \star\, d\left(\star\frac{i}{2}\mathrm{dvol}_g\right) \\ &= \star\, d\left(\frac{i}{2}\right) \\ &= 0 \end{aligned}$$

so A is a Yang-Mill's connections on the Hopf fibration.

3.1.3 Source Free Electromagnetism

For the moment, we fix $P = \mathbb{R}^{1,3} \times U(1)$[36], and in any global gauge $s : \mathbb{R}^{1,3} \to P$ we write A_s as:

$$A_s = i\left(-V dt + M_i dx^i\right)$$

where V, M_i are all smooth functions on $\mathbb{R}^{1,3}$. We also choose an orientation, such that in the standard coordinates, the global four form:

$$dt \wedge dx \wedge dy \wedge dz$$

is our orientation/volume form. Our goal is to first show that the Yang-Mill's equation, and the Bianchi identity yield Maxwell's field equations in a vacuum, implying that Electromagnetism is a $U(1)$ gauge theory. We will then move onwards to incorporating various types of matter into our field equations, so that we obtain a full theory of classical electromagnetism.

First note that since $U(1)$ is abelian, we have the following simplifications:

$$\mathrm{Ad}(P) = \mathbb{R}^{1,3} \times i\mathbb{R} \qquad \text{and} \qquad d_A = d$$

This means that the curvature form can be thought of as a regular two form one $\mathbb{R}^{1,3}$ multiplied by the imaginary constant i, and that the Yang-Mill's equation, and Bianchi Identity simplify to:

$$d \star F_{\mathbb{R}^{1,3}} = 0 \qquad \text{and} \qquad dF_{\mathbb{R}^{1,3}} = 0$$

Furthermore, if F is the curvature form of a connection A, then any global gauge determines the element $F_{\mathbb{R}^{1,3}} \in \Omega(\mathbb{R}^{1,3}, i\mathbb{R})$, as if $s' = s \cdot h$ for some physical gauge transformation $h : \mathbb{R}^{1,3} \to U(1)$, we have that:

$$F_{s'} = \mathrm{Ad}_{h^{-1}} \circ F_s = F_s$$

so F_s is independent of our choice of gauge.

[36] By $\mathbb{R}^{1,3}$ we mean $\mathbb{R}^4$ with the Minkowski metric of signature $(-,+,+,+)$.

Proposition 3.1.4. *Let $s : U \to \mathbb{R}^{1,3}$ be a global gauge, and A a connection on P. If we identify A_s as the electromagnetic four potential, multiplied by i then the components F_s are the components of the physical fields* $\mathbf{E}$ *and* $\mathbf{B}$*. In particular, if we write F_s as an antisymmetric matrix, then F_s is exactly i multiplied by the* ***electromagnetic field tensor****:*

$$F = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}$$

Proof. Note that:

$$\mathbf{E} = -\nabla V - \partial_t \mathbf{M} \qquad \text{and} \qquad \mathbf{B} = \nabla \times \mathbf{M}$$

We need to calculate the coefficients of F_s in the standard (t, x, y, z) coordinates on $\mathbb{R}^{1,3}$. We see that:

$$F_s = dA_s$$

hence, if we set:

$$F_{\mu\nu} = F_s(\partial_\mu, \partial_\nu)$$

we have that by **Proposition 2.1.22**:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

We calculate:

$$\begin{aligned} F_{tx} &= i\left(-\partial_x V - \partial_t M_x\right) = -iE_x \\ F_{ty} &= i\left(-\partial_y V - \partial_t M_y\right) = -iE_y \\ F_{tz} &= i\left(-\partial_z V - \partial_t M_z\right) = -iE_z \\ F_{xy} &= i\left(\partial_x M_y - \partial_y M_x\right) = iB_z \\ F_{xz} &= i\left(\partial_x M_z - \partial_z M_x\right) = -iB_y \\ F_{yz} &= i\left(\partial_y M_z - \partial_z M_y\right) = iB_x \end{aligned}$$

which implies the claim. $\square$

Since $F_s \in \Omega(\mathbb{R}^{1,3}, i\mathbb{R})$ we see that the Yang-Mill's Lagrangian in coordinates is given by:

$$\mathscr{L}_{YM}[A] = -\frac{1}{2} F_{\mu\nu} F^{\mu\nu}$$

As usual, we are summing over ordered indices, $\mu < \nu$, and our convention in these coordinates is:

$$t < x < y < z$$

From **Definition 3.1.7** it should be clear for $P = \mathbb{R}^{1,3} \times U(1)$, that the Yang-Mill's Lagrangian is Lorentz invariant, as the inner product of any two tensors is a Lorentz invariant quantity. It is mildly less obvious that the Lagrangian is actually one of the simplest Lorentz invariant combinations of $\mathbf{E}$ and $\mathbf{B}$.

Proposition 3.1.5. *The Yang-Mill's Lagrangian on $\mathbb{R}^{1,3} \times U(1)$ is the fundamental Lorentz invariant quantity:*

$$\mu = \frac{1}{2}\left(\mathbf{E}^2 - \mathbf{B}^2\right)$$

where $\mathbf{E}^2 = \mathbf{E} \cdot \mathbf{E}$ *and* $\mathbf{B}^2 = \mathbf{B} \cdot \mathbf{B}$ *is the usual Euclidean dot product on* $\mathbb{R}^3$*.*

Proof. We see that:

$$F_{\mu\nu} F^{\mu\nu} = F_{\mu\nu} \eta^{\mu\sigma} \eta^{\nu\alpha} F_{\sigma\alpha}$$

where η is the Minkowski metric of signature $(-,+,+,+)$. We calculate, with no implied summation:

$$\begin{aligned}F_{\mu\nu}F^{\mu\nu} =&F_{tx}\eta^{tt}\eta^{xx}F_{tx} + F_{ty}\eta^{tt}\eta^{yy}F_{ty} + F_{tz}\eta^{tt}\eta^{zz}F_{tz}\\ &+F_{xy}\eta^{xx}\eta^{yy}F_{xz} + F_{xz}\eta^{xx}\eta^{zz}F_{xz} + F_{yz}\eta^{yy}\eta^{zz}F_{yz}\\ =&-(iE_x)(-iE_x) - (iE_y)(-iE_y) - (iE_z)(iE_z)\\ &\quad + (iB_z)(-iB_z) + (iB_y)(-iB_y) + (iB_z)(-iB_z)\\ =&-\mathbf{E}^2 + \mathbf{B}^2\end{aligned}$$

hence:

$$\mathscr{L}_{YM}[A] = \frac{1}{2}\left(\mathbf{E}^2 - \mathbf{B}^2\right)$$

□

We now wish to show our first main result of the section, that the Yang-Mill's equation, and the Bianchi identity yield Maxwell's field equations vacuum.

Theorem 3.1.5. *The Yang-Mill's equation and the Bianchi identity on $P = \mathbb{R}^{1,3} \times P$ are equivalent to the source free Maxwell's Field equations in a vacuum:*

$$\begin{aligned}\nabla\cdot\mathbf{E} =&0 \qquad \nabla\times\mathbf{E} =-\partial_t\mathbf{B}\\ \nabla\cdot\mathbf{B} =&0 \qquad \nabla\times\mathbf{B} = \ \ \partial_t\mathbf{E}\end{aligned}$$

As a result, source free classical electromagnetism in a vacuum is a $U(1)$ Yang-Mill's gauge theory.

Proof. Let s be a global gauge, and F_s be the global curvature form on $\mathbb{R}^{1,3}$ with components in the standard coordinates given by **Proposition 3.1.4**. We first wish to calculate the components $d\star F$. First note that since F is a two form on $\mathbb{R}^{1,3}$, that $d\star F$ is a three form, and thus has four components. We begin by calculating $d\star$ of the electric field components of F:

$$d(\star - iE_x dt\wedge dx) =d(iE_x dy\wedge dz) = i\partial_x E_x dx\wedge dy\wedge dz + i\partial_t E_x dt\wedge dy\wedge dz \tag{3.1.11}$$
$$d(\star - iE_y dt\wedge dy) =d(-iE_y dx\wedge dz) = i\partial_y E_y dx\wedge dy\wedge dz - i\partial_t E_y dt\wedge dx\wedge dz \tag{3.1.12}$$
$$d(\star - iE_z dt\wedge dz) =d(iE_z dx\wedge dy) = i\partial_z E_z dx\wedge dy\wedge dz + i\partial_t E_z dt\wedge dx\wedge dy \tag{3.1.13}$$

Examining the purely spacial components of (3.1.11), (3.1.12), (3.1.13) we see that:

$$i\left(\partial_x E_x + \partial_y E_y + \partial_z E_z\right) dx\wedge dy\wedge dz =i\nabla\cdot\mathbf{E}dx\wedge dy\wedge dz$$

Since $d\star F$ must vanish identically, we find that:

$$\nabla\cdot\mathbf{E} = 0$$

which is the first Maxwell equation. Continuing, with the magnetic field components of F:

$$\begin{aligned}d(\star iB_x dy\wedge dz) =&d(iB_x dt\wedge dx) = i\partial_y B_x dt\wedge dx\wedge dy + i\partial_z B_x dt\wedge dx\wedge dz\\ d(\star - iB_y dx\wedge dz) =&d(iB_y dt\wedge dy) = -i\partial_x B_y dt\wedge dx\wedge dy + i\partial_z B_y dt\wedge dy\wedge dz\\ d(\star iB_z dx\wedge dy) =&d(iB_z dt\wedge dz) = -i\partial_x B_z dt\wedge dy\wedge dz - i\partial_y B_z dt\wedge dy\wedge dz\end{aligned}$$

Looking at the $dt\wedge dy\wedge dz$ component,

$$i\left(\partial_z B_y - \partial_y B_z + \partial_t E_x\right) dt\wedge dy\wedge dz = i\left(-(\nabla\times\mathbf{B})_x + \partial_t E_x\right) dt\wedge dy\wedge dz$$

Since $d\star F$ vanishes identically:

$$(\nabla\times\mathbf{B})_x = \partial_t E_x \tag{3.1.14}$$

For the $dt\wedge dx\wedge dz$ component,:

$$i\left(\partial_z B_x - \partial_x B_z - \partial_t E_y\right) dt\wedge dx\wedge dz = i\left((\nabla\times\mathbf{B})_y - \partial_t E_y\right) dt\wedge dx\wedge dz$$

hence:

$$(\nabla \times \mathbf{B})_y = \partial_t E_y \tag{3.1.15}$$

Finally, for the $dt \wedge dx \wedge dy$ component,

$$i\,(\partial_y B_x - \partial_x B_y + \partial_t E_z)\,dt \wedge dx \wedge dy = i\,(-(\nabla \times \mathbf{B})_z + \partial_t E_z)$$

thus:

$$(\nabla \times \mathbf{B})_z = \partial_t E_z \tag{3.1.16}$$

Combining (3.1.14), (3.1.15), and (3.1.16) we obtain the fourth Maxwell equation:

$$\nabla \times \mathbf{B} = \partial_t \mathbf{E}$$

We now wish to calculate the components of dF. We have that dF is a three form, and thus as four components. We begin by calculating d of the magnetic field components of F:

$$d(iB_x dy \wedge dz) = i\partial_x B_x dx \wedge dy \wedge dz + i\partial_t B_x dt \wedge dy \wedge dz \tag{3.1.17}$$
$$d(-iB_y dx \wedge dz) = i\partial_y B_y dx \wedge dy \wedge dz - i\partial_t B_y dt \wedge dx \wedge dz \tag{3.1.18}$$
$$d(iB_z dx \wedge dy) = i\partial_z B_z dx \wedge dy \wedge dz + i\partial_t B_z dt \wedge dx \wedge dy \tag{3.1.19}$$

Combining the purely spatial components of (3.1.17), (3.1.18), and (3.1.19) we see:

$$i(\partial_x B_x + \partial_y B_y + \partial_z B_z) dx \wedge dy \wedge dz = \nabla \cdot \mathbf{B} dx \wedge dy \wedge dz$$

Since dF must vanish identically, we obtain the third Maxwell equation:

$$\nabla \cdot \mathbf{B} = 0$$

We continue with the electric field components of F:

$$\begin{aligned}
d(-iE_x dt \wedge dx) &= -i\partial_y E_x dt \wedge dx \wedge dy - i\partial_z E_x dt \wedge dx \wedge dz \\
d(-iE_y dt \wedge dy) &= i\partial_x E_y dt \wedge dx \wedge dy - i\partial_z E_y dt \wedge dy \wedge dz \\
d(-iE_z dt \wedge dz) &= i\partial_x E_z dt \wedge dx \wedge dz + i\partial_y E_z dt \wedge dy \wedge dz
\end{aligned}$$

Examining the $dt \wedge dy \wedge dz$ component:

$$i\,(\partial_y E_z - \partial_z E_y + \partial_t B_x)\,dt \wedge dy \wedge dz = i\,((\nabla \times \mathbf{E})_x + \partial_t B_x)$$

so we have:

$$(\nabla \times \mathbf{E})_x = -\partial_t B_x \tag{3.1.20}$$

Examining the $dt \wedge dx \wedge dz$ component:

$$i\,(\partial_x E_z - \partial_z E_x - \partial_t B_y)\,dt \wedge dx \wedge dz = i\,(-(\nabla \times \mathbf{E})_y - \partial_t B_y)\,dt \wedge dx \wedge dz$$

so we have:

$$(\nabla \times \mathbf{E})_y = -\partial_t B_y \tag{3.1.21}$$

Examining the $dt \wedge dx \wedge dy$ component:

$$i\,(\partial_x E_y - \partial_y E_x + \partial_t B_z)\,dt \wedge dx \wedge dy = i\,((\nabla \times \mathbf{E})_z + \partial_t B_z)\,dt \wedge dx \wedge dy$$

so we have:

$$(\nabla \times \mathbf{E})_z = -\partial_t B_z \tag{3.1.22}$$

Combining (3.1.20), (3.1.21), and (3.1.22) we obtain the second Maxwell equation:

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B}$$

implying the claim. □

3.1.4 Scalar Electrodynamics Formalism

We now wish to incorporate matter into our field theory. Let us first work less precisely; in the differential forms formalism of electromagnetism, one prescribes a current one form:

$$J = -\rho dt + j_x dx + j_y dy + j_z dz \tag{3.1.23}$$

where ρ is the charge density, and $\mathbf{j}$ is the current density. We necessitate that:

$$\begin{aligned} d \star J =& d\left(\rho dx \wedge dy \wedge dz - j_x dt \wedge dy \wedge dz + j_y dt \wedge dx \wedge dz - j_z dt \wedge dx \wedge dy\right) \\ =& (\partial_t \rho + \nabla \cdot \mathbf{j}) dt \wedge dx \wedge dy \wedge dz \\ =& 0 \end{aligned}$$

as this implies the continuity equation:

$$\nabla \cdot \mathbf{j} = -\partial_t \rho$$

One then writes that:

$$\mathscr{L}_{EM} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + J^\mu A_\mu \tag{3.1.24}$$

We notice that for any $\lambda : \mathbb{R}^{1,3} \to \mathbb{R}$:

$$\begin{aligned} \star d \star (J\lambda) &= \star\, d(\lambda \star J) \\ &= \star\,(d\lambda \wedge \star J) + \star(\lambda d \star J) \\ &= \star\,(\rho \partial_t \lambda + j_x \partial_x \lambda + j_y \partial_y \lambda + j_z \partial_z \lambda) dt \wedge dx \wedge dy \wedge z \\ &= -\, J^\mu \partial_\mu \lambda \end{aligned}$$

hence under the gauge transformation $A \to A + d\lambda$:

$$J^\mu (A_\mu + \partial_\mu \lambda) = J^\mu A_\mu - \star d \star (J\lambda)$$

Integrating over $\mathbb{R}^{1,3}$[37] we obtain by Stokes theorem:

$$\int_{\mathbb{R}^{1,3}} (J^\mu \partial_\mu \lambda) \mathrm{dvol}_g = -\int_{\mathbb{R}^{1,3}} \star d \star (J\lambda) \mathrm{dvol}_g = \int_{\mathbb{R}^{1,3}} d(\lambda \star J) = \int_{\partial \mathbb{R}^{1,3}} \lambda \star J = 0$$

since $\mathbb{R}^{1,3}$ has empty boundary. So this extra term changes nothing when we find the stationary points of the action, implying that this formalism is gauge invariant.

Ultimately, (3.1.24) is a flawed Lagrangian. Indeed, we are exploiting the fact that global sections of $\mathbb{R}^{1,3} \times U(1)$ exist, so any hope of writing such a Lagrangian on a non trivial principal bundle is immediately lost. Furthermore, the continuity equation should be a consequence of our field theory, not a priori. We fix this by introducing a "matter field" (i.e. a section of a vector bundle associated to some principal bundle P) which is then coupled to the gauge field via the exterior covariant derivative. The matter field will then partially determine the gauge field, and vice versa. But this is expected, as in electromagnetism, matter determines the electromagnetic potential, and the electromagnetic potential determines how matter moves. Furthermore, when we vary this new action, the matter field, and the gauge field A, will determine a current J which satisfies the continuity equation.

We now fix the following the data:

- An n-dimensional orientable (pseudo) Riemannian manifold (M, g)
- A principal G bundle over M with compact structure group G
- An Ad-invariant inner product on $\mathfrak{g}$, and an orthonormal basis T^i for $\mathfrak{g}$.
- A complex representation $\rho : G \to GL(W)$ with associated complex vector bundle $E = P \times_\rho W$.
- A Hermitian G invariant inner product $\langle \cdot, \cdot \rangle_W$ on W, with associated bundle metric $\langle \cdot, \cdot \rangle_E$.

[37] $\mathbb{R}^{1,3}$ is not compact, but we make the vague justification that our functions go to zero fast enough so that the integral converges.

Definition 3.1.11. If the dimension of W is one, then a smooth section of E is a **complex scalar field** and if the dimension of W is greater than one, then a smooth section of E is called a **multiplet of complex scalar fields**.

Furthermore, recall that we called non-trivial representations of $\mathfrak{g}$ charged; we have a similar definition for sections of E:

Definition 3.1.12. Sections Φ of an associated vector bundle $E = P \times_\rho W$, with:

$$\rho_* : \mathfrak{g} \longrightarrow \mathrm{End}(W)$$

are called **charged scalars**.

Indeed, the formalism of Electromagnetism we are about to derive is often called **scalar electrodynamics**, or **scalar QED**. Though we are treating this theory purely classically, we note that the Lagrangian we are about to write down corresponds to a relativistic quantum theory, i.e. a quantum field theory. We modify the Yang-Mill's Lagrangian in the following to incorporate such charged scalar fields:

Definition 3.1.13. The **Yang-Mills-Higgs Lagrangian** for a multiplet scalar field coupled to a gauge field is defined by:

$$\mathscr{L}_{YMH}[\Phi, A] = -\frac{1}{2}\langle F_M^A, F_M^A\rangle_{\mathrm{Ad}(P)} - V(\langle\Phi,\Phi\rangle_E) + \langle d_A\Phi, d_A\Phi\rangle_E$$

where $V : \mathbb{R} \to \mathbb{R}$ is a smooth function.

Theorem 3.1.6. *The Yang-Mills-Higgs Lagrangian is gauge invariant, i.e.*

$$\mathscr{L}_{YMH}[f^{-1}\cdot\Phi, f^*A] = \mathscr{L}_{YMH}[\Phi, A]$$

for all $f \in \mathscr{G}(P)$.

We need the following lemma:

Lemma 3.1.3. *Let $\Phi \in \Gamma(E)$, and $f \in \mathscr{G}(P)$, then:*

$$d_{f^*A}\left(f^{-1}\cdot\Phi\right) = f^{-1}\cdot(d_A\Phi)$$

Proof. First recall that there exists a $\sigma_f \in C^\infty(P, G)$ such that for all $p \in P$:

$$f(p) = p\cdot\sigma_f(p)$$

Since d_A is independent of choice of gauge, it suffices to check this locally. Let $s : U \to P_U$ be a local gauge, then:

$$f(s(x)) = s(x)\cdot\sigma_f(s(x))$$

Note that $\sigma_f(s(x))$ is a map $U \to G$, and thus a physical gauge transformation, which we now denote by h. Let $\phi : U \to W$ be the smooth map such that:

$$\Phi|_U = [s, \phi]$$

then by **Theorem 2.1.6**:

$$f^{-1}\cdot\Phi_U = [f^{-1}(s), \phi] = [s\cdot h^{-1}, \phi] = [s, \rho(h^{-1})\phi]$$

where $h^{-1} : U \to G$ denotes the map satisfying $h\cdot h^{-1} = e$. Furthermore, we have that by **Theorem 2.1.8:**

$$f^*A = \mathrm{Ad}_{\sigma_f^{-1}}\circ A + \sigma_f^*\mu_G$$

Thus in our local gauge:

$$(f^*A)_s = \mathrm{Ad}_{h^{-1}}\circ A_s + h^*\mu_G$$

For any vector field $X \in \mathfrak{X}(M)$:

$$d_{f^*A}(f^{-1}\cdot\Phi)(X)|_U = [s, d(\rho(h^{-1})\phi)(X) + \rho_*(\mathrm{Ad}_{h^{-1}})\circ A_s(X) + h^*\mu_G(X))\rho(h^{-1})\phi]$$

Our work from **Theorem 2.1.13** demonstrates that for all $x \in U$:

$$d(\rho(h^{-1})\phi)(X_x) = \rho(h^{-1})d\phi(X) - \rho_*(\mu_G(D_x h(X_x))) \cdot \rho(h^{-1})\phi \tag{3.1.25}$$

while:

$$\begin{aligned}\rho_*(\mathrm{Ad}_{h^{-1}} \circ A_s(X)+h^*\mu_G(X))\rho(h^{-1})\phi = \\ \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi + \rho_*(\mu_G(D_x h(X_x))) \cdot \rho(h^{-1})\phi\end{aligned} \tag{3.1.26}$$

Adding (3.1.25) and (3.1.26) we obtain:

$$\begin{aligned}d(\rho(h^{-1})\phi)(X) + \rho_*(\mathrm{Ad}_{h^{-1}} \circ A_s(X)+h^*\mu_G(X))\rho(h^{-1})\phi = \\ \rho(h^{-1})d\phi(X) + \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi\end{aligned}$$

hence:

$$\begin{aligned}d_{f^*A}(f \cdot \Phi)(X)|_U =& [s, \rho(h^{-1})d\phi(X) + \rho(h^{-1}) \cdot \rho_*(A_s(X_x))\phi] \\ =& [s \cdot h^{-1}, (d_A\phi)_s] \\ =& f^{-1} \cdot (d_A\Phi)|_U\end{aligned}$$

as desired. □

We can now prove **Theorem 3.1.6**:

Proof. Note that the Yang-Mill's part of the $\mathscr{L}_{YMH}$ is gauge invariant, so we need only argue the other terms are gauge invariant. By **Lemma 3.1.3** we have that:

$$\mathscr{L}_H[f^{-1} \cdot \Phi, f^*A] = -V(\langle f^{-1} \cdot \Phi, f^{-1} \cdot \Phi\rangle_E) + \langle f^{-1} \cdot d_A\Phi, f^{-1} \cdot d_A\Phi\rangle_E$$

Since $\langle \cdot, \cdot\rangle_E$ is an associated bundle metric, defined by the G invariant inner product $\langle \cdot, \cdot\rangle_W$, we have that:

$$\langle f^{-1} \cdot \Phi, f^{-1} \cdot \Phi\rangle_E = \langle \Phi, \Phi\rangle_E \qquad \text{and} \qquad \langle f^{-1} \cdot d_A\Phi, f^{-1} \cdot d_A\Phi\rangle_E = \langle d_A\Phi, d_A\Phi\rangle_E$$

hence:

$$\mathscr{L}_{YMH}[f^{-1} \cdot \Phi, f^*A] = \mathscr{L}_{YMH}[\Phi, A]$$

as desired. □

Definition 3.1.14. The **Yang-Mills-Higgs Action** is defined as:

$$S_{YMH}[\Phi, A] = \int_M \mathscr{L}_{YMH}[\Phi, A]\mathrm{dvol}_g$$

Note that since $\mathscr{L}_{YMH}$ depends on two fields, the gauge field, and a complex scalar field we can vary the action with respect to either field. We first vary the action with respect Φ.

Theorem 3.1.7. *In addition to the aforementioned fixed data, let (M, g) be a closed manifold, i.e. compact and without boundary. The variation of S_{YMH} with respect to Φ yields the following field equation:*

$$d_A^\star d_A\Phi = V'(\langle \Phi, \Phi\rangle_E)\Phi \tag{3.1.27}$$

where V' is the derivative of $V : \mathbb{R} \to \mathbb{R}$.

Proof. Let Ψ be any other section of E, then:

$$\mathscr{L}_{YMD}[\Phi + t\Psi, A] = \mathscr{L}_{YM} - V(\langle \Phi + t\Psi, \Phi + t\Psi\rangle_E) + \langle d_A(\Phi + t\Psi), d_A(\Phi + t\Psi)\rangle_E$$

We see first that:

$$\langle \Phi + t\Psi, \Phi + t\Psi\rangle_E = \langle \Phi, \Phi\rangle_E + t\langle \Phi, \Psi\rangle_E + t\langle \Psi, \Phi\rangle_E + \mathcal{O}(t^2)$$

hence:

$$\frac{d}{dt}\bigg|_{t=0}\langle\Phi+t\Psi,\Phi+t\Psi\rangle_E = \langle\Phi,\Psi\rangle_E+\langle\Psi,\Phi\rangle_E$$
$$=2\operatorname{Re}(\langle\Phi,\Psi\rangle_E)$$

So by the chain rule:

$$\frac{d}{dt}\bigg|_{t=0}V(\langle\Phi+t\Psi,\Phi+t\Psi\rangle_E)=2V'(\langle\Phi,\Phi\rangle_E)\operatorname{Re}(\langle\Phi,\Psi\rangle_E)$$

Furthermore, we see that:

$$\frac{d}{dt}\bigg|_{t=0}\langle d_A(\Phi+t\Psi),d_A(\Phi+t\Psi)\rangle_E = \langle d_A\Phi,d_A\Psi\rangle_E+\langle d_A\Psi,d_A\Phi\rangle_E$$
$$=2\operatorname{Re}(\langle d_A\Phi,d_A\Psi\rangle_E)$$

Therefore, by **Theorem 3.1.2**:

$$\begin{aligned}\frac{d}{dt}\bigg|_{t=0}S_{YMH}[\Phi+t\Psi,A] =& 2\int_M \operatorname{Re}(\langle d_A\Phi,d_A\Psi\rangle_E)-V'(\langle\Phi,\Phi\rangle_E)\operatorname{Re}(\langle\Phi,\Psi\rangle_E)\mathrm{dvol}_g\\ =&2\operatorname{Re}\left(\int_M\langle d_A\Phi,d_A\Psi\rangle_E+V'(\langle\Phi,\Phi\rangle_E)\langle\Phi,\Psi\rangle_E\mathrm{dvol}_g\right)\\ =&2\operatorname{Re}\left(\int_M\langle d_A^\star d_A\Phi,\Psi\rangle_E-V'(\langle\Phi,\Phi\rangle_E)\langle\Phi,\Psi\rangle_E\mathrm{dvol}_g\right)\\ =&2\operatorname{Re}\left(\int_M\langle d_A^\star d_A\Phi-V'(\langle\Phi,\Phi\rangle_E)\Phi,\Psi\rangle_E\mathrm{dvol}_g\right)\end{aligned}$$

Since the L^2 inner product of twisted forms is non degenerate, it follows that if Φ is a critical point of S_{YMH} then:

$$d_A^\star d_A\Phi-V'(\langle\Phi,\Phi\rangle_E)\Phi=0$$

i.e.,

$$d_A^\star d_A\Phi=V'(\langle\Phi,\Phi\rangle_E)\Phi$$

as desired. □

To vary the action with respect to A we will need the following two lemmas:

Lemma 3.1.4. *Let $\alpha_M\in\Omega^1(M,Ad(P))$, and $\Phi\in\Gamma(E)$, then there exists a canonical twisted one form $\alpha_M\cdot\Phi\in\Omega^1(M,E)$.*

Proof. Note that $E=P\times_\rho W$ comes equipped with a representation of G on W; we will use the induced representation ρ_* of $\mathfrak{g}$ on W to define this form. First, note that for all $x\in M$, we have that for some $w\in W$:

$$\Phi(x)=[p,w]$$

where $p\in P_x$. Furthermore, we have that there exists a unique $\alpha\in\Omega^1_{\text{hor}}(P,\mathfrak{g})^{\text{Ad}}$ such that for all $X\in T_xM$:

$$\alpha_{Mx}(X)=[p,\alpha_p(Y)]$$

where $\pi(p)=x$ and $\pi_*Y=X$. We thus define $\alpha_M\cdot\Phi$ point wise by:

$$(\alpha_M\cdot\Phi)_x(X)=[p,\rho_*(\alpha_p(Y))\cdot w] \tag{3.1.28}$$

This is independent of our choice of Y by our work in **Theorem 2.1.18**, and it is independent of our choice of p as for any $q=p\cdot g$ we have that:

$$\begin{aligned}(\alpha_M\cdot\Phi)_x(X)=&[p\cdot g,\rho_*(\alpha_{p\cdot g}(Y))\cdot\rho(g^{-1})w]\\ =&[p\cdot g,\rho_*(\mathrm{Ad}_{g^{-1}}\circ\alpha_p(Y))\cdot\rho(g^{-1})w]\\ =&[p\cdot g,\rho(g^{-1})\cdot\rho_*(\alpha_p(Y))\cdot w]\\ =&[p,\rho_*(\alpha_p(Y))\cdot w]\end{aligned}$$

hence (3.1.27) is well defined. Finally, let $s : U \to P$ be a local gauge, and $\phi : U \to W$ be a smooth map, such that:

$$\Phi = [s, \phi]$$

Denote $s^*\alpha$ by α_s, then for all $X \in \mathfrak{X}(M)$:

$$(\alpha_M \cdot \Phi)(X) = [s, \rho_*(\alpha_s(X)) \cdot \phi]$$

which is smooth section of E, hence $\alpha_M \cdot \Phi \in \Omega^1(M, E)$, as desired. $\square$

Lemma 3.1.5. *There exists a unique twisted one form $J_H(A, \Phi) \in \Omega^1(M, Ad(P))$ such that for all $\alpha_M \in \Omega^1(M, Ad(P))$:*

$$\langle \alpha_M, J_H(A, \Phi)\rangle_{Ad(P)} = 2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E)$$

Proof. Suppose that J_H exists and is not unique. Then there exists an $\omega \in \Omega^1(M, \mathrm{Ad}(P))$ such that for all $\alpha_M \in \Omega^1(M, \mathrm{Ad}(P))$:

$$\langle \alpha_M, \omega\rangle_{\mathrm{Ad}(P)} = 2\,\mathrm{Re}\,(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E)$$

Then we see that for all α_M:

$$\begin{aligned}\langle \alpha_M, J_H - \omega\rangle_{\mathrm{Ad}(P)} =&\langle \alpha_M, J_H\rangle_{\mathrm{Ad}(P)} - \langle \alpha_M, \omega\rangle_{\mathrm{Ad}(P)}\\ =&2\,\mathrm{Re}\,(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E) - 2\,\mathrm{Re}\,(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E)\\ =&0\end{aligned}$$

Since $\langle \cdot, \cdot\rangle_{\mathrm{Ad}(P)}$ is nondegenerate, it follows that $J_H = \omega$, hence J_H is uniquely determined.

We now wish to prove existence. Note that:

$$2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E) =\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E + \langle \alpha_M \cdot \Phi, d_A\Phi\rangle_E$$

Furthermore, the assignment:

$$\begin{aligned}\Lambda : \Omega^1(M, \mathrm{Ad}(P)) &\longrightarrow C^\infty(M)\\ \alpha_M &\longmapsto \langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E + \langle \alpha_M \cdot \Phi, d_A\Phi\rangle_E\end{aligned}$$

is clearly a $C^\infty(M)$ linear map, thus Λ is a global section of $TM \otimes \mathrm{Ad}(P)^*$, where $\mathrm{Ad}(P)^*$ is bundle dual to $\mathrm{Ad}(P)$. Since the bundle metric $\langle \cdot, \cdot\rangle_{\mathrm{Ad}(P)}$ on $T^*M \otimes \mathrm{Ad}(P)$ is non degenerate, it follows that it induces a bundle isomorphism:

$$F : T^*M \otimes \mathrm{Ad}(P) \longrightarrow TM \otimes \mathrm{Ad}(P)^*$$

that satisfies:

$$F(\omega)(\eta) = \langle \omega, \eta\rangle_{\mathrm{Ad}(P)}$$

for all $\omega, \eta \in \Omega^1(M, \mathrm{Ad}(P))$. Setting $J_H = F^{-1}(\Lambda)$, then implies the claim as for all $\alpha_M \in \Omega^1(M, \mathrm{Ad}(P))$:

$$\begin{aligned}\langle J_H, \alpha_M\rangle_{\mathrm{Ad}(P)} =&F(J_H)(\alpha_M)\\ =&\Lambda(\alpha_M)\\ =&2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E)\end{aligned}$$

$\square$

The one form J_H is the analogue of the current one form mentioned earlier, and it is entirely determined by the fields A and Φ. To see this explicitly, we wish to find a local expression for J_H. Let $s : U \to P_U$ be a local gauge, and e_i be an orthonormal frame for E_U, which for some orthonormal basis $\{w_i\}$ of W can be written as:

$$e_i = [s, w_i]$$

We then see that for some smooth functions $\phi^i : U \to \mathbb{C}$:

$$\Phi = [s, \phi^i w_i]$$

The exterior covariant derivative is then given by:

$$\begin{aligned} d_A\Phi =&[s, d\phi^i \otimes w_i + \phi^i \rho_*(A_s)w_i] \\ =&[s, \partial_\mu \phi^i dx^\mu \otimes w_i + \phi^i A^a_\mu dx^\mu \otimes \rho_*(T_a)w_i] \end{aligned}$$

Furthermore,

$$\alpha_M \cdot \Phi = [s, \phi^i \alpha^a_\mu dx^\mu \otimes \rho_*(T_a)w_i]$$

Hence:

$$\begin{aligned} \langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E =&\phi^j(\partial_\mu \bar{\phi}^i)\alpha^b_\nu \langle dx^\mu, dx^\nu\rangle \langle w_i, \rho_*(T_b)w_j\rangle_W + \bar{\phi}^i \bar{A}^a_\mu \phi^j \alpha^b_\nu \langle dx^\mu, dx^\nu\rangle\langle \rho_*(T_a)w_i, \rho_*(T_b)w_j\rangle_W \\ =&\phi^j(\partial_\mu \bar{\phi}^i)\alpha^{\mu b}\langle w_i, \rho_*(T_b)w_j\rangle_W + \bar{\phi}^i \bar{A}^a_\mu \phi^j \alpha^{\mu b}\langle \rho_*(T_a)w_i, \rho_*(T_b)w_j\rangle_W \end{aligned}$$

We see that $\rho_*(T_a)$ is an endomorphism of finite dimensional vector spaces, hence:

$$\rho_*(T_a) = \Gamma^k_{al} w_k \otimes w^l$$

where w^l is the covector dual to w_l determined by $\langle\cdot,\cdot\rangle_W$, and each $\Gamma^k_{al} \in \mathbb{C}$. We see that:

$$\rho_*(T_a)w_i = \Gamma^k_{ai} w_k$$

so:

$$\begin{aligned} \langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E =&\phi^j(\partial_\mu \bar{\phi}^i)\alpha^{\mu b}\langle w_i, \Gamma^l_{bj} w_l\rangle_W + \bar{\phi}^i \bar{A}^a_\mu \phi^j \alpha^{\mu b}\langle \Gamma^m_{ai} w_m, \Gamma^l_{bj} w_l\rangle_W \\ =&\phi^j(\partial_\mu \bar{\phi}^i)\alpha^{\mu b}\Gamma^i_{bj} + \bar{\phi}^i \bar{A}^a_\mu \alpha^{\mu b}\phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} \end{aligned}$$

implying that:

$$2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E) =\phi^j(\partial_\mu \bar{\phi}^i)\alpha^{\mu b}\Gamma^i_{bj} + \bar{\phi}^j(\partial_\mu \phi^i)\bar{\alpha}^{\mu b}\bar{\Gamma}^i_{bj} + \bar{\phi}^i \bar{A}^a_\mu \alpha^{\mu b}\phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} + \phi^i A^a_\mu \bar{\alpha}^{\mu b}\bar{\phi}^j \Gamma^m_{ai}\bar{\Gamma}^m_{bj}$$

where there is still a summation over i and m despite both indices being upper. Note that we assume $\mathfrak{g}$ to be a real Lie algebra, hence:

$$\bar{A}^a_\mu = A^a_\mu \qquad \text{and} \qquad \bar{\alpha}^{\mu b} = \alpha^{\mu b}$$

With the above in mind:

$$2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M \cdot \Phi\rangle_E) =(\phi^j(\partial_\mu \bar{\phi}^i)\Gamma^i_{bj} + \bar{\phi}^j(\partial_\mu \phi^i)\bar{\Gamma}^i_{bj})\alpha^{\mu b} + (\bar{\phi}^i A^a_\mu \phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} + \phi^i \bar{\phi}^j A^a_\mu \Gamma^m_{ai}\bar{\Gamma}^m_{bj})\alpha^{\mu b}$$

hence we define:

$$J_H = [s, (\phi^j(\partial_\mu \bar{\phi}^i)\Gamma^i_{bj} + \bar{\phi}^j(\partial_\mu \phi^i)\bar{\Gamma}^i_{bj} + \bar{\phi}^i A^a_\mu \phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} + \phi^i \bar{\phi}^j A^a_\mu \Gamma^m_{ai}\bar{\Gamma}^m_{bj})dx^\mu \otimes T_b] \tag{3.1.29}$$

where there is an implied summation over the index b, despite it being lower. Note that the above is a one form written in terms of the local gauge field. For any:

$$\alpha_M = [s, \alpha^c_\nu dx^\nu \otimes T_c]$$

we then have that:

$$\begin{aligned} \langle \alpha_M, J_H\rangle_{\mathrm{Ad}(P)} =&\langle J_h, \alpha_M\rangle_{\mathrm{Ad}(P)} \\ =&(\phi^j(\partial_\mu \phi^i)\Gamma^i_{bj} + \bar{\phi}^j(\partial_\mu \phi^i)\bar{\Gamma}^i_{bj} + \bar{\phi}^i A^a_\mu \phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} + \phi^i \bar{\phi}^j A^a_\mu \Gamma^m_{ai}\bar{\Gamma}^m_{bj})\alpha^c_\nu \langle dx^\mu, dx^\nu\rangle\langle T_b, T_c\rangle_{\mathfrak{g}} \\ =&(\phi^j(\partial_\mu \phi^i)\Gamma^i_{bj} + \bar{\phi}^j(\partial_\mu \phi^i)\bar{\Gamma}^i_{bj})\alpha^{\mu b} + (\bar{\phi}^i A^a_\mu \phi^j \bar{\Gamma}^m_{ai}\Gamma^m_{bj} + \phi^i \bar{\phi}^j A^a_\mu \Gamma^m_{ai}\bar{\Gamma}^m_{bj})\alpha^{\mu b} \end{aligned}$$

So (3.1.29) is a local expression for J_H as desired.

Theorem 3.1.8. *Let (M, g) be a closed (pseudo)-Riemannian manifold. The variation of S_{YMH} with respect to the connection A yields the following field equation:*

$$d_A^\star F_M^A = J_H(A, \Phi) \tag{3.1.30}$$

Proof. Recall that:

$$L_{YMH}[\Phi, A] = -\frac{1}{2}\langle F_M^A, F_M^A\rangle_{\mathrm{Ad}(P)} - V(\langle\Phi,\Phi\rangle_E) + \langle d_A\Phi, d_A\Phi\rangle_E$$

Note that the potential V will go to zero when we vary the action with respect to A, so we elect to ignore it. Let $\alpha \in \Omega^1_{\mathrm{hor}}(P,\mathfrak{g})^{\mathrm{Ad}}$, then for any local gauge $s : U \to P_U$, and and smooth map $\phi : U \to W$ such that:

$$\Phi = [s, \phi]$$

we have that by **Lemma 3.1.4**:

$$\begin{aligned} d_{A+t\alpha}\Phi =& [s, d\phi + \rho_*(A_s)\phi + t\rho_*(\alpha_s)\phi] \\ =& [s, d\phi + \rho_*(A_s)\phi] + t[s, \rho_*(\alpha_s)\phi] \\ =& d_A\Phi + t\alpha_M \cdot \Phi \end{aligned}$$

Therefore:

$$\langle d_{A+t\alpha}\Phi, d_{A+t\alpha}\Phi\rangle_E = \langle d_A\Phi, d_A\Phi\rangle_E + 2t\,\mathrm{Re}(\langle d_A\Phi, \alpha_M\cdot\Phi\rangle_E) + \mathcal{O}(t^2)$$

By **Lemma 3.1.5**:

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0}\langle d_{A+t\alpha}\Phi, d_{A+t\alpha}\Phi\rangle_E =& 2\,\mathrm{Re}(\langle d_A\Phi, \alpha_M\cdot\Phi\rangle_E) \\ =& \langle\alpha_M, J_H(A,\Phi)\rangle_{\mathrm{Ad}(P)} \end{aligned}$$

From our work **Theorem 3.1.4** we then find that:

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} S_{YMH} &= \int_M \langle\alpha_M, J_H(A,\Phi)\rangle_{\mathrm{Ad}(P)} - \langle\alpha_M, d_A^\star F_M^A\rangle_{\mathrm{Ad}(P)}\mathrm{dvol}_g \\ &= \int_M \langle\alpha_M, J_H(A,\Phi) - d_A^\star F_M^A\rangle_{\mathrm{Ad}(P)}\mathrm{dvol}_g \end{aligned}$$

Since the L_2 inner product is nondegenerate, we have that if A is a stationary point of S_{YMH} then:

$$J_H(A,\Phi) - d_A^\star F_M^A = 0 \Rightarrow d_A^\star F_M^A = J_H(A,\Phi)$$

as desired. □

Equations (3.1.30) and (3.1.27) are often referred to as the **Yang-Mills-Higgs equations**, and they play a crucial role in studying the Higgs mechanism of the Standard Model. For our purposes however, we will use them to build a theory of scalar electrodynamics.

We now go back to the case where $P = \mathbb{R}^{1,3}\times U(1)$, but in addition we set $W = \mathbb{C}$, and choose the representation of $U(1)$ on $\mathbb{C}$ given by standard complex multiplication. This implies that:

$$E = P\times_\rho \mathbb{C} = \mathbb{R}^{1,3}\times\mathbb{C}$$

as every vector bundle over $\mathbb{R}^n$ is trivial. Furthermore, we now choose to make explicit the imaginary character of the connection one form and the curvature, i.e. we write in some global gauge:

$$iA_s = i(-V dt + M_i dx^i) \qquad \text{and} \qquad iF_s = i(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu\wedge dx^\nu$$

Hence for $\phi\in\Gamma(E)$ we write:

$$(d_A\phi)_\mu = \partial_\mu\phi + iA_\mu\phi$$

and for $\bar{\phi} \in \Gamma(\bar{E})$ we have:

$$(d_A\bar{\phi})_\mu = \partial_\mu\bar{\phi} - iA_\mu\bar{\phi}$$

where the induced representation of ρ_* on $\bar{E}$ is given by conjugation. We also fix the following inner product induced by the standard Hermitian inner product on $\mathbb{C}$:

$$\langle \phi, \psi \rangle_E = \bar{\phi}\psi$$

for all $\phi, \psi \in \Gamma(E)$. Finally, we fix the following potential:

$$V(\langle \phi, \phi \rangle_E) = -m^2|\phi|^2 = -m^2\bar{\phi} \cdot \phi$$

for some positive constant m, which can be interpreted as the rest mass of ϕ.

Theorem 3.1.9. *A complex scalar field ϕ which leaves L_{YMH} stationary satisfies the following field equation:*

$$d_A^\star d_A \phi = -m^2\phi \tag{3.1.31}$$

Given a global gauge, and the global coordinates (t, x, y, z) we can write this as:

$$\left[(\partial_\mu + iA_\mu)(\partial^\mu + iA^\mu) - m^2\right]\phi = 0$$

Proof. We see that by **Theorem 3.1.7**:

$$d_A^\star d_A \phi = -m^2\phi$$

Fix a global gauge s, and the global coordinates (t, x, y, z):

$$\begin{aligned}(d_A\phi)_s =& d\phi + iA_s\phi \\ =& (\partial_\mu\phi + iA_\mu\phi)dx^\mu\end{aligned} \tag{3.1.32}$$

With $n = 4$ and $k = t = 1$ we have that by **Definition 3.1.5**:

$$d_A^\star = \star d_A \star$$

Taking the Hodge star of (3.1.32):

$$\begin{aligned}\star(d_A\phi)_s =& -(\partial_t\phi + iA_t\phi)dx \wedge dy \wedge dz - (\partial_x\phi + iA_x\phi)dt \wedge dy \wedge dz \\ &+ (\partial_y\phi + iA_y\phi)dt \wedge dx \wedge dz - (\partial_z\phi + iA_z\phi)dt \wedge dx \wedge dy\end{aligned}$$

Applying the exterior covariant derivative:

$$(d_A \star d_A\phi)_s = d(\star d_A\phi)_s + iA_s \wedge (\star d_A\phi)_s$$

We will calculate each term separately, beginning with the exterior derivative:

$$\begin{aligned}d(\star d_A\phi)_s =& \left[-(\partial_t^2\phi + i\partial_t A_t\phi) + (\partial_x^2\phi + i\partial_x A_x\phi)\right. \\ &\left. + (\partial_y^2\phi + i\partial_y A_y\phi) + (\partial_z^2\phi + i\partial_z A_z\phi)\right] dt \wedge dx \wedge dy \wedge dz \\ =& (\partial_\mu\partial^\mu\phi + i\partial_\mu(A^\mu\phi))\, dt \wedge dx \wedge dy \wedge dz\end{aligned}$$

Calculating the second term:

$$\begin{aligned}iA_s \wedge (\star d_A\phi)_s =& \langle -iA_s, d_A\phi\rangle dt \wedge dx \wedge dy \wedge dz \\ =& iA_\mu(\partial^\mu\phi + iA^\mu\phi)dt \wedge dx \wedge dy \wedge dz\end{aligned}$$

Taking the Hodge star of both terms we obtain:

$$\begin{aligned}(d_A^\star d_A\phi)_s =& -(\partial_\mu\partial^\mu\phi + i\partial_\mu(A^\mu)\phi + 2iA^\mu\partial_\mu\phi - A^\mu A_\mu\phi) \\ =& -(\partial_\mu + iA_\mu)(\partial^\mu + iA^\mu)\phi\end{aligned}$$

hence:

$$\left[(\partial_\mu + iA_\mu)(\partial^\mu + iA^\mu) - m^2\right]\phi = 0$$

as desired. □

We have already calculated a local expression for the current one form $J_H(A, \Phi)$, however, in the $P = \mathbb{R}^{1,3} \times U(1)$ case, the unwieldy expression given by (3.1.29) reduces tremendously. Indeed, since $\phi \in C^\infty(\mathbb{R}^{1,3}, \mathbb{C})$, and:

$$\rho_*(i\alpha_M)\phi = i\alpha_M\phi$$

for all $i\alpha_M \in \Omega^1(\mathbb{R}^{1,3}, i\mathbb{R})$, we have no need to introduce the gamma coefficients, as everything is just scalar multiplication. With this in mind we see that:

$$\langle d_A\phi, i\alpha_M \cdot \phi\rangle_E = i\phi(\partial_\mu\bar{\phi})\alpha^\mu + A_\mu\alpha^\mu\bar{\phi}\phi$$

hence:

$$\begin{aligned} 2\,\mathrm{Re}\left(\langle d_A\phi, i\alpha_M \cdot \phi\rangle_E\right) =& i\phi(\partial_\mu\bar{\phi})\alpha^\mu + 2A_\mu\alpha^\mu\bar{\phi}\phi - i\bar{\phi}(\partial_\mu\phi)\alpha^\mu \\ =& \left(i\phi(\partial_\mu\bar{\phi}) + 2A_\mu\bar{\phi}\phi - i\bar{\phi}(\partial_\mu\phi)\right)\alpha^\mu \end{aligned}$$

implying that the coefficients of iJ_H are given by:

$$\begin{aligned} J_\mu =& i\phi(\partial_\mu\bar{\phi}) + 2A_\mu\bar{\phi}\phi - i\bar{\phi}(\partial_\mu\phi) \\ =& i\phi\left(\partial_\mu\bar{\phi} - iA_\mu\bar{\phi}\right) - i\bar{\phi}\left(\partial_\mu\phi + iA_s\phi\right) \\ =& i\left(\phi(d_A\bar{\phi})_\mu - \bar{\phi}(d_A\phi)_\mu\right) \end{aligned}$$

Each J_μ is real, as is easily verified, so:

$$iJ_H = iJ_\mu dx^\mu \in \Omega^1(\mathbb{R}^{1,3}, i\mathbb{R})$$

For our purposes, we have little need to further deal with the complex field ϕ, hence we set:

$$J = J_\mu dx^\mu = -\rho dt + j_x dx + j_y dy + j_z dz$$

where we are interpreting J_t as the charge density ρ, and J_i as the components of the current density $\mathbf{j}$.

Corollary 3.1.1. *The current one form $iJ_H(A, \phi) \in \Omega^1(\mathbb{R}^{1,3}, i\mathbb{R})$ satisfies the continuity equation:*

$$d \star J_H = 0$$

Proof. First recall that for any $i\alpha_M \in \Omega^k(\mathbb{R}^{1,3}, i\mathbb{R})$:

$$d_A(i\alpha_M) = id\alpha_M$$

as $U(1)$ is an abelian Lie group. Therefore, by **Theorem 3.1.8**

$$d^\star iF_M^A = iJ_H$$

Take the Hodge star of both sides to obtain:

$$\star\star(d \star iF_M^A) = \star(J_H) \Longrightarrow d \star (iF_M^A) = \star(J_H)$$

Then, since $d \circ d = 0$:

$$d \star (iJ_H) = 0$$

as desired. $\square$

The corollary above gives us exactly what we had set out to obtain: a continuity equation that is a consequence of our theory, not a priori. Furthermore, we can extract our initial continuity equation by noting that:

$$d \star (iJ_H) = i\,(d \star J) = 0 \Longrightarrow d \star J = 0$$

We end the section by demonstrating that this Lagrangian reproduces Maxwell's fields equations due to a source.

Theorem 3.1.10. *With $P = \mathbb{R}^{1,3} \times U(1)$, and $E = \mathbb{R}^{1,3} \times \mathbb{C}$, the Bianchi identity, and the second Yang-Mills-Higgs equation:*

$$d(iF_M^A) = 0 \qquad \text{and} \qquad d^\star(iF_M^A) = iJ_H(A, \Phi)$$

are equivalent to Maxwell's field equations in a vacuum with an arbitrary source:

$$\begin{aligned} \nabla \cdot \mathbf{E} =& \rho \qquad \nabla \times \mathbf{E} = && -\partial_t \mathbf{B} \\ \nabla \cdot \mathbf{B} =& 0 \qquad \nabla \times \mathbf{B} = && \mathbf{j} + \partial_t \mathbf{E} \end{aligned}$$

Proof. Let s be a global gauge, and F_s be the global curvature form on $\mathbb{R}^{1,3}$ with components in the standard coordinates given by **Proposition 3.1.4**. Our work in **Theorem 3.1.5** already tells us that the Bianchi identity implies:

$$\nabla \times \mathbf{E} = -\,\partial_t \mathbf{B} \qquad \text{and} \qquad \nabla \cdot \mathbf{B} = 0$$

Furthermore:

$$\star d^\star(iF) = \star \star d \star (iF) = d \star (iF)$$

Hence we wish to show that:

$$d \star (iF) = \star(iJ)$$

implies the other two Maxwell equations. From our work at the beginning of the section, we have that:

$$\star(iJ) = i\left(\rho dx \wedge dy \wedge dz - j_x dt \wedge dy \wedge dz + j_y dt \wedge dx \wedge dz - j_z dt \wedge dx \wedge dy\right) \tag{3.1.33}$$

While our work form **Theorem 3.1.5** implies that:

$$\left(d \star (iF)\right)_{xyz} = i\nabla \cdot \mathbf{E} dx \wedge dy \wedge dz$$

Comparing the above with (3.1.33) we have that:

$$\nabla \cdot \mathbf{E} = \rho$$

which is the first Maxwell equation. Furthermore, we have:

$$\begin{aligned} \left(d \star (iF)\right)_{tyz} =& i\left(-(\nabla \times \mathbf{B})_x + \partial_t E_x\right) = -ij_x \\ \left(d \star (iF)\right)_{txz} =& i\left((\nabla \times \mathbf{B})_y - \partial_t E_y\right) = ij_y \\ \left(d \star (iF)\right)_{txy} =& i\left(-(\nabla \times \mathbf{B})_z + \partial_t E_z\right) = -ij_z \end{aligned}$$

which implies the fourth Maxwell equation:

$$\nabla \times \mathbf{B} = \mathbf{j} + \partial_t \mathbf{E}$$

as desired. □

3.1.5 QED Formalism

Now that we understand how scalar matter fields are incorporated to our field theory, we are ready to explore how to write field theories which incorporate fermions, i.e. spinor fields. It is important to note that spin is a completely quantum phenomenon, but we can offer a mathematical heuristic in the classical limit. For integer spin, mathematically, all spin refers to is how the fields isometries on $\mathbb{R}^{t,s}$. For example spin 0 particles are scalar fields, whose value at any point in space time can't depend on the chosen coordinates. Furthermore, spin 1 particles are one forms (or vector fields), and transform covariantly (contravariantly) under a change of coordinates. Higher order integer spin n particles, then transform like rank $(0, n)$ (or $(n, 0)$) tensor fields.

The case of half integer spin particles is mildly more intricate, and related to the spinor representation of the group $\text{Spin}^+(t, s)$ on Δ_{t+s}. For example, results from physics demonstrate that fermions[38], particles which have spin 1/2, are spinor fields, so sections of $S = \mathrm{Spin}^+(\mathbb{R}^{t,s}) \times_\kappa \Delta_{t+s}$. Then since any bundle over $\mathbb{R}^{t,s}$ is trivial, we have the any spinor field Ψ can be written as:

$$\Psi(x) = (x, \psi(x))$$

for some smooth map $\psi : \mathbb{R}^{t,s} \to \Delta_{t+s}$. Spin 1/2 particles then transform under the spin representation by:

$$\psi(x) \longmapsto \psi'(x) = \kappa(g) \cdot \psi(\lambda(g)^{-1}x)$$

Note that this is a different transformation property than the bundle automorphism on S induced by $f \in \mathscr{G}(\mathrm{Spin}^+(\mathbb{R}^{t,s}))$. Indeed we have that any $g \in \mathrm{Spin}^+(t,s)$ induces a transformation on the base space as well given by the double covering homomorphism, so $\mathrm{Spin}^+(t,s)$ can act on $\mathbb{R}^{t,s}$, as well as the spinor part of the field. This is clearly necessary given our discussion in the previous paragraph on integer spin.

Higher order half integer spin $n/2$ particles can then be obtained by the induced spinor representation of $\mathrm{Spin}^+(t,s)$ on tensor products of Δ_{t+s}, however no higher order half integer spin particles are known to exist. It follows that integer spin particles are fields which transform under the induced standard representations of $SO^+(t,s)$, and half integer spin fields transform under the induced spinor representations of $\mathrm{Spin}^+(t,s)$. When M is not $\mathbb{R}^{t,s}$, but instead an arbitrary pseudo Riemannian spin manifold, the above discussion only holds locally.

Obviously, to incorporate spinor fields into a Lagrangian, we need extra constraints on the base manifold M, implying that such Lagrangians are harder to come by than the previous two we have studied. We thus fix the following data:

- An n dimensional (pseudo) Riemannian spin manifold (M, g).
- A spin structure $\mathrm{Spin}^+(M)$ on M.
- A principal G bundle over M with compact structure group G.
- An Ad-invariant inner product on $\mathfrak{g}$, and an orthonormal basis T^i for $\mathfrak{g}$.
- A complex representation $\rho : G \to GL(V)$ with associated complex vector bundle $E = P \times_\rho V$.
- A Hermitian G invariant inner product $\langle \cdot, \cdot \rangle_V$ on V, with associated bundle metric $\langle \cdot, \cdot \rangle_E$
- A Dirac form $\langle \cdot, \cdot \rangle$ on the Dirac spinor space Δ_n with Dirac bundle metric $\langle \cdot, \cdot \rangle_S$.

We note that on the twisted spinor bundle $S \otimes E$, we can obtain a Hermitian scalar product $\langle \cdot, \cdot \rangle_{S\otimes E}$ given for all $\Psi, \Phi \in (S \otimes E)_x$ at all $x \in M$ by:

$$\langle \Psi, \Phi \rangle_{S\otimes E} = \langle \Psi^i, \Phi^j \rangle_S \cdot \langle \tau_i, \tau_j \rangle_E$$

where τ_i is any basis for E_x. We also set the following notation:

$$\bar{\Psi}\Phi = \langle \Psi, \Phi \rangle_{S\otimes E}$$

Let $\epsilon \times_M s$ be a local section $U \to S \otimes E_U$, and let τ_i be the local orthonormal frame of $S \otimes E_U$, then, we can write that:

$$\bar{\Psi}\Phi = \sum_i \bar{\Psi}^i \Phi^i$$

Definition 3.1.15. The **Yang-Mills-Dirac Lagrangian** for a twisted spinor field of mass m coupled to a gauge field is:

$$\begin{aligned}\mathscr{L}_{YMD}[\Psi, A] &= -\frac{1}{2}\langle F_M^A, F_M^A \rangle_{\mathrm{Ad}(P)} - m\langle \Psi, \Psi \rangle_{S\otimes E} + \mathrm{Re}\left(\langle \Psi, D_A \Psi \rangle_{S\otimes E}\right) \\ &= -\frac{1}{2}\langle F_M^A, F_M^A \rangle_{\mathrm{Ad}(P)} - m\bar{\Psi}\Psi + \mathrm{Re}\left(\bar{\Psi} D_A \Psi\right)\end{aligned}$$

Just as before, this Lagrangian is gauge invariant:

[38] In full generality, fermions must be elements of $S \otimes E$ so that they can be coupled to a gauge field, but in a free theory of fermions the discussion holds. Generalizing this statement to section of $S \otimes E$ is not difficult.

Theorem 3.1.11. *The Yang-Mills-Dirac Lagrangian satisfies:*

$$\mathscr{L}_{YMD}[f^{-1}\cdot\Phi, f^*A] = \mathscr{L}_{YMD}[\Phi, A]$$

for all $f\in\mathscr{G}(P)$.

Proof. Note that Yang-Mills, and mass terms, are automatically gauge invariant by our work in previous sections. We thus need to show that:

$$D_{f^*A}(f^{-1}\Psi) = f^{-1}D_A\Psi$$

We see that since $f\in\mathscr{G}(P)$, f acts on $\mathrm{Spin}^+(M)\times P$ by:

$$f((p,q)) = (p, f(q)) = (p, q\cdot\sigma_f(q)) = (p,q)\cdot(e,\sigma_f(q))$$

Since D_A is independent of our choice of gauge, it suffices to check this locally. Let $\epsilon\times_M s: U\to (\mathrm{Spin}^+(M)\times_M P)_U$ be a local gauge, then for some basis v_i of V, and smooth maps $\psi^i: U\to\Delta_n$, we can write:

$$\Psi|_U = [\epsilon\times_M s, \psi^i\otimes v_i]$$

Then:

$$f(\epsilon\times_M s(x)) = (\epsilon\times_M s(x))\cdot(e,\sigma(s(x)))$$

Denote the physical gauge transformation $\sigma(s(x)): U\to G$ by h. Then by **Theorem 2.1.6**:

$$\begin{aligned} f^{-1}\cdot\Psi_U =&[\epsilon\times_M s, (\kappa\otimes\rho)(e,h^{-1})\psi^i\otimes v_i] \\ =&[\epsilon\times_M s, \psi^i\otimes\rho(h^{-1})v_i] \end{aligned}$$

By **Theorem 2.2.18**, and our work in **Lemma 3.1.3**, we have that:

$$(f^*A)_s = \mathrm{Ad}_{h^{-1}}\circ A_s + h^*\mu_G$$

We thus have for any local oriented and time oriented orthonormal frame e_a of TM_U that:

$$\begin{aligned} D_{f^*A}(f^{-1}\cdot\Psi)|_U =&D_{f^*A}\left[\epsilon\times_M s, \psi^i\otimes\rho(h^{-1})\phi\right] \\ =&\left[\epsilon\times_M s, D_{f^*A}\psi^i\otimes\rho(h^{-1})\phi\right] \end{aligned}$$

We have that:

$$\begin{aligned} &D_{f^*A}(\psi^i\otimes\rho(h^{-1})v_i) = \\ &\quad i\Gamma^a\left(d(\psi^i\otimes\rho(h^{-1})v_i(e_a)) - \frac{1}{4}\xi_{bc}(e_a)\Gamma^{bc}\psi^i\otimes\rho(h^{-1})v_i + \psi^i\otimes\rho_*(\mathrm{Ad}_{h^{-1}}\circ A_s(e_a) + h^*\mu_G(e_a))\rho(h^{-1})v_i\right) \end{aligned}$$

We note that by our work in **Lemma 3.1.3**, for all $x\in U$ we have:

$$d(\psi^i\otimes v_i)(e_a) = d\psi^i(e_a)\otimes\rho(h^{-1})v_i + \psi^i\otimes\rho_*(\mu_G(D_x h(e_a)))\cdot\rho(h^{-1})v_i$$

We can thus split $D_{f^*A}(\psi^i\otimes\rho(h^{-1}v_i))$ into a "spin" part:

$$\Omega_S = i\Gamma^a\left(d\psi^i(e_a) - \frac{1}{4}\xi_{bc}(e_a)\Gamma^{bc}\psi^i\right)\otimes\rho(h^{-1})v_i$$

and a gauge part:

$$\Omega_A = i\Gamma^a\psi^i\otimes\left(\rho_*(\mu_G(D_x h(e_a))) + \rho_*(\mathrm{Ad}_{h^{-1}}\circ A_s(e_a) + h^*\mu_G(e_a))\right)\rho(h^{-1})v_i$$

Our work in **Lemma 3.1.3** tells us that the gauge part can be reduced to:

$$\Omega_A = i\Gamma^a\psi^i\otimes\rho(h^{-1})\rho_*(A_s(e_a))v_i$$

hence:

$$\Omega_A + \Omega_S = (\kappa \otimes \rho)(e, h^{-1}) i\Gamma^a \cdot \left(d\psi^i(e_a) \otimes v_i - \frac{1}{4}\xi_{bc}(e_a)\Gamma^{bc}\psi^i \otimes v_i + \psi_i \otimes \rho_*(A_s(e_a))v_i \right)$$

Since:

$$d(\psi^i) \otimes v_i = d(\psi^i \otimes v_i)$$

it follows that:

$$\Omega_A + \Omega_S = (\kappa \otimes \rho)(e, h^{-1}) D_A(\psi^i \otimes v_i)$$

Therefore:

$$\begin{aligned} D_{f^*A}(f^{-1} \cdot \Psi)|_U =& [(\epsilon \times_M s), (\kappa \otimes \rho)(e, h^{-1}) D_A(\psi^i \otimes v_i)] \\ =& f^{-1} \cdot [\epsilon \times_M s, D_A(\psi^i \otimes v_i)] \\ =& f^{-1} \cdot D_A \Psi|_U \end{aligned}$$

implying the claim. □

Denote the sections of $S \otimes E$ with compact support by $\Gamma(S \otimes E)^c$. To define the Yang-Mills-Higgs action, we need the following construction:

Definition 3.1.16. The L^2 inner product on $(S \otimes E)^c$ with compact support denoted by $\langle \cdot, \cdot \rangle_{S\otimes E, L^2}$, is given on all $\Phi, \Psi \in \Gamma(S \otimes E)^c$ by:

$$\langle \Psi, \Phi \rangle_{L^2, S\otimes E} = \int_M \langle \Psi, \Phi \rangle_{S \otimes E} \mathrm{dvol}_g$$

Note that this is essentially the same construction as the L^2 inner product on differential k forms, and twisted k forms. In particular, a similar argument to **Proposition 3.1.3** demonstrates that this inner product is nondegenerate.

Given that $d_A^\star$ is the formal adjoint of d_A, it should not surprise the reader that we obtain a similar, if not more convenient result for the twisted Dirac operator D_A:

Theorem 3.1.12. *Let M be a manifold without boundary. If the Dirac form satisfies $\delta = -1$, then the twisted Dirac operator is formally self adjoint with respect to the L^2 inner product on $\Gamma(S \otimes E)^c$. In other words, for all $\Phi, \Psi \in \Gamma(S \otimes E)^c$ we have that:*

$$\int_M \langle D_A \Psi, \Phi \rangle_{S\otimes E} \, dvol_g = \int_M \langle \Psi, D_A \Phi \rangle_{S \otimes E} \, dvol_g$$

We will need the following sequence of lemmas to prove this claim:

Lemma 3.1.6. *Let ∇^A be a covariant derivative on an associated vector bundle E, then there exists a canonical covariant derivative on E^* defined implicitly for all $\Phi^* \in \Gamma(E^*)$ and all $\Psi \in \Gamma(E)$ by:*

$$(\nabla_X \Phi^*)(\Psi) = \mathscr{L}_X(\Phi^*(\Psi)) - \Phi^*(\nabla_X(\Psi))$$

Proof. Note that if ρ is the representation of G on V, then there is an induced representation ρ' on V^* defined implicitly by:

$$(\rho'(g) \cdot \omega)(v) = \omega(\rho(g^{-1}) \cdot v)$$

for all $g \in G$, $\omega \in V^*$, and $v \in V$. We then have that:

$$E^* = P \times_{\rho'} V^* \tag{3.1.34}$$

It follows by differentiating $(\rho(\exp(tX)) \cdot \omega)(v)$ at $t = 0$, that the induced representation $\rho'_* : \mathfrak{g} \to \mathrm{End}(V^*)$ is also defined implicitly by:

$$(\rho'_*(X) \cdot \omega)(v) = -\omega(\rho'_*(X) \cdot v)$$

We also see, for some:

$$\Phi^* = [p, \phi^*] \qquad \text{and} \qquad \Psi = [p, \psi]$$

where $\pi(p) = x$, $\phi^* \in V^*$ and $\psi \in V$, that the pairing:

$$\Phi^*(\Psi) = \phi^*(\psi)$$

is well defined. Indeed, for some $g \in G$, we have that:

$$\Phi^* = [p \cdot g, \rho'(g^{-1}) \cdot \phi^*] \qquad \text{and} \qquad \Psi = [p \cdot, \rho(g) \cdot \psi]$$

so:

$$\Phi^*(\Psi) = \rho'(g^{-1})\phi^*(\rho(g)\psi) = \phi^*(\psi)$$

It suffices to prove the claim in a local gauge, as the covariant derivative is independent of our choice of a local gauge. Let $s : U \to P_U$ be a local gauge, and $\phi^* : U \to V^*$, $\psi : U \to V$ be smooth maps. Then:

$$\begin{aligned}(\nabla^A_X \phi^*)(\psi) =&(d\phi_*(X))(\psi) + (\rho'_*(A_s(X))\phi^*)(\psi) \\ =&(d\phi^*(X))(\psi) - \phi^*(\rho_*(A_s(X))\psi)\end{aligned}$$

We note that:

$$d(\phi^*(\psi))(X) = (d\phi^*(X))(\psi) + \phi^*(d\psi(X))$$

hence:

$$\begin{aligned}(\nabla^A_X \phi^*)(\psi) =&d(\phi^*(\psi))(X) - \phi^*(d\psi(X)) - \phi^*(\rho_*(A_s(X))\psi) \\ =&d(\phi^*(\psi))(X) - \phi^*(\nabla^A_X \psi)\end{aligned}$$

It follows that:

$$\begin{aligned}(\nabla^A_X [s, \phi^*]) =&d(\phi^*(\psi))(X) - \phi^*(\nabla^A_X \Psi) \\ =&\mathscr{L}_X(\Phi^*(\Psi)) - \Phi^*(\nabla^A_X \Psi))\end{aligned}$$

as desired. $\square$

Lemma 3.1.7. *Let ∇ be the Levi-Civita connection on (M, g), and $\omega \in \Omega^1(M)$, then if $\nabla\omega = 0$, ω is closed, i.e $d\omega = 0$.*

Proof. Note that by the preceding lemma, for all $X, Y \in \mathfrak{X}(M)$:

$$(\nabla_X \omega)(Y) =\mathscr{L}_X(\omega(Y)) - \omega(\nabla_X Y)$$

Furthermore, we have that:

$$d\omega(X, Y) = \mathscr{L}_X(\omega(Y)) - \mathscr{L}_Y(\omega(X)) - \omega\left(\mathscr{L}_X Y\right)$$

Since the Levi-Civita connection is torsion free:

$$\omega(\nabla_X Y) = \omega\left(\nabla_Y X\right) + \omega\left(\mathscr{L}_X Y\right)$$

However, since $\nabla\omega = 0$ we have that:

$$\omega(\nabla_Y X) = \mathscr{L}_Y(\omega(X)) - \left(\nabla_Y \omega\right)(X) = \mathscr{L}_Y(\omega(X))$$

hence:

$$\begin{aligned}(\nabla_X \omega)(Y) =&\mathscr{L}_X(\omega(Y)) - (\mathscr{L}_Y(\omega(X)) + \omega\left(\mathscr{L}_X Y\right)) \\ =&\mathscr{L}_X(\omega(Y)) - \mathscr{L}_Y(\omega(X)) - \omega\left(\mathscr{L}_X Y\right) \\ =&d\omega(X, Y)\end{aligned}$$

but $\nabla\omega = 0$ thus for all $X, Y \in \mathfrak{X}(M)$:

$$d\omega(X, Y) = 0$$

therefore ω is closed. $\square$

Lemma 3.1.8. *Let $\omega \in \Omega^1(M)$, $x \in M$ a point, and $e_1, \dots, e_n$ a local oriented and time oriented orthonormal frame for an open neighborhood U of x such that $(\nabla e_i)(x) = 0$ for all i. Let $\omega_i = \omega(e_i)$, and define $\omega^i = \eta^{ii}\omega_i$, where there is no summation as, in this frame $g = \eta$, i.e. standard metric on $\mathbb{R}^{t,s}$. Then at the point x:*

$$(\star d \star \omega)(x) = (-1)^t \sum_i^n L_{e_i}\omega^i(x)$$

Proof. Let e^i be the frame g dual to e_i, and is thus an orthonormal coframe. It follows that since in this frame $g = \eta$:

$$g_{ij} = \eta_{ij} = \eta^{ij}$$

Furthermore,

$$\omega = \tilde{\omega}_i e^i$$

We can find $\tilde{\omega}_i$ by:

$$\begin{aligned}\omega_j = \omega(e_j) =& \tilde{\omega}_i e^i(e_j)\\ =& \tilde{\omega}_i \eta_{ij}\\ =& \tilde{\omega}_j \eta_{jj}\end{aligned}$$

implying that in this coframe:

$$\omega = \sum_{i=1}^n \omega_i \eta^{ii} e^i$$

In a coordinate neighborhood around $x \in M$:

$$d\omega_i = \partial_j(\omega_i)dx^j$$

which for some matrix of functions such that:

$$dx^j = X^j_k e^k$$

can be written as:

$$d\omega_i = \partial_j(\omega_i)X^j_k e^k \tag{3.1.35}$$

It follows from **Lemma 3.1.1** that:

$$\begin{aligned}\star\omega =& \sum_{i=1}^n \eta^{ii}\eta^{ii}\omega_i \epsilon_{i1\cdots\hat{\imath}\cdots n} e^1 \wedge \cdots \wedge \hat{e}^i \wedge \cdots e^n\\ =& \sum_{i=1}^n \omega_i \epsilon_{i1\cdots\hat{\imath}\cdots n} e^1 \wedge \cdots \wedge \hat{e}^i \wedge \cdots e^n\end{aligned}$$

We now note that, by the Leibniz property of ∇, for all $Y \in \mathfrak{X}(M)$:

$$\begin{aligned}(\nabla_X e^i)(Y) =& \mathscr{L}_X(e^i(Y)) - e^i(\nabla_X Y)\\ =& \mathscr{L}_X(Y^j e^i(e_j)) - e^i(\nabla_X Y^j e_j)\\ =& Y^j e^i(\nabla_X e_j)\end{aligned}$$

hence at $x \in M$:

$$(\nabla_X e^i)(Y) = 0$$

So by **Lemma 3.1.7**, $(de^i)(x) = 0$. It follows that at x:

$$d(\star\omega) = \sum_{i=1}^n \frac{\partial \omega_i}{\partial x^j} X^j_k \epsilon_{i1\cdots\hat{\imath}\cdots n} e^k \wedge e^1 \wedge \cdots \wedge \hat{e}^i \wedge \cdots e^n$$

unless $k = i$, the term is zero hence:

$$d(\star\omega) = \sum_{i=1}^{n} \frac{\partial \omega_i}{\partial x^j} X_i^j \epsilon_{i1\cdots\hat{i}\cdots n} e^i \wedge e^1 \wedge \cdots \wedge \hat{e}^i \wedge \cdots e^n$$

where there is no second implied summation over i. Note that:

$$e^i \wedge e^1 \wedge \cdots \wedge \hat{e}^i \wedge \cdots e^n = (-1)^{i-1} \mathrm{dvol}_g$$

while:

$$\epsilon_{i1\cdots\hat{i}\cdots n} = (-1)^{i-1}$$

hence:

$$d(\star\omega) = \sum_{i=1}^{n} \frac{\partial \omega_i}{\partial x^j} X_i^j \mathrm{dvol}_g$$

It then again follows from **Lemma 3.1.1** that at x:

$$(\star d \star \omega) = (-1)^t \sum_{i=1}^{n} \frac{\partial \omega_i}{\partial x^j} X_i^j$$

We thus need to show:

$$\left(\sum_{i=1}^{n} \frac{\partial \omega_i}{\partial x^j} X_i^j \right)(x) = \left(\sum_{i}^{n} \mathscr{L}_{e_i} \omega^i \right)(x)$$

By (3.1.34) we have that since each ω_i just a function:

$$d\omega_i(e_i) = \mathscr{L}_{e_i} \omega_i = \eta_{ii} \frac{\partial \omega_i}{\partial x^j} X_k^j e^k$$

multiplying both sided by η^{ii} implies that:

$$\frac{\partial \omega_i}{\partial x^j} X_i^j = \eta^{ii} \mathscr{L}_{e_i}(\omega_i)$$

and thus the claim. □

Lemma 3.1.9. *Let $\langle \cdot, \cdot \rangle$ be a Dirac form such that $\delta = -1$. Then for fixed $\Phi, \Psi \in \Gamma(S \otimes E)$, there exists a one form $\omega \in \Omega^1(M, \mathbb{C})$ such that:*

$$\langle D_A \Phi, \Psi \rangle_{S\otimes E} - \langle \Phi, D_A \Psi \rangle_{S\otimes E} = (-1)^t (\star d \star \omega)$$

Proof. We define ω by:

$$\omega(X) = \langle X \cdot \Phi, \Psi \rangle_{S\otimes E}$$

It is clear that this defines a smooth, $C^\infty(M)$ linear map $\mathfrak{X}(M) \to C^\infty(M, \mathbb{C})$, and thus ω is indeed a one form.

We will prove the result by showing that for all $x \in M$ the equality holds. It is a standard fact from Riemannian geometry that for all $x \in M$ there exists a local oriented orthonormal frame for a neighborhood of x such that $(\nabla e_i)(x) = 0$[39]. Let e be such a frame for some $x \in M$, and for some smooth section $\epsilon \times_M s$, such that $\Lambda \circ \epsilon = e$ set:

$$\Phi = [\epsilon \times_M s, \phi^i \otimes v_i] \qquad \text{and} \qquad \Psi = [\epsilon \times_M s, \psi^i \otimes v_i]$$

for some smooth maps $\phi^i, \psi^i : U \to \Delta_n$, and an orthornormal basis v_i for V. We then have that:

$$\omega_i = \sum_j \langle \gamma_i \cdot \phi^j, \psi^j \rangle$$

[39]This result can be found in any text on Riemannian geometry.

hence by **Definition 2.2.20**:

$$\omega^i = \sum_j \langle \gamma^i \cdot \phi^j, \psi^j \rangle = -\sum_j \langle \phi^j, \gamma^i \psi^j \rangle$$

Note that:

$$\begin{aligned}\partial_k \omega^i &= \sum_j \partial_k \langle \gamma^i \cdot \phi^j, \psi^j \rangle \\ &= \sum_j \left(\langle \gamma^i \cdot \partial_k \phi^j, \psi^j \rangle - \langle \phi^j, \gamma^i \cdot \partial_k \psi^j \rangle \right)\end{aligned}$$

From **Lemma 3.1.8** we have that:

$$\begin{aligned}(-1)^t (\star d \star \omega) &= \sum_{i=1}^n \frac{\partial \omega^i}{\partial x^k} X_i^k \\ &= \sum_{i=1}^n \sum_j \left(\langle \gamma^i \cdot \partial_k \phi^j, \psi^j \rangle - \langle \phi^j, \gamma^i \cdot \partial_k \psi^j \rangle \right) X_i^k \\ &= \sum_j \left(\langle \gamma^i \cdot d\phi^j(e_i), \psi^j \rangle - \langle \phi^j, \gamma^i \cdot d\psi^j(e_i) \rangle \right)\end{aligned}$$

where there is now an implied summation over i. We see that at $x \in M$:

$$\langle D_A \Phi, \Psi \rangle_{S \otimes E} = \sum_j \left\langle \gamma^i \cdot d\phi^j(e_i), \psi^j \right\rangle + \langle \gamma^i \cdot \phi^j, \psi^j \rangle \left\langle \rho_*(A_a(e_i)) v_j, v_j \right\rangle_V$$

while:

$$\langle \Phi, D_A \Psi \rangle_{S \otimes E} = \sum_j \left\langle \phi^j, \gamma^i \cdot d\psi^j(e_i) \right\rangle - \langle \gamma^i \cdot \phi^j, \psi^j \rangle \left\langle v_j, \rho_*(A_a(e_i)) v_j \right\rangle_V$$

Since $\langle \cdot, \cdot \rangle_V$ is G invariant, we have that at x:

$$\begin{aligned}\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} = &\sum_j \left\langle \gamma^i \cdot d\phi^j(e_i), \psi^j \right\rangle - \left\langle \phi^j, \gamma^i \cdot d\psi^j(e_i) \right\rangle \\ = &(-1)^t (\star d \star \omega)\end{aligned}$$

Since this holds for all $x \in M$ we thus have the claim. □

We can now prove **Theorem 3.1.12**:

Proof. We have that by **Lemma 3.1.9**:

$$\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} = (-1)^t (\star d \star \omega)$$

Note that since $\star(1) = \mathrm{dvol}_g$, we have that:

$$\star \left(\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} \right) = \left(\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} \right) \mathrm{dvol}_g$$

So:

$$\left(\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} \right) \mathrm{dvol}_g = (-1)^t \star (\star d \star \omega)$$

Note that $d \star \omega$ is an n form, so by **Proposition 3.1.2**:

$$\star\star = (-1)^t$$

hence:

$$\left(\langle D_A \Phi, \Psi \rangle_{S \otimes E} - \langle \Phi, D_A \Psi \rangle_{S \otimes E} \right) \mathrm{dvol}_g = d(\star \omega)$$

It thus follows by Stokes' theorem that:

$$\begin{aligned}\int_M \langle D_A\Psi, \Phi\rangle_{S\otimes E}\mathrm{dvol}_g - \int_M \langle \Psi, D_A\Phi\rangle_{S\otimes E}\mathrm{dvol}_g =& \int_M \langle D_A\Psi, \Phi\rangle_{S\otimes E} - \langle \Psi, D_A\Phi\rangle_{S\otimes E}\mathrm{dvol}_g\\ =& \int_M (d\star\omega)\\ =& 0\end{aligned}$$

implying the claim. ☐

We can now proceed in a similar manner to the Yang-Mills-Higgs Lagrangian. We define the following action:

Definition 3.1.17. The **Yang-Mills-Dirac Action** is given by:

$$S_{YMD}[\Psi, A] = \int \mathscr{L}_{YMD}\mathrm{dvol}_g$$

The action can be viewed as smooth map from $\Gamma(\psi) \times \mathscr{A}(P) \to \mathbb{R}$. It should be clear that the action is real valued as the Lagrangian is real valued.

We want to obtain the field equation of the twisted spinor field Ψ. To do this, we vary action with respect to Ψ, just as we did with a multiplet scalar field in **Theorem 3.1.7**.

Theorem 3.1.13. *In addition to the the aforementioned fixed data, let (M, g) be a closed manifold, and $\delta = -1$ for the Dirac bundle metric $\langle\cdot,\cdot\rangle_S$. The variation of S_{YMD} with respect to a twisted spinor field Ψ yields the following field equation:*

$$D_A\Psi = m\Psi$$

*This is known as the **Dirac Equation**.*

Proof. Let Φ be any other section of $S \otimes E$, then:

$$\mathscr{L}_{YMH}[\Psi + t\Phi, A] = \mathscr{L}_{YM} - m\langle\Psi + t\Phi, \Psi + t\Phi\rangle_{S\otimes E} + \mathrm{Re}(\langle\Psi + t\Phi, D_A\Psi + tD_A\Phi\rangle_{S\otimes E})$$

We first see that:

$$\begin{aligned}\langle\Psi + t\Phi, \Psi + t\Phi\rangle_{S\otimes E} =& \langle\Psi, \Psi\rangle_{S\otimes E} + t\langle\Psi\Phi\rangle_{S\otimes E} + t\langle\Phi\Psi\rangle_{S\otimes E} + \mathcal{O}(t^2)\\ =& \langle\Psi, \Psi\rangle_{S\otimes E} + 2t\,\mathrm{Re}(\langle\Psi, \Phi\rangle_{S\otimes E})\mathcal{O}(t^2)\end{aligned}$$

While:

$$\langle\Psi + t\Phi, D_A\Psi + tD_A\Phi\rangle_{S\otimes E} = \langle\Psi, D_A\Psi\rangle_{S\otimes E} + t\langle D_A\Psi, \Phi\rangle_{S\otimes E} + t\langle\Psi, D_A\Phi\rangle_{S\otimes E} + \mathcal{O}(t^2)$$

It follows that by **Theorem 3.1.12**:

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0} S_{YMD}[\Psi + t\Phi, A] =& \int_M \mathrm{Re}(\langle D_A\Psi, \Phi\rangle_{S\otimes E} + \langle\Psi, D_A\Phi\rangle_{S\otimes E} - 2m\langle\Psi, \Phi\rangle_{S\otimes E})\mathrm{dvol}_g\\ =& \mathrm{Re}\left(\int_M \langle D_A\Psi, \Phi\rangle_{S\otimes E} + \langle\Psi, D_A\Phi\rangle_{S\otimes E} - 2m\langle\Psi, \Phi\rangle_{S\otimes E}\mathrm{dvol}_g\right)\\ =& 2\,\mathrm{Re}\left(\int_M \langle D_A\Psi, \Phi\rangle_{S\otimes E} - m\langle\Psi, \Phi\rangle_{S\otimes E}\mathrm{dvol}_g\right)\\ =& 2\,\mathrm{Re}\left(\int_M \langle D_A\Psi - m\Psi, \Phi\rangle\mathrm{dvol}_g\right)\end{aligned}$$

Since the L^2 inner product is nondegenerate, we have if Ψ leaves S_{YMD} stationary, then:

$$D_A\Psi = m\Psi$$

as desired. ☐

To vary the action with respect to A we need analogues of **Lemma 3.1.4** and **Lemma 3.1.5**.

Lemma 3.1.10. *Let $\alpha_M \in \Omega^1(M, Ad(P))$ and $\Psi \in \Gamma(S \otimes E)$, then there exists a canonical section $\alpha_M \cdot \Phi \in \Gamma(S \otimes E)$.*

Proof. This proof is similar to **Lemma 3.1.4**, but instead also incorporates the Clifford multiplication of forms. Note that any $\psi \in \Delta_n \otimes E$ can be written as the linear combination:

$$\psi = \psi^i \otimes v_i$$

for some $\psi^i \in \Delta_n$, and $v_i \in V$. It follows that for all $x \in M$:

$$\Psi(x) = [(p, q), \psi^i \otimes v_i]$$

where $(p, q) \in (\mathrm{Spin}^+(M) \times_M P)_x$. Note that $\alpha_M \in \Gamma(T^*M \otimes \mathrm{Ad}(P))$, and that $T^*M \otimes \mathrm{Ad}(P)$ is the associated vector bundle:

$$T^*M \otimes \mathrm{Ad}(P) = (SO^+(M) \times_M P) \times_{\rho'_{SO^+} \otimes \mathrm{Ad}} (\mathbb{R}^{t,s*} \otimes \mathfrak{g})$$

where ρ'_{SO^+} is the representation of $SO^+(t, s)$ on $\mathbb{R}^{t,s*}$ induced by the standard representation of $SO^+(t, s)$ on $\mathbb{R}^{t,s}$. It follows that:

$$\alpha_M(x) = [(\Lambda(p), q), \omega^i \otimes X_i]$$

for some $\omega^i \in \mathbb{R}^{t,s*}$ and $X_i \in \mathfrak{g}$. Recall that there exists a bundle isomorphism $F : T^*M \to TM$, given by:

$$\omega \longmapsto \omega \lrcorner g^{-1}$$

As associated vector bundles, it follows that this map is given by:

$$F([\Lambda(p), \omega]) = [\Lambda(p), \omega \lrcorner \eta^{-1}]$$

where η is the pseudo Euclidean inner product on $\mathbb{R}^{t,s}$. Denote $\omega \lrcorner \eta^{-1}$ by v_ω, we thus obtain an induced bundle isomorphism:

$$\begin{aligned} F : T^*M \otimes \mathrm{Ad}(P) &\longrightarrow TM \otimes \mathrm{Ad}(P) \\ [(\Lambda(p), q), \omega^i \otimes X_i] &\longmapsto [(\Lambda(p)p, q), v_{\omega^i} \otimes X_i] \end{aligned}$$

where there is still an implied summation over i. We define $\alpha_M \cdot \Psi$ point wise by:

$$\begin{aligned} (\alpha_M \cdot \Psi)_x =& (F(\alpha_M) \cdot \Psi)_x \\ =& [(p, q), v_{\omega^i} \cdot \psi^j \otimes \rho_*(X_i)v_j] \in (S \otimes E)_x \end{aligned}$$

Since α_M is Ad invariant:

$$F(\alpha_M)(x) = [(\Lambda(p \cdot s), q \cdot g), \rho_{SO^+}(\lambda(s^{-1}))v_{\omega^i} \otimes \mathrm{Ad}_{g^{-1}}(X_i)]$$

We see that this construction is well defined as by **Proposition 2.2.11**:

$$\begin{aligned} &[(p \cdot s, q \cdot g), (\rho_{SO^+}(\lambda(s))^{-1}v_{\omega^i}) \cdot (\kappa(s^{-1})\psi^j) \otimes \rho_*(\mathrm{Ad}_{g^{-1}}X_i)\rho(g)^{-1}v_j] \\ &\qquad =[(p, q) \cdot (s, g), \kappa(s^{-1}) \cdot (v_{\omega^i} \cdot \psi^j) \otimes \rho(g^{-1}) \cdot \rho_*(X_i)v_j] \\ &\qquad =[(p, q) \cdot (s, g), (\kappa \otimes \rho)(s, g)^{-1}(v_{\omega^i} \cdot \psi^j \otimes \rho_*(X_i)v_j)] \\ &\qquad =[(p, q), v_{\omega^i} \cdot \psi^j \otimes \rho_*(X_i)v_j] \end{aligned}$$

It then follows that $\alpha_M \cdot \Psi$ by examining smooth sections of $\mathrm{Spin}^+(M) \times_M P$, implying the claim. ☐

Lemma 3.1.11. *There exists a unique twisted one form $J_D(\Psi) \in \Omega^1(M, Ad(P))$ such that:*

$$\langle \alpha_M, J_D(\Psi) \rangle_{Ad(P)} = \mathrm{Re}(\langle \Psi, \alpha_M \cdot \Psi \rangle)_{S \otimes E}$$

Proof. This follows from the same exact argument as **Lemma 3.1.5**. ☐

We will not find a general local expression for $J_D(\Psi)$, as we did for $J_H(A,\Psi)$, but instead will wait until we study this Lagrangian over $\mathbb{R}^{1,3}$ with $G=U(1)$. We can now vary the action with respect to A:

Theorem 3.1.14. *In additions to the aforementioned fixed data, let (M,g) be closed. Then the variation of S_{YMD} with respect to the connection form A yields the following field equation:*

$$d_A^\star F_M^A = J_D(\Psi) \tag{3.1.36}$$

called the ***Yang-Mills-Dirac Equation****.*

Proof. Let $\alpha\in\Omega^1_{\text{hor}}(P,\mathfrak{g})^{\text{Ad}}$, then our work in **Theorem 3.1.4** demonstrates that:

$$\begin{aligned}\mathscr{L}_{YMH}[\Psi,A+t\alpha]=&-\frac{1}{2}\left\langle F_M^A,F_M^A\right\rangle_{\text{Ad}(P)}-t\left\langle d_A\alpha_M,F_M^A\right\rangle_{\text{Ad}(P)}\\&-m\langle\Psi,\Psi\rangle_{S\otimes E}+\text{Re}(\langle\Psi,D_{A+t\alpha}\Psi\rangle_{S\otimes E})+\mathcal{O}(t^2)\end{aligned}$$

In any local gauge we have that:

$$D_{A+t\alpha}\Psi=[\epsilon\times_M s, D_{A+t\alpha}\psi]$$

where:

$$\begin{aligned}D_{A+t\alpha}\psi=&\gamma^a\left(d\psi(e_a)+\frac{1}{4}\xi_{bc}(e_a)\gamma^{bc}\psi+\rho_*(A_s(e_a)+t\alpha_s(e_a))\psi\right)\\=&D_A\psi+t\gamma^a\rho_*(\alpha_s(e_a))\psi\end{aligned}$$

Note that if $\{e_i\}$ is the standard basis for $\mathbb{R}^{t,s}$, and $\{e^i\}$ is the basis η dual to it then we can write:

$$\alpha_s=e^i\otimes X_i$$

where:

$$e^i=[\Lambda(\epsilon),e^i]$$

so:

$$\begin{aligned}\gamma^a\alpha_s(e_a)=&\gamma^a e^i(e_a)\otimes X_i\\=&\sum_{i=1}^n\gamma^a\eta_{ai}\otimes X_i\\=&\sum_{i=1}^n\gamma_i X_i\end{aligned}$$

It then follows that by **Lemma 3.1.10**:

$$\begin{aligned}[\epsilon\times_M s,\gamma^a\rho_*(\alpha_s(e_a))\psi]=&\sum_{i=1}^n[\epsilon\times_M s,\gamma_i\cdot\psi^j\otimes\rho_*(X_i)v_j]\\=&(\alpha_M\cdot\Psi)|_U\end{aligned}$$

as $e^i\lrcorner\eta^{-1}=e_i$, and multiplication by e_i is given by mathematical clifford multiplication. We thus have that:

$$D_{A+t\alpha}\Psi=D_A\Psi+t(\alpha_M\cdot\Psi)$$

so:

$$\langle\Psi,D_{A+t\alpha}\Psi\rangle_{S\otimes E}=\langle\Psi,D_A\Psi\rangle+t\langle\Psi,\alpha_M\cdot\Psi\rangle_\Psi$$

We now calculate, and use **Lemma 3.1.11** and **Theorem 3.1.2** to obtain:

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0}S_{YMH}[\Psi,A+t\alpha]=&\int_M-\langle d_A\alpha_M,F_M^A\rangle_{\text{Ad}(P)}+\text{Re}(\langle\Psi,\alpha_M\cdot\rangle_{S\otimes E})\text{dvol}_g\\=&\int_M-\langle\alpha_M,d_A^\star F_M^A\rangle_{\text{Ad}(P)}+\langle\alpha_M,J_D(\Psi)\rangle_{\text{Ad}(P)}\text{dvol}_g\\=&\int_M\langle\alpha_M,J_D(\Psi)-d_A^\star F_M^A\rangle_{\text{Ad}(P)}\text{dvol}_g\end{aligned}$$

Since the L^2 inner product is nondegenerate, we thus have that in order for A to be a stationary point of S_{YMH}:

$$d_A^\star F_M^A = J_D(\Psi)$$

as desired. $\square$

Now that we have derived the Dirac equation, and the Yang-Mills-Dirac equation in full generality, we are ready to restrict ourselves to the case where $P = \mathbb{R}^{1,3} \times U(1)$, and $S \otimes E = \mathbb{R}^{t,s} \times (\Delta_4 \otimes \mathbb{C})$. This restriction yields the classical Lagrangian and Field equations for QED; unfortunately, we are not particularly well equipped to deal with this in great detail, as there is not much classically to be done here. For those interested in the quantum aspects of this theory, we recommend *Peskin and Schroeder's An Introduction to Quantum Field Theory*, and Ticciati's *Quantum Field Theory for Mathematicians*.

Note that we have chosen $V = \mathbb{C}$ as $U(1)$ admits a representation on $\mathbb{C}$ which lines up with our theory electromagnetism for scalar fields. In other words, the charged representation is one dimensional, which lines up with electric charge in electromagnetism. Furthermore, we have have $\Delta_4 \otimes \mathbb{C} \cong \Delta_4$, as Δ_4 is the complex vector space $\mathbb{C}^4$ by **Theorem 2.2.4**.

We use the spinor representation of $\mathrm{Cl}(1,3)$ on $\mathbb{C}^4$ given by **Example 2.2.5**, and a Dirac form A given by:

$$A = i\Gamma_1\Gamma_2\Gamma_3$$

We also have a decomposition into Weyl spinors:

$$\psi = \psi_R + \psi_L$$

given by the eigenspaces of the Chirality operator:

$$\Gamma_5 = \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix}$$

Finally, as in the previous section, we write in some global gauge:

$$iA_s = i(-Vdt + M_i dx^i) \qquad \text{and} \qquad iF_s = i(\partial_\mu A_\nu - \partial_\nu A_\mu) dx^\mu \wedge dx^\nu$$

where $(t, x^i) = (t, x, y, z)$ are the global standard coordinates on $\mathbb{R}^{1,3}$. Note that since all bundles are trivial, every $\Psi\Gamma(\mathbb{R}^{1,3} \times \Delta_4)$ can be written as:

$$\Psi(x) = (x, \psi(x))$$

for some smooth $\psi : \mathbb{R}^{1,3} \to \Delta_4$. This justifies examining solely maps like ψ, in the following discussion.

Proposition 3.1.6. *In the global coordinates (t, x, y, z) the Dirac operator is given by:*

$$\begin{aligned} D_A\psi =& \gamma^\mu \left(\partial_\mu \psi + iA_\mu \psi\right) \\ =& i\Gamma^\mu \left(\partial_\mu + iA_\mu \psi\right) \end{aligned}$$

Proof. This follows from the fact that the one forms ξ_{ab} vanish per **Example 2.2.8**. $\square$

We wish to calculate the Dirac equation in this simplified set up.

Proposition 3.1.7. *The Dirac equation:*

$$D_A\psi = m\psi$$

for $\psi : \mathbb{R}^{1,3} \to \Delta_4$ is equivalent to:

$$\left[i\Gamma^\mu \left(\partial_\mu + iA_\mu\right) - m\right]\psi = 0$$

Proof. This follows from **Proposition 3.1.6**. $\square$

We now calculate the current one form:

Proposition 3.1.8. *The current one form $iJ_D(\psi)$ is given by:*

$$J_D(\psi)^\mu = -\bar{\psi}\Gamma^\mu\psi$$

Proof. Let $i\alpha_M \in \Omega(\mathbb{R}^{1,3}, i\mathbb{R})$, then we see that:

$$\alpha_M = i\alpha_\mu dx^\mu$$

It follows that:

$$\begin{aligned}\alpha_M \lrcorner \eta^{-1} =& i\alpha_i\eta^{\mu\nu}\partial_\mu(dx^i)\otimes\partial_\nu \\ =& i\alpha_i\eta^{\mu i}\partial_\mu\end{aligned}$$

Therefore:

$$\begin{aligned}\alpha_M\cdot\psi =& i\eta^{\mu i}\gamma_\mu\cdot\alpha_i\psi \\ =& i\gamma^i\alpha_i\psi\end{aligned}$$

The inner product satisfies:

$$\langle\psi, i\alpha_i\gamma^i\psi\rangle = i\alpha_i\bar{\psi}\gamma^i\psi$$

We see that this real as by **Definition 2.2.20**:

$$\begin{aligned}\overline{\langle\psi, i\alpha_i\Gamma^i\psi\rangle} =& \langle i\alpha_i\gamma^i\cdot\psi,\psi\rangle \\ =& -\langle i\alpha_i\psi, \gamma^i\cdot\psi\rangle \\ =& \langle\psi, i\alpha_i\gamma^i\cdot\psi\rangle\end{aligned}$$

So since $i\alpha_i$ is purely imaginary, it follows that $\bar{\psi}\gamma^i\psi$ is also purely imaginary. We note that:

$$i\alpha_i\bar{\psi}\gamma^i\psi = -\alpha_i\bar{\psi}\Gamma^i\psi$$

where $-\bar{\psi}\Gamma^i\psi$ must be real. We define:

$$iJ_D(\psi)^\mu = -i\bar{\psi}\Gamma^\mu\psi$$

and find that:

$$\begin{aligned}\langle i\alpha_M, iJ_D(\psi)\rangle =& -\alpha_\mu\bar{\psi}\Gamma^\mu\psi \\ =& i\alpha_\mu\bar{\psi}\gamma^\mu\psi\end{aligned}$$

implying the claim. □

Note that this implies that:

$$\begin{aligned}J_D(\psi)_\mu =& -\eta_{\mu\nu}\bar{\psi}\Gamma^\nu\psi \\ =& -\eta_{\mu\nu}\eta^{\nu\rho}\bar{\psi}\Gamma_\rho\psi \\ =& -\delta^\rho_\mu\bar{\psi}\Gamma_\rho\psi \\ =& -\bar{\psi}\Gamma_\mu\psi\end{aligned}$$

The same argument to **Corollary 3.1.1** tells us that $iJ_D(\psi)$ satisfies the usual continuity equation:

Corollary 3.1.2. *The current one form $iJ_D(\psi) \in \Omega^1(\mathbb{R}^{1,3}, i\mathbb{R})$ satisfies the continuity equation:*

$$d \star iJ_D(\psi) = 0$$

Finally, our work in **Theorem 3.1.10** gives the following result:

Theorem 3.1.15. *The Bianchi identity, and the Yang-Mills-Dirac equation can be expressed as:*

$$\begin{aligned}\nabla\cdot\mathbf{E} =& J_D(\psi)_0 \qquad & \nabla\times\mathbf{E} =& -\partial_t\mathbf{B} \\ \nabla\cdot\mathbf{B} =& 0 \qquad & \nabla\times\mathbf{B} =& \mathbf{j}+\partial_t\mathbf{E}\end{aligned}$$

where $\mathbf{j} = (J_D(\psi)_x, J_D(\psi)_y, J_D(\psi)_z)$.

References

[1] Mark JD Hamilton. *Mathematical Gauge Theory: With Applications to the Standard Model.* Springer, 2017.

[2] John M. Lee. *Introduction to Smooth Manifolds.* Springer, 2003.

[3] M.L. Michelson; H.B. Lawson. *Spin Geometry.* Princeton University Press, 1989.

[4] M.F. Atiyah.; R. Bott; A. Shapiro. "Clifford Modules". In: (1964).

[5] S. Kobayashi; K. Nomizu. *Foundations of Differential Geometry Vol. I.* Wiles Classics Library, 1969.

[6] Kirill Krasnov. *Formulations of General Relativity: Gravity, Spinors, and Differential Forms.* Cambridge University Press, 2020.

[7] Jürgen Jost. *Riemannian Geometry and Geoemtric Analysis.* Springer, 2011.

[8] Thomas Friedrich. *Dirac Operators in Riemannian Geometry.* American Mathematical Society, 1997.

[9] D.V. Schroeder M.E. Peskin. *An Introduction to Quantum Field Theory.* Avalon Publishing, 1995.

[10] Robin Ticciati. *Quantum Field Theory for Mathematicians.* Cambridge University Press, 1999.