# Curvature on Principal $G$-Bundles:
## Differential Geometric Formalisms of Modern Physics

An Honors Thesis
Department of Mathematics
Northwestern University

Malcolm Spilka Lazarow
Completed on
April 20, 2017
Revised on
October 1, 2017

supervised by
Professor Santiago Cañez

# Contents

# 1 Introduction

The purpose of this thesis, in the most general sense, is to support the notion that many modern abstract frameworks in differential geometry can be used to construct elegant descriptions of the physical universe. I believe that the interplay between pure math and physics is not only beautiful, but also allows one to uncover fundamental truths hidden within physical phenomena.

I began learning differential geometry with the intention of studying Einsteinian gravity, though my path gradually turned to the study of Lie groups and Lie algebras. Thus, this thesis is dedicated to describing these groups, algebras, and related concepts, which will tie in very nicely with Riemannian geometry.

Lie groups were introduced by Sophus Lie, but it was his doctoral student Élie Cartan who really began to expand the theories. The study began about 150 years ago and continues today. Along the way, in the first half of the 20th century, physicists began describing quantum physics from the perspective of symmetries, which gave a natural introduction of Lie groups to physics; one such notion of this type of Lie-symmetry is called a *gauge theory*.

We will begin by describing Maxwell's equations in the framework of differential forms. This introduction is seemingly irrelevant to and is not a prerequisite for the remainder of the thesis; however, this material will pop up again when we discuss Yang-Mills in section 6. This thesis can be thought of as three chapters, ignoring the introduction: sections 2-3 describing Lie groups and algebras, section 4 describing Riemannian geometry, and sections 5-6 describing the geometry of principal $G$-bundles and Yang-Mills, thereby giving an introduction to gauge theories. These three, seemingly separate, sections will coalesce with Maxwell's equations and Yang-Mills in section 6, which is why their introduction is an appropriate motivation.

Though the ultimate goal of the thesis is to describe how Maxwell's equations arise as a consequence of curvature on principal $U(1)$-bundles and the Yang-Mills equations, along the way I will mention anecdotes to other topics of interest in pure math.

One of the first books I read when I began studying with Professor Cañez was Gravitation by Miser, Thorne, and Wheeler [1]; the book book begins with a short poem, which I presume was written by the authors. I found it quite inspiring when I began my studies, so I wanted to include it here:

We dedicate this book
To our fellow citizens
Who, for love of truth,
Take from their own wants
By taxes and gifts,
And now and then send forth
One of themselves
As dedicated servant,
To forward the search
Into the mysteries and marvelous simplicities
Of this strange and beautiful Universe,
Our home.

Lastly, I would like to thank Professor Cañez for taking the time to guide me through all of this. I truly could not have mastered everything in this thesis without his help. Having taken over 10 courses with Professor Cañez, it is bittersweet to be finishing our studies together. I would also like to thank my family and friends, who have put up with my rants about how incredible differential geometry is and that fact that I'm constantly reading math and physics textbooks and papers on my phone... though I do not think this habit will end.

REVISION NOTES: the purpose of the revision is simply to fix grammatical errors and awkward phrasing. When writing this thesis, I was much more focused on whether or not the proofs were correct than if my English was the best it could be. No content has been added or removed and no proofs have been altered (except to fix grammar and unclear wording).

## 1.1 Notations

1. $F$ stands for "field" and is used interchangeably with $\mathbb{R}$ or $\mathbb{C}$.
2. When $\mathbb{1}$ appears as a map, it is the identity map. $\mathbb{1}$ also denotes the identity matrix; however, arbitrary identity elements of a group are denoted by $e$.
3. There are many propositions that have multiple parts. For example; Proposition 2.2.1 has 3 parts to it. In order to reference the second part, I will write 2.2.1(2).
4. Einstein summation notation is always employed
5. If I say something like "$\{U_i\}$ forms an open cover for a manifold $M$", then $i$ is an index in an indexing set $I = \{1, \dots, n\}$; I tend to also use $j, k, l$, and $m$ as indices when there are multiple indexing variables needed.
6. Maps like $proj_A : A \times B \to A$ simply denotes a projection that takes a tuple in $A \times B$ and maps it to the $A$ component.
7. Definitions are typically explicitly labeled, but the actual term will always be defined is in bold. At times, a term will be introduced or defined outside of a declared definition, in which case it will still be in bold.
8. Typically $\Gamma(E)$ is used to denote smooth sections of a vector bundle; in the Riemannian geometry section, I employ $T^k_l(M) = \times_k \Gamma(TM) \times_l \Gamma(T^*M)$ ($k$ and $l$ direct products, respectively).
9. Qualifiers carry unless stated otherwise; for example, if an entire section has assumed that $G$ is a Lie group, then the propositions do too, unless stated otherwise.

## 1.2 The Hodge Star Operator

Suppose $(M, g)$ is an oriented pseudo-Riemannian $n-$manifold, and that $\omega^1, \dots, \omega^k, \eta^1, \dots, \eta^k$ are 1-forms on $M$. We recall a few basic constructions associated with the Riemannian metric $g$:

1. The metric $g$ can be written in terms of its components, $g_{ij}$, which are the matrix elements $g(X_i, X_j)$ ($\{X_i\}$ is a local frame on $T_pM$ at some point $p \in M$). $g^{ij}$ represents the matrix elements of $g^{-1}$; in other words: $g_{ij}g^{kl} = \delta_{jk}$.

2. The map $\omega \mapsto \omega^{\#}$, given by $\omega \mapsto g^{ij}\omega_j$, is one of the *musical isomorphisms*. This musical isomorphism gives us contravariant tensors from covariant tensors. The other musical isomorphism, $\omega \mapsto \omega^{\flat}$, is the inverse of $\omega \mapsto \omega^{\#}$.

**Proposition 1.2.1.** *The metric $g$ determines an inner product that is frame independent on $\Lambda^k(T_p^*M)$ (the vector bundle of alternating $k$-tensors) by:*
$\langle \omega^1 \wedge \cdots \wedge \omega^k, \eta^1 \wedge \cdots \wedge \eta^k \rangle = det(\langle (\omega^i)^\sharp, (\eta^j)^\sharp \rangle)$

*Proof.* Suppose $\{E_i\}$ is an orthonormal basis on $T_pM$, and let $\{\varepsilon^i\}$ be the orthonormal basis dual to $\{E_i\}$. Any element $\alpha \in \Lambda^k T^*M$ can be written as a linear combination of elements of the form $\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}$. Since the Riemannian metric is a tensor, it is multilinear, and therefore we need only prove the proposition for two arbitrary $\varepsilon^I = \varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}$ and $\varepsilon^J = \varepsilon^{j_1} \wedge \cdots \wedge \varepsilon^{j_l}$ (due to multilinearity, the proposition will then hold for all linear combinations of elements in $\Lambda^k T^*M$).

Following the formula above, we define $\langle \alpha_I \varepsilon^I, \beta_J \varepsilon^J \rangle = \alpha_I \beta_J \delta_{IJ}$ where $\alpha_I, \beta_J$ are arbitrary constants. We must now show that if $\{e^i\}$ is another orthonormal dual basis, we get the same result from the above formula.

Suppose that $\{e^i\}$ is orthonormal, and $\omega = \sum \alpha_I \varepsilon^I = \sum a_I e^I$ and $\eta = \sum \beta_J \varepsilon^J = \sum b_J e^J$. Since $\{e^i\}$ and $\{\varepsilon^i\}$ are orthonormal, we then have an orthogonal transformation, $T$, such that $T(\sum a_I e^I) = \sum \alpha_I \varepsilon^I$ and $T(\sum b_J e^J) = \sum \beta_J \varepsilon^J$. We then have that:
$a_I b_J \delta_{IJ} = \langle a_I e^I, b_J e^J \rangle = TT^T \langle a_I e^I, b_J e^J \rangle = \langle T(a_I e^I), T(b_J e^J) \rangle = \langle \alpha_I \varepsilon^I, \beta_J \varepsilon^J \rangle = \alpha_I \beta_J \delta_{IJ}$, and thus the formula is frame independent. □

We now define the Hodge Star operator:

**Definition.** *The* ***Hodge Star*** *operator, $\star : \Gamma(\Lambda^k T^*M) \to \Gamma(\Lambda^{n-k} T^*M)$, is given by*

$$\omega \wedge \star\omega = \langle \omega, \omega \rangle_g dV_g$$

*(where $dV_g$ is the Riemannian volume form, defined in Appendix C).*

As an example, we will compute the Hodge star for all elements of $\Gamma(\Lambda^2 T^*\mathbb{R}^4)$, where $\mathbb{R}^4$ is equipped with the **Minkowski metric**, defined by:

$$\eta = -dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$$

This is called the $(-+++)$ signature; in the context of relativistic physics, $x_1$ typically represents time and $x_2, x_3, x_4$ the three spatial coordinates. For shorthand, I will write $dx_i \wedge dx_j$ as $i \wedge j$. All these computations boil down to calculating what coordinates have to be added and in what order, such that the original tensor wedged with its Hodge Star is equal to the volume element:

1. $\star(1 \wedge 2) = 3 \wedge 4$. We see this by computing the permutation $(1342) \mapsto (1234)$ by $(1342) \mapsto (1324) \mapsto (1234)$ by $\sigma = (23)(42) \Rightarrow sgn(\sigma) = 1$.
2. $\star(1 \wedge 3) = 4 \wedge 2$, we have that $(1423) \mapsto (1234)$ by $\sigma = (32)(43)$
3. $\star(1 \wedge 4) = 2 \wedge 3$, we have that $(1423) \mapsto (1234)$ by $\sigma = (32)(43)$
4. $\star(2 \wedge 3) = 1 \wedge 4$, we have that $(2341) \mapsto (1234)$ by $\sigma = (32)(12)(41)$

5. $\star(2 \wedge 4) = 3 \wedge 1$, we have $(2413) \mapsto (1234)$ by $\sigma = (31)(14)(12)$
6. $\star(3 \wedge 4) = 1 \wedge 2$, we have $(3412) \mapsto (1234)$ by $\sigma = (24)(13)$

## 1.3 Maxwell's Equations in terms of Differential Forms

Recall Maxwell's Equations in their classical statement (Gaussian units are used):

1. $\nabla \cdot \vec{E} = \rho$
2. $\nabla \cdot \vec{B} = 0$
3. $\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$
4. $\nabla \times \vec{B} = \vec{j} + \frac{\partial \vec{E}}{\partial t}$

where $\vec{E}$ is the electric field (with components $E_x, E_y, E_z$), $\vec{B}$ is the magnetic field (with components $B_x, B_y, B_z$), $\rho$ is the electric density, and $\vec{j}$ is the current density.

If one seeks to rewrite these equations in terms of differential forms, then one defines the electric 1-form as

$$E = E_x dx + E_y dy + E_z dz$$

and the magnetic 2-form as

$$B = B_z dx \wedge dy + B_x dy \wedge dz + B_y dz \wedge dx$$

We define the **Faraday tensor** as

$$F = B + E \wedge dt$$

The general setting for electromagnetism is, in light of special relativiy (pun intended), **Minkowski space**, which is $\mathbb{R}^4$ equipped with the Minkowski metric.

We thus compute:

$$dF = (\partial_t B_z - \partial_y E_x + \partial_x E_y) dt \wedge dx \wedge dy + (\partial_x B_x + \partial_y B_y + \partial_z B_z) dx \wedge dy \wedge dz$$

$$+(\partial_x E_z - \partial_z E_x - \partial_t B_y) dt \wedge dx \wedge dz + (\partial_y E_z - \partial_z E_y + \partial_t B_x) dt \wedge dy \wedge dz$$

We see, by that by requiring $dF = 0$, we get the following:

1. $\partial_x B_x + \partial_y B_y + \partial_z B_z = 0$
2. $\partial_x E_y - \partial_y E_x = -\partial_t B_z$
3. $\partial_z E_x - \partial_x E_z = -\partial_t B_y$
4. $\partial_y E_z - \partial_z E_y = -\partial_t B_x$

which are the second and third of Maxwell's equations! We now consider compute $\star d \star F$:

$$\star F = B_z dt \wedge dz + E_x dz \wedge dy + B_x dt \wedge dx + E_y dx \wedge dz + B_y dt \wedge dy + E_z dy \wedge dx \Rightarrow$$

$$d\star F = (\partial_t E_y + \partial_z B_x - \partial_x B_z) dt \wedge dx \wedge dz + (\partial_y B_x - \partial_x B_y - \partial_t E_z) dt \wedge dx \wedge dy + (\partial_z B_y - \partial_y B_z - \partial_t E_x) dt \wedge dy \wedge dz \Rightarrow$$

$\star d\star F = (\partial_x E_x+\partial_y E_y+\partial_z E_z)dt+(\partial_z B_y-\partial_y B_z-\partial_t E_x)dx+(\partial_x B_z-\partial_z B_x-\partial_t E_y)dy+(\partial_y B_x-\partial_x B_y-\partial_t E_z)dz$

We see that by defining requiring the current 1-form, $J$, as

$$J = \rho dt + j_x dx + j_y dy + j_z dz$$

and requiring that $\star d \star F = J$, we recover the first and fourth of Maxwell's equations.

The real point here is the pair of equations:

1. $dF = 0$
2. $\star d \star F = J$

The climactic moment of this thesis comes when we recover these two equations from a seemingly unrelated abstract setting in the final section. These equations come from a framework in differential geometry called *Yang-Mills Theory*. We will uncover them as a constraint on the curvature of a type of manifold entitled a *principal G-bundle*. In this abstract setting, $F$ corresponds to some differential form quantifying the curvature of $P$, a principal $G$-bundle. The "$G$" in principal $G$-bundle is a Lie group and in the case of electromagnetism, $G = U(1)$. The above two equations, called the *Yang-Mills Equations* constrain the curvature on $P$ and are thus relevant to numerous gauge theories, which we will build up to describe in a rigorous sense.

In fact, $dF = 0$ implies that $F = dA$ (this is true by the Poincare Lemma, which can be found in [2]), for some 1-form $A$. We thus define $A$ as the electromagnetic potential:

$$A = \phi dt + A_x dx + A_y dy + A_z dz$$

where $\phi$ is the *electric potential* and the components $A_x dx + A_y dy + A_z dz$ can be thought of as a 1-form such that $(A_x dx+A_y dy+A_z dz)^\sharp = A_x\partial_x + A_y\partial_y + A_z\partial_z$, the magnetic vector potential. We note that by replacing $A$ with $A+df$ (where $f : \mathbb{R}^4 \to \mathbb{R}$), $F$ remains constant, thereby maintaining a solution to Maxwell's Equations. This is a gauge theory: some alteration to $A$ that leaves $F$, which encodes the "physically observable quantities", unchanged.

The fundamental object of study in particle physics is the standard model, which is a gauge theory. One such gauge theory is quantum electrodynamics; this thesis serves as an introduction to the mathematical constructions of gauge theories and how Maxwell's equations can be seen as a special case of a more abstract mathematical construction. The next time we will see Maxwell's equations is when they are "derived" from the Yang-Mills equations with a $U(1)$ gauge. Of course, physical laws cannot be derived, but indeed we can derive them as mathematical equations (and one can apply what we do here to gauges such as $SU(2)$ for the electroweak force and $SU(3)$ for the strong force, whose study is called quantum chromodynamics).

# 2 Lie Groups

## 2.1 Basic Definitions

**Definition.** *$G$ is a Lie group if the following conditions apply:*

1. *$G$ is a topological group with the structure of a smooth manifold.*

2. *The multiplication map, $m : G \times G \to G$ by $m(g, h) = gh$, is smooth.*

Typically, one also requires that the inversion map $i : G \to G$ (by $i(g) = g^{-1}$) is smooth. This leads us to our first result:

**Proposition 2.1.1.** *If multiplication in $G$ is smooth, then inversion is necessarily smooth.*

*Proof.* Because $m$ is smooth, we can define a smooth function $\varphi : G \times G \to G \times G$ by $\varphi(g, h) = (g, gh)$. We have, by Theorem C.1, that $dm_{(e,e)}(X, Y) = (X, X + Y)$, thus $d\varphi_{(e,e)} = (X, X + Y)$ has full rank, thus it is a local diffeomorphism. Therefore, it has a *smooth* inverse: we seek to define $\varphi^{-1}$ such that $\varphi^{-1} \circ \varphi = \mathbb{1}$. $\varphi^{-1}(g, gh) = (e, e) \Rightarrow \varphi^{-1}(\alpha, \beta) = (\alpha, i(\alpha)\beta)$ is smooth, thus inversion is smooth. $\square$

## 2.2 Generated Lie Subgroups and the Identity Component

Most of the information in this section will be relevant in proofs in sections 3.3 and 3.4.

**Definition.** *Suppose $W \subseteq G$ is a neighborhood in $G$ (not necessarily a subgroup of $G$). We say that $S$ is the* ***subgroup generated by*** *$W$ if it is the smallest subgroup of $G$ that contains $W$; we say "$W$ generates $S$".*

**Proposition 2.2.1.** *Suppose $W \subseteq G$ is any neighborhood of $e$.*

1. *$W$ generates an open subgroup of $G$*
2. *If $W$ is connected, then $W$ generates a connected open subgroup of $G$*
3. *If $G$ is connected, then $W$ generates $G$.*

A proof of the above is given in [2]. We now mention a few remarks about a particular Lie subgroup.

**Definition.** *The component of $G$ containing $e$ is the* ***identity component of*** *$G$, hereafter referred to as $G_e$.*

**Proposition 2.2.2.** *$G_e$ has the following properties:*

1. *$G_e$ is the only connected open subgroup of $G$*
2. *$G_e$ is a normal subgroup of $G$*
3. *If $H$ is a connected component of $G$, then it is diffeomorphic to $G_e$.*

*Proof.*

1. Let $W$ be any neighborhood of $e \in G$. We then have, by Proposition 2.2.1(3), that $W$ generates $G_e$, an open subgroup, since $G_e$ is connected by the definition of a component. Suppose there is another open subgroup of $G$, $S$; moreover, suppose that $S$ is not equal to $G_e$. We then have that $W$ *also* generates $S \Rightarrow S = G_e$ or $G_e \subseteq S$; however, by hypothesis $G_e$ is a component, so it is the largest connected subset of $G$ so $G_e \nsubseteq S$.
2. $G_e$ is connected. Since $L_g$ and $R_g$ are continuous, $L_g G_e$ is connected, and $R_{g^{-1}} L_g G_e$ is connected *and* a neighborhood of $e \in G$. Therefore $R_{g^{-1}} L_g G_e \subseteq G_e \Rightarrow g G_e g^{-1} \subseteq G_e$.

3. Let $H$ be a component of $G$ (note that components are connected by definition), and $h \in H$ be arbitrary; we construct $\varphi : H \to G_e$ by $\varphi = L_{h^{-1}} \Rightarrow L_{h^{-1}}H \subseteq G_e$. Similarly we can construct an inverse $L_h : G_e \to H$ and we have that $L_h G_e \subseteq H$; we thus have that $G_e$ and $H$ are diffeomorphic since left translation is a diffeomorphism.

$\square$

The next ingredient we will add to our Lie vocabulary is the Lie group homomorphism:

## 2.3 Lie Group Homomorphisms

**Definition.** *If $G$ and $H$ are both Lie groups, then $\varphi : G \to H$ is a Lie group homomorphism if it is a group homomorphism that is also smooth; thus it is a Lie group isomorphism if it is a Lie homomorphism and diffeomorphism.*

**Proposition 2.3.1.** *Lie group homomorphisms are of constant rank*

The proposition above gives us the following corollaries, which will allow us to build a list of Lie groups (the proposition and both corollaries are proven in [2]):

**Corollary 2.3.1.1.** *The kernel of a Lie group homomorphsim is a properly embedded Lie subgroup of the domain.*

**Corollary 2.3.1.2.** *The image of a Lie group homomorphsim is a Lie subgroup of the codomain.*

**First Examples of Lie Groups**

1. Trivially, $\mathbb{R}^n$ and $\mathbb{C}^n$ form Lie groups.
2. $GL(n,\mathbb{R})$. We show that $GL(n,\mathbb{R})$ is a smooth manifold and that multiplication is smooth:

   First, recall that an open subset of a smooth manifold is a smooth manifold. Next, we have that $\mathfrak{g} = \mathbb{R}^{n^2}$ (where $\mathfrak{g}$ is the space of all $n$-by-$n$ matrices with real entries) is a smooth manifold; thus we can consider the determinant $det : \mathbb{R}^{n^2} \to \mathbb{R}$, which is a smooth map because it is merely a polynomial with entries in $\mathbb{R}^{n^2}$. We have that $\forall A \in GL(n,\mathbb{R})$, $det(A) \neq 0 \Rightarrow GL(n,\mathbb{R}) = det^{-1}(U)$ where $U = (-\infty, 0) \cup (0, \infty)$. $U$ is an open subset of $\mathbb{R}$ and $det$ is continuous, so $det^{-1}(U)$ is open in $\mathbb{R}^{n^2}$, thus it is a smooth manifold. Multiplication in $GL(n,\mathbb{R})$ is smooth because every matrix multiplication is a dot product of its columns or rows, and thus a polynomial in $\mathbb{R}^{2n}$, so it is a Lie group.
3. $SL(n,F)$ is a properly embedded Lie subgroup of $GL(n,F)$; this follows from the determinant: $det : GL(n,F) \to F^{\times}$. The determinant is a group homomorphism, and is smooth, as it is a polynomial; the kernel of the determinant is $SL(n,F)$.
4. Nearly any matrix group one can think of ($O(n), SO(n), U(n), SU(n)$, etc.) is a Lie group; the definitions of all these matrix Lie groups are (re)stated in the glossary) is a Lie group. Here, we will explicitly show that $U(n)$ and $SU(n)$ are properly embedded Lie subgroups.

   $U(n)$: $U(n)$ consists of matrices such that $\bar{A}^T A = \mathbb{1}$. Consider the map $\varphi : GL(n,\mathbb{C}) \to GL(n,\mathbb{C})$ by $\varphi(A) = \bar{A}^T A$. It is a Lie group homomorphism: $\varphi(AB) = (\bar{AB})^T (AB) = \bar{B}^T \bar{A}^T AB = \bar{B}^T \varphi(A) B = \varphi(B)\varphi(A)$, and it is clearly smooth because multiplication in $GL(n,\mathbb{C})$ is smooth (note $A \in GL(n,\mathbb{C}) \Rightarrow \bar{A}^T \in GL(n,\mathbb{C})$). Thus $ker(\varphi) = \{A \in GL(n,\mathbb{C}) | \bar{A}^T A = \mathbb{1}\} = U(n) \Rightarrow U(n)$ is a properly embedded Lie subgroup of $GL(n,\mathbb{C})$. We have shown that $\forall n \geq 1, U(n)$ is an embedded Lie subgroup. A special case is when $n = 1$; for this case, we have $U(1) = \{\alpha \in \mathbb{C} : |\alpha| = 1\}$. In other words, elements of $U(1)$ are complex numbers with unit modulus: it is simply the complex unit circle $\mathbb{S}$. This is the most straightforward Lie group to visualize that is not "trivial" (such as $\mathbb{R}^n$). Elements in $U(1)$ are of the form $e^{2\pi i\phi}$ with $\phi \in [0, 1)$.

$SU(n)$: $SU(n) = ker(det)$ where $det : U(n) \to \mathbb{R}^*$. $det$ is clearly a Lie group homomorphism because $det(AB) = det(A)det(B)$.

5. $U(1) \cong \mathbb{S}^1$ so $\mathbb{S}^1$ is a Lie group
6. The direct product of groups is a group, and the direct product of smooth manifolds is a smooth manifold; thus, the direct product of Lie groups is a Lie group. In particular the $n$-torus, $\mathbb{T}^n = \mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ is a Lie group.

## 2.4 Lie Group Actions

### 2.4.1 Representations

We start this section by giving very brief exploration into the deep field of representation theory, which we will revisit when we study Lie algebras (where it will be of significantly more important than here). The purpose of representation theory in this thesis is just to build up to the adjoint representation. A representation of a Lie group is a homomorphism $\rho : G \to GL(V)$ where $GL(V)$ is the set of all linear transformations of a vector space $V$.

**Definition.** *A representation is said to be* ***faithful*** *if the homomorphism, $\rho$, is injective.*

**Definition.** *When one considers about a Lie group that is a matrix group, one naturally thinks of the* ***fundamental*** *or* ***defining representation of*** *$G$; this is given by the inclusion map $H \hookrightarrow GL(\mathbb{R}^{n^2})$ where $H \subseteq GL(n,\mathbb{R})$ is a Lie subgroup (this is, trivially, a homomorphism).*

### 2.4.2 Lie Group Actions

Recall the definition of a group action:

**Definition.** *A left action of a group $G$ on a set $M$ is defined to be a map $G \times M \to M$ such that:*

1. $g_1 \cdot (g_2 \cdot p) = (g_1 g_2) \cdot p$
2. $e \cdot p = p$

Hereafter, group actions are written as $g \cdot p$; right actions are defined analogously. We say that an action is continuous or smooth if $\theta : G \times M \to M$, defined by the action, is continuous or smooth. Hereafter, we again assume $G$ is a Lie group and $M$ is a smooth manifold. Before moving onto some propositions, we list some basic definitions that will be used throughout this thesis:

**Definition.** *Given a $p \in M$, we define the* ***orbit of*** *$p$, denoted $\Theta_p$, as the set $G \cdot p = \{g \cdot p : g \in G\}$.*

**Definition.** *Given a $p \in M$, we define the* ***isotropy group of*** $p$*, denoted* $G_p$*, as the subset of* $G$ *that fixes* $p$*; namely,* $G_p = \{g \in G : g \cdot p = p\}$

**Definition.** *We say that an action is* ***transitive*** *if there is one orbit; in other words, for any pair* $p, q \in M$*, there is an element* $g \in G$ *such that* $g \cdot p = q$*. A simple example to keep in mind is* $SO(3)$ *acting on* $\mathbb{S}^2$ *by rotating a point about a specified vector.*

**Definition.** *We say that an action is* ***free*** *if the only element of* $G$ *that fixes an element of* $M$ *is the identity element; in other words, the action is free if* $g \cdot p = p \Leftrightarrow g = e$.

**Definition.** *Let* $F : M \to N$ *be a map; if it is invariant under actions then it is said to be* ***equivariant****; in other words,* $F$ *is equivariant if* $F(g \cdot p) = g \cdot F(p)$ *or* $F(p \cdot g) = F(p) \cdot g$.

**Examples of Lie Group Actions**

1. Consider $[x] \in \mathbb{RP}^n$ (the definition of $\mathbb{RP}^n$ and related sets can be found in Appendix C); we can define an action of $GL_{n+1}(\mathbb{R})$ on $\mathbb{RP}^n$ by $A \cdot [x] \coloneqq [Ax]$. In fact, this defines a smooth, transitive action:

   First, we confirm that it is a group action: suppose $A \in GL_{n+1}(\mathbb{R})$, $[x] \in \mathbb{RP}^n \Rightarrow$ $[\mathbb{1} \cdot x] = [x] = \mathbb{1}[x]$. We also have that $A \cdot B[x] = [A \cdot B \cdot x] = [(AB) \cdot x] = (AB)[x]$, so it is indeed a group action. Note that $GL_{n+1}(\mathbb{R})$ acts smoothly on $\mathbb{R}^{n+1}$ with two orbits: $\{0\}$ and $\mathbb{R}^n/\{0\}$ (this is not a quotient, it is just $\mathbb{R}^n$ without the origin); therefore we have that $GL_{n+1}(\mathbb{R})$ acts transitively on $\mathbb{R}^{n+1}/\{0\}$. We have the natural quotient map $\pi : \mathbb{R}^{n+1}/\{0\} \to \mathbb{RP}^n$ by $\pi(x) = [x]$.

   Thus we can write $A \cdot [x] = [Ax]$ as $A \cdot \pi(x) = \pi(Ax)$, where $x \in \mathbb{R}^{n+1}/\{0\}$. Therefore, $\pi(Ax)$ has one orbit, and thus $A \cdot \pi(x)$ is a smooth, transitive action.

2. Identify $\mathbb{S}^{2n+1} \hookrightarrow \mathbb{C}^{n+1}$ as the complex unit sphere. We define the action of $\mathbb{S}^1$ on $\mathbb{S}^{2n+1}$, entitled the **Hopf action**, by $z \cdot (w_1, \ldots, w_{n+1}) = (zw_1, \ldots, zw_{n+1})$. We will show that this is a smooth action, and its orbits are disjoint unit circles in $\mathbb{C}^{n+1}$ whose union is $\mathbb{S}^{2n+1}$. This will prove to be relevant in sections 2.4.3 and 2.4.4

   Let $z \in \mathbb{S}^1 \Rightarrow z = e^{2\pi i\phi}$, we will denote $\alpha = 2\pi i$ for convenience. Now, let $(w_1, \ldots, w_{n+1}) \in \mathbb{S}^{2n+1} \in \mathbb{C}^{n+1}$; note that we can write every complex number $w_i$ as $c_i e^{\alpha\phi}$ for some $c_i \in \mathbb{R}$ and $\phi \in [0, 1)$, thus we can write $(w_1, \ldots, w_{n+1}) = (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}})$ without loss of generality (note that the equation for the sphere is thus $Re(w_1)^2 + Im(w_1^2) + Re(w_2^2) + \cdots + Re(w_{n+1})^2 + Im(w_{n+1})^2 = |w_1|^2 + \cdots + |w_{n+1}|^2 = c_1^2 + \cdots + c_{n+1}^2 = 1$). We have that the action is given by $e^{\alpha\phi} \cdot (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}}) = (c_1 e^{\alpha(\phi+\phi_1)}, \ldots, c_{n+1} e^{\alpha(\phi+\phi_{n+1})})$ (which is clearly smooth, since addition in $\mathbb{R}$ is smooth); we now check that it is a group action:

   (a) $1 \in \mathbb{S}^1$ is given by $1 = e^{\alpha 0}$, thus $1 \cdot (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}}) = (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}})$.

   (b) Let $x, y \in \mathbb{S}^1$ with $x = e^{\alpha\phi_x}$ and $y = e^{\alpha\phi_y}$. Thus $x \cdot y \cdot (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}}) = (c_1 e^{\alpha(\phi_1+\phi_x+\phi_y)}, \ldots, c_{n+1} e^{\alpha(\phi_{n+1}+\phi_x+\phi_y)}) = (e^{\alpha(\phi_x+\phi_y)}) \cdot (c_1 e^{\alpha\phi_1}, \ldots, c_{n+1} e^{\alpha\phi_{n+1}})$

   Now, consider $\phi_1, \ldots, \hat{\phi}_i, \ldots, \phi_{n+1} = \beta$ (note the hat means that the item is omitted), and $\phi_i = \varphi_i$, thus $(w^1, \ldots, w^{n+1}) = (e^{\alpha\beta}, \ldots, e^{\alpha\varphi_i}, \ldots, e^{\alpha\beta})$. Let $z \in \mathbb{S}^1$ be $z = e^{\alpha\phi} \Rightarrow$

$z \cdot w = (e^{\alpha(\beta+\phi)}, \ldots, e^{\alpha(\varphi_i+\phi)}, \ldots, e^{\alpha(\beta+\phi)})$. Since $\phi \in [0,1)$, this clearly carves out a unit circle.

We will now show that the orbits are disjoint. Let $\tilde{w} = (e^{\alpha\beta}, \ldots, e^{\alpha\varphi_j}, \ldots, e^{\alpha\beta})$ where $i \neq j$ (namely, $\phi_1, \ldots, \hat{\phi}_j, \ldots, \phi_{n+1} = \beta$). Then $z \cdot \tilde{w} = (e^{\alpha(\beta+\phi)}, \ldots, e^{\alpha(\varphi_j+\phi)}, \ldots, e^{\alpha(\beta+\phi)}$. Now suppose $\exists \phi$ such that $z \cdot \tilde{w} = z \cdot w$ (note $\tilde{w} \neq w$ since $i \neq j$). In particular, if $z \cdot \tilde{w} = z \cdot w$ then $(e^{\alpha(\beta+\phi)}, \ldots, e^{\alpha(\varphi_j+\phi)}, \ldots, e^{\alpha(\beta+\phi)}) = (e^{\alpha(\beta+\phi)}, \ldots, e^{\alpha(\varphi_i+\phi)}, \ldots, e^{\alpha(\beta+\phi)})$. Then clearly $\varphi_j = \beta$ or $\varphi_i = \beta$, which is a contradiction, so the orbits are disjoint. Since every point is in an orbit, we clearly have that the union of all orbits is $\mathbb{S}^{2n+1}$.

### 2.4.3 Quotient Manifolds

In topology, one speaks about quotient topological spaces. If $X$ is a topological space, and $\sim$ is an equivalence relation, then we can define a quotient topological space, $X/\sim$ as the space of all equivalence classes. With the quotient topology, $X/\sim$ also forms a topological space. We seek to construct an analogous definition for smooth manifolds: if we have a smooth manifold, $M$, and a Lie group $G$, then when is the orbit space $M/G$ a smooth manifold and what structure can we determine via this type of quotient space? Say that we have a left Lie group action $\theta : G \times M \to M$. Much like we can define a quotient topological space associated with an equivalence relation, which would define $X/\sim$ as the space of equivalence classes, we can define a topological space $M/G$ as the space of orbits given by a Lie group action of $G$ on $M$ (more explicitly, if $p, q \in M$ are in the same orbit, then their projections into $M/G$ are equal); indeed, being in the same orbit is an equivalence relation. The culmination of this section is to specify when this space $M/G$ is necessarily a smooth manifold:

**Definition.** *If a Lie group $G$ acts on a smooth manifold $M$, we call it a **proper action** if the map $\theta \times \mathbb{1} : G \times M \to M \times M$ given by $(g, p) \mapsto (g \cdot p, p)$ is a proper map. Recall that a map $f : A \to B$ is a proper map if for any compact $K \in B$, the preimage of $K$ is compact.*

**Proposition 2.4.1.** *$M$ is a smooth manifold, and $G$ is a Lie group acting on $M$; the following are equivalent:*

1. *The action is proper*
2. *If $(p_i)$ is a convergent sequence in $M$ and $(g_i)$ is a sequence in $G$ such that $(g_i \cdot p_i)$ converges, then a subsequence of $(g_i)$ converges*
3. *For every compact subset $U \subseteq M$, the set $G_U = \{g \in G : (g \cdot U) \cap U \neq \emptyset\}$ is compact*

**Corollary 2.4.1.1.** *If a compact Lie group acts continuously on $M$, then the action is proper.*

**Corollary 2.4.1.2.** *If $G$ acts properly on $M$, then each orbit is a closed subset of $M$*

With the notion of a proper action, we have the pinnacle theorem:

**Proposition 2.4.2.** *If $G$ acts smoothly, freely, and properly on $M$, then the orbit space $M/G$ is a smooth manifold of dimension $dim(M) - dim(G)$; its smooth structure is unique with the property that $\pi : M \to M/G$ is a smooth submersion (where $\pi$ sends a point to its orbit).*

The four above results are proven in [2]. We have analogous definitions of properness of actions and quotient manifolds by discrete Lie groups given via Theorems $C.2$ and $C.3$. We can now consider two examples:

1. Quotient is not a smooth manifold:

   Let $\mathbb{R}^+$ act on $\mathbb{R}^n$ by $r \cdot (x^1, \ldots, x^n) = (rx^1, \ldots, rx^n)$. Note that the orbits are infinite rays and 0, thus the orbits are not closed subsets so the action is not proper (the contrapositive of Corollary 2.4.1.2). However, there is a subspace $X \subseteq \mathbb{R}^n/\mathbb{R}^+$ that is homeomorphic to $\mathbb{S}^{n-1}$:

   Let $X$ be the set of all orbits except for the orbit of the origin; namely, $X$ contains all infinite rays. Let $v \in \mathbb{R}^n$, we will denote its orbit as $[v]$; we can define a map $f : X \to \mathbb{S}^{2n-1}$ by $f([v]) = \frac{v}{|v|}$ (which is defined since $[0] \notin X$). $f$ is clearly continuous and surjective, thus it is a quotient map. We also have another quotient map: $\pi : \mathbb{R}^n - \{0\} \to X$. By Theorem $C.4$, we have that since $\pi$ and $f$ are constant on each others fibers, $\exists \varphi$, a homeomorphism, such that $\varphi \circ f = \pi \Rightarrow X$ is homeomorphic to $\mathbb{S}^{2n-1}$.

   $f$ and $\pi$ are constant on each other's fibers: $f(p) = f(q) \Leftrightarrow [p] = [q] \Leftrightarrow \pi(p) = \pi(q)$

2. Quotient is a smooth manifold:

   Recall the Hopf action defined earlier; we will now show that the action is free and proper (recall that smoothness has already been shown), and thus $\mathbb{S}^{2n+1}/\mathbb{S}^1$ is a smooth manifold (by proposition 2.4.2):

   (a) Freeness of $\mathbb{S}^1 \times \mathbb{S}^{2n+1} \to \mathbb{S}^{2n+1}$.
   Suppose $e^{2\pi i\phi} \cdot (e^{2\pi i\phi_1}, \ldots, e^{2\pi i\phi_{n+1}}) = (e^{2\pi i(\phi_1+\phi)}, \ldots, e^{2\pi i(\phi_{n+1}+\phi)}) = (e^{2\pi i\phi_1}, \ldots, e^{2\pi i\phi_{n+1}})$
   We then have that $e^{2\pi i(\phi_i+\phi)} = e^{2\pi i\phi_i} \Rightarrow \phi \in \mathbb{Z} \Rightarrow e^{2\pi i\phi} = 1$; thus the action is free.

   (b) Properness: $\mathbb{S}^1$ is compact and the action is continuous, thus the action is proper by Corollary 2.4.1.1

   Consider the natural quotient map $\pi : \mathbb{C}^{n+1} - \{0\} \to \mathbb{CP}^n$, but restricted to $\mathbb{S}^{2n+1}$ in $\mathbb{C}^{n+1}$, which is well known to be a smooth submersion. We make a quick note about the fibers of $\mathbb{CP}^n$ in $\mathbb{C}^{n+1}$: Let $\pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$ be the projection $\pi(q) = [q]$.

   Consider $\pi^{-1}([q]) = \{\lambda q : \lambda \in \mathbb{C}\}$; we can thus write $\lambda = e^{2\pi i\phi} \in \mathbb{S}^1$ and notice that the fibers of $\mathbb{CP}^n$ in $\mathbb{S}^{2n+1}$ are disjoint circles.

   Because the Hopf action's obits are disjoint circles in $\mathbb{C}^{n+1}$ and the fibers of $\mathbb{CP}^n$ are disjoint circles in $\mathbb{C}^{n+1}$, it is straightforward to see that they are constant on each other's fibers, thus there is a diffeomorphism $\varphi : \mathbb{S}^{2n+1}/\mathbb{S}^1 \to \mathbb{CP}^n$ such that we can construct the following diagram:

$$\begin{tikzcd} \mathbb{S}^{2n+1} \arrow[d, "\varphi \circ \pi"'] \arrow[dr, "\pi"] & \\ \mathbb{CP}^n & \mathbb{S}^{2n+1}/\mathbb{S}^1 \arrow[l, "\varphi"] \end{tikzcd}$$

**Definition of Principal $G$-Bundles**

Before we give the definition of a principal $G$-bundle, we first describe the more general case of a *fiber bundle*:

**Definition.** *A **fiber bundle over** $M$ is a topological space $E$ with a surjective and continuous $\pi : E \to M$ such that for each $x \in M$, there is a neighborhood $x \in U \subseteq M$ and a homeomorphism, entitled the **local trivialization**, $\Phi : \pi^{-1}(U) \to U \times F$, thus we have the following diagram:*

$$\begin{tikzcd} \pi^{-1}(U) \subseteq E \arrow[rr, "\Phi"] \arrow[dr, "\pi"'] & & U \times F \arrow[dl, "\mathbb{1}_U"] \\ & U \subseteq M & \end{tikzcd}$$

*We say that $E$ is the total space, $M$ is the base space, and $F$ is the fiber (where $\pi^{-1}(x) \in E$ is homeomorphic to $F$).*

**Proposition 2.4.3.** *Suppose $G$ acts smoothly, freely, and properly on $M$. We have that $M$ is the total space of a smooth fiber bundle with base $M/G$ and projection $\pi : M \to M/G$*

*Proof.* By Proposition 2.4.2, we immediately know that $M/G$ is a smooth manifold of dimension $dim(M) - dim(G)$. Let $U \subseteq M/G$, we have that $\pi^{-1} : U \to \pi^{-1}(U) \subseteq M$ and we can define another projection $\mathbb{1}_U : U \times G \to U$ as the projection onto the $U$ component. We have that $\varphi^{-1}$ from the proof of the previous proposition is a homeomorphism, thus we have:

$$\begin{tikzcd} \pi^{-1}(U) \subseteq P \arrow[rr, "\varphi^{-1}"] \arrow[dr, "\pi"'] & & U \times G \arrow[dl, "\mathbb{1}_U"] \\ & U \subseteq M/G & \end{tikzcd}$$

□

**Definition.** *A **principal $G$-bundle** is a fiber bundle, $P$, over $M$ (with projection $\pi : P \to M$) such that*

1. *The model fiber is a Lie group $G$, which acts on $P$ (with $\pi$ invariant and the trivialization, hereafter denoted $\varphi_i$, equivariant with respect to the action)*
2. *$G$ acts on the fibers of $M$ freely and transitively*
3. *$P/G$ forms a smooth manifold where $P/G = M$ (this does imply that $M$ acts on $P$ properly)*

*We thereby have the following diagram:*

$$\begin{tikzcd} \pi^{-1}(U_i) \subseteq P \arrow[rr, "\varphi_i"] \arrow[dr, "\pi"'] & & U_i \times G \arrow[dl, "proj_{U_i}"] \\ & U_i \subseteq M = P/G & \end{tikzcd}$$

The first two points above give us that the fibers, $\pi^{-1}(x) \subset P$, are homeomorphic to $G$ (at times, the fiber over $x$, $\pi^{-1}(x)$ will be referred to as $P_x$). Thus, if $\{U_i\}$ is an open cover for $M$, then we have $G$-equivariant diffeomorphisms $\psi_i : \pi^{-1}(U_i) \to U_i \times G$.

If $G$ acts properly, then the quotient $P/G$ is a manifold, which implies that $G$ acts on the fibers of $M$ freely and transitively. Thus if the action is smooth, free, and proper, then we are guaranteed the second item in the above definition and $P$ forms a principal $G$-bundle over $P/G$.

One quick example is the following: we have already proven that $\mathbb{S}^1$ acts on $\mathbb{S}^{2n+1}$ such that the quotient, $\mathbb{S}^{2n+1}/\mathbb{S}^1$, forms a smooth manifold. In fact, if we identify $\mathbb{S}^1$ as $U(1)$, we see that $\mathbb{S}^{2n+1}$ forms a principal $U(1)$-bundle over $\mathbb{CP}^n$.

In addition, suppose that for some $i, j$, we have that $U_i \cap U_j \neq \emptyset$; then we have to viable ways to locally trivialize our bundle:

$$\begin{tikzcd}
U_{ij} \times G \arrow[dr, "proj_{U_{ij}}"'] & \pi^{-1}(U_{ij}) \arrow[l, "\varphi_j"'] \arrow[r, "\varphi_i"] \arrow[d, "\pi"'] & U_{ij} \times G \arrow[dl, "proj_{U_{ij}}"] \\
 & U_{ij} &
\end{tikzcd}$$

Let $p \in \pi^{-1}(m), m \in U_{ij}$; then we must have

$$proj_{U_{ij}} = proj_{U_{ij}} \circ \varphi_j \circ \varphi_i^{-1} \Rightarrow \varphi_j \circ \varphi_i^{-1}(m, g) = \varphi_j(\pi^{-1}(m), \gamma_i^{-1}(m)) = (m, \sigma(m, g))$$

for some smooth map $\sigma : U_{ij} \times G \to G$ and an equivariant $\gamma_i : \pi^{-1}(U_i) \to G$. Since $\psi_j$ and $\psi_i$ are local diffeomorphisms, we must have that $g \mapsto \sigma(m, g)$ can be written as $g \mapsto \tau \cdot g$ for some $\tau \in G$ that depends on $U_{ij}$ and $p$. Therefore, we have that $\gamma_i(p) = \tau_{ij}(p)\gamma_j(p)$. We thus have that $\tau_{ij}(p) = \gamma_i(p)\gamma_j(p)^{-1}$

**Definition.** *These functions $\tau_{ij}(p)$ are called the **principal transition functions***

We have a few results about these principal transition functions:

1. $\tau_{ij}(p)$ is constant on fibers of $m \in M$: $\tau_{ij}(pg) = \gamma_i(pg)\gamma_j(pg)^{-1} = \gamma_i(p)gg^{-1}\gamma_j(p) = \gamma_i(p)\gamma_j(p) = \tau_{ij}(p)$
2. $\tau_{ij}(p)\tau_{ji}(p) = \gamma_i(p)\gamma_j(p)^{-1}\gamma_j(p)\gamma_i(p)^{-1} = e$
3. $\tau_{ij}(p)\tau_{jk}(p)\tau_{ki}(p) = \gamma_i(p)\gamma_j(p)^{-1}\gamma_k(p)^{-1}\gamma_k(p)\gamma_i(p)^{-1} = e$ where $p \in U_{ijk} = U_i \cap U_j \cap U_k$

These constraints are called the **cocycle constraints**. In fact, we are able to recover our original principal $G$-bundle $P$ by defining $P = \coprod_i (U_i \times G)/\sim$ where $\sim$ is the equivalence relation $(m, g) \sim (m, \alpha_{ij}(m)g)$. In order to see this, we review the analogous fact for vector bundles.

### The Cocycle Constraints

**Proposition 2.4.4.** *Let $\pi : E \to M$ be a smooth vector bundle of rank $k$ over $M$. Suppose that $\{U_i\}_{i \in I}$ is an open cover of $M$ and for each $i \in I$ we are given a smooth local trivialization $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$ of $E$. For each $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$, let $\tau_{ij} : U_i \cap U_j \to GL_k(\mathbb{R})$ be the transition function defined by* 3 *above. We have that $\tau_{ij}(p)\tau_{jk}(p) = \tau_{ik}(p)$ where $p \in U_i \cap U_j \cap U_k$*

*Proof.* We immediately have $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$, where $i$ is arbitrary, so we can define analogous $\Phi_j, \Phi_k$. We assume that $U_i \cap U_j \cap U_k \neq \emptyset$ and proceed:

$\Phi_i \circ \Phi_j^{-1}(p,v) = (p, \tau_{ij}(p))$ and $\Phi_j \circ \Phi_k^{-1}(p,v) = (p, \tau_{jk}(p)v)$. We thus have $(\Phi_i \circ \Phi_j^{-1}) \circ (\Phi_j \circ \Phi_k^{-1})(p,v) = (p, \tau_{ij}(p)\tau_{jk}(p)v)$. On the other hand, $(\Phi_i \circ \Phi_j^{-1}) \circ (\Phi_j \circ \Phi_k^{-1})(p,v) = \Phi_i \circ \Phi_k^{-1}(p,v) = (p, \tau_{ik}(p)v) \Rightarrow \tau_{ij}(p)\tau_{jk}(p)v = \tau_{ik}(p)v \forall v \Rightarrow \tau_{ij}(p)\tau_{jk}(p) = \tau_{ik}(p)$ $\square$

We can now prove the main result of this section:

**Proposition 2.4.5.** *Let $M$ be a smooth manifold and $\{U_i\}$ be an open cover for $M$. Suppose that for each $i, j$, we are given a smooth map $\tau_{ij} = U_i \cap U_j \to GL(k, \mathbb{R})$ such that $\tau_{ij}(p)\tau_{jk}(p) = \tau_{ik}(p)$ for $p \in U_i \cap U_j \cap U_k$ for all $i, j, k$. We then have that there is a smooth rank $k$ vector bundle, $E$, over $M$ with smooth local trivializations $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$ whose transition functions are $\tau_{ij}$*

*Proof.* Let $U = \coprod_{i \in I}(U_i \times \mathbb{R}^k)$ where $\{U_i\}_{i \in I}$ is an open cover of $M$. Define an equivalence relation on $U$ by $(p,v)_i \sim (q,w)_j$ (where the subscript denotes which set in the open cover the point belongs to) if and only if $p = q$ and $v = \tau_{ij}(p)w$. Note that because of 3, we have that $\tau_{ii}^2(p) = \tau_{ii}(p) \Rightarrow \tau_{ii} \equiv \mathbb{1}$ (reflexivity); $v = \tau_{ij}(p)w \Rightarrow w = \tau_{ji}(p)v$ (symmetry); and transitivity is, in fact, the entire point of Proposition 2.4.4; therefore the relation is indeed an equivalence relation.

The argument of the proof is to employ Theorem C.5, so we will need to construct vector space $E_p$ and bijections $\Phi_i$ satisfying the various requirements of Theorem C.5.

Let $E = U/\sim$ be the quotient space of equivalence classes and $\varphi : U \to E$ be the quotient map. We define $E_p = \varphi(\{p\} \times \mathbb{R}^k)$ for any given point $p \in M$. We seek to define a vector space structure for each $E_p$, so we aim to define scalar multiplication and vector addition. For multiplication, consider the map $\mathbb{R} \times \varphi^{-1}(E_p) \mapsto E_p$ by $(c, (p,v)_i) \mapsto \varphi((p, cv)_i)$. To show it is well defined, consider $(p,v)_i \sim (p,w)_j$, we have that $\tau_{ij}(p)w = v$ and $\tau_{ij}(p)cw = c\tau_{ij}(p)w = cv$, thus we have scalar multiplication. For vector addition, we define the map $\varphi^{-1}(E_p) \times \varphi^{-1}(E_p) \mapsto \varphi^{-1}(E_p)$ by $((p,v)_i, (p,w)_i) \mapsto (p, v+w)_i$.

By assumption, we have an open cover for $M$, thus we need only check criteria 2 and 3. To attain these bijective maps, $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$ with the properties 2 an 3, let's write out the maps we currently have:

$$\begin{tikzcd}
U_i \times \mathbb{R}^k \subseteq U \arrow[d, "\varphi"'] \arrow[dr, "\mathbb{1}_{U_i \times \mathbb{R}^k}"] & \\
\pi^{-1}(U_i) \subseteq E \arrow[r, "\Phi_i"] \arrow[d, "\pi"'] & U_i \times \mathbb{R}^k \arrow[dl, "\mathbb{1}_{U_i}"] \\
U_i \subseteq M &
\end{tikzcd}$$

The map $\pi$ is defined as follows: we can define $\Pi : U \to M$ by $\Pi((p,v)_i) = p$ , thus we can view $\Pi$ as the composition of $\pi \circ \varphi$ where $\pi$ takes an equivalence class $[p,v] \mapsto p$. We thus define $\Phi_i$ as the unique bijection satisfying $\Phi_i \circ \varphi = \mathbb{1}_{U_i \times \mathbb{R}^k}$. We thus check criteria 2 and 3:

2. Because $\varphi$ is linear, we have that $\Phi_i$ is linear as well, clearly it is thus a vector space isomorphism when restricted to $E_p$.
3. Consider $\Phi_i \circ \Phi_j^{-1}(p, v) = \Phi_i(q(p, v)_j) = \Phi_i(q(p, \tau_{ij}(p)v)_i) = (p, \tau_{ij}(p)v)$, thus criterion 3 is confirmed

We can therefore apply Theorem $C.5$ and confirm that $E$ is a rank $k$ vector bundle over $M$.

□

We can thus construct an analogous definition of a principal $G$-bundle:

**Definition.** *$P$ is a* ***principal $G$-bundle****, if we have the following data:*

1. *A locally finite, open cover $\{U_i\}_{i\in I}$ of $M$*
2. *$\{\tau_{ij} : U_i \cap U_j \to G\}$ where each map $\tau_{ij}$ is smooth satisfies the cocycle constraints:*
   (a) *$\tau_{ii} \equiv \mathbb{1} \in G$*
   (b) *$\tau_{ij}^{-1} = \tau_{ji}$ in a multiplicative sense. Namely, $\tau_{ij}^{-1}\tau_{ij} \equiv \mathbb{1}$ (of course, $\tau_{ij} \circ \tau_{ji}$ is not even defined).*
   (c) *If $U_i \cap U_j \cap U_k \neq \emptyset$, then we must have $\tau_{ij}\tau_{jk}\tau_{ki} \equiv \mathbb{1}$*

The $\tau_{nm}$ listed above are the *principal bundle transition functions* introduced just before this section on the cocycle constraints. This definition of a principal $G$-bundle is, in fact, synonymous with our original definition; a proof of this is given in [3].

#### 2.4.4 Examples of principal $G$-bundles: The Frame bundle and Generalized Hopf Fibrations

1. The Frame bundle, $F(TM)$ over $M$:

   Suppose $\{U\}_i$ is an open cover for $M$; we will construct the diagram:

$$\begin{tikzcd} P_{GL(TU_i)} \arrow[r, "\psi_i"] \arrow[d, "\pi"'] & U_i \times GL(n, \mathbb{R}) \arrow[dl, "proj"] \\ U_i \subseteq M & \end{tikzcd}$$

   A point $p \in P_{GL(TU_i)}$ is of the form $(v_1\big|_p, \ldots, v_n\big|_p)$, where $v_i\big|_p \in TM$ and $(v_1\big|_p, \ldots, v_n\big|_p)$ is a tuple of linearly independent vectors.

   Recall that $TM$ is a vector bundle over $M$, with local trivializations $\varphi : TM \to M \times \mathbb{R}^n$, or $\varphi_p : T_pM \to \{p\} \times \mathbb{R}^n$; thus $\varphi(v_i) \in \mathbb{R}^n$.

   We then define $\pi : F(TM) \to U_i$ by $\pi(v_1\big|_p, \ldots, v_n\big|_p) = p$; furthermore, we define $\psi_i : F(TM) \to U_i \times GL(n, \mathbb{R})$ by $\psi_i(v_1\big|_p, \ldots, v_n\big|_p) = (p, A)$ where $A \in GL(n, \mathbb{R})$ is given by:

$$A = \begin{bmatrix} \varphi_p(v_1\big|_p) & \ldots & \varphi_p(v_n\big|_p) \end{bmatrix}$$

We have that $GL(n, \mathbb{R})$ acts on $\pi^{-1}(x)$ transitively is trivial, because if we have another frame $(w_1\big|_p, \ldots, w_n\big|_p)$ and a matrix $B$ defined analogously to $A$ above, then we can find a change of basis matrix $C \in GL(n, \mathbb{R})$ such that $A = CB$.

2. $\mathbb{S}^n$ as a principal $O(1)$-bundle over $\mathbb{RP}^n$:

   First, recall that $O(1) = \{A \in GL(1, \mathbb{R}) : det(A) = \pm 1\} \cong \mathbb{Z}_2$. Next, recall the projection $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1} \to \mathbb{RP}^n$ given by $p \in \mathbb{S}^n$ and considering the line connection $p \in \mathbb{R}^{n+1}$ to $-p \in \mathbb{R}^{n+1}$. This line lies in $\mathbb{RP}^n$ by definition. Therefore, we can show that $\mathbb{S}^n$ forms a principal $O(1)$-bundle over $\mathbb{RP}^n$ by showing that $O(1)$ acts freely and transitively on the fibers of $\mathbb{RP}^n$ in $\mathbb{S}^n$. It is trivial that the projection $\mathbb{S}^n \mapsto \mathbb{S}^n/O(1)$ and the projection $\mathbb{S}^n$ are constant along each others fibers because they are the same map.

   The action $O(1) \times \mathbb{S}^n \mapsto \mathbb{S}^n$ is given by $(\pm 1, p) \mapsto \pm p$. Given $[p] \in \mathbb{RP}^n$, we have that its fiber $\pi^{-1}([p]) = \{p \in \mathbb{S}^n : \pi(p) = [p] \Leftrightarrow \pi(p) = \pm p\}$. Clearly $\mathbb{Z}_2$ acts freely on $\pi^{-1}([p])$ (let $x \in \mathbb{Z}_2$ such that $x \cdot p = \pm 1p = p \Rightarrow x = 1$). Furthermore, the action if clearly transitive since the fiber of $[p]$ is made up of points $\pm p$ and $\mathbb{Z}_2 = \pm 1$.

   A way to see that $\mathbb{S}^n/O(1)$ forms a smooth manifold is to employ theorems $C.2$ and $C.3$; we will check that the action is proper here, though smoothness is trivial (since multiplication in $\mathbb{R}^{n+1}$ is smooth and the action is merely $p \in \mathbb{S}^{n+1}$ multiplied by $\pm 1$) and freeness has already been shown. Theorem $C.2(1)$ is fulfilled so long as $U$ is fully contained in the same hemisphere as $p$; analogously, theorem $C.2(2)$ is fulfilled so long as $V$ and $W$ are contained within balls on the surface of $\mathbb{S}^n$ of radius smaller than $min($ distance between $p$ and $q$, distance between $p$ and $-q)$. We thus have, by Theorem $C.3$, that $\mathbb{S}^n$ forms a principal $O(1)$-bundle over the quotient.

3. $\mathbb{S}^{2n+1}$ as a principal $U(1)$-bundle over $\mathbb{CP}^n$:

   We have already seen that this, the Hopf action, is smooth, free, and proper, so the quotient manifold is well defined and is diffeomorphic to $\mathbb{CP}^n$. Another way of seeing $\mathbb{S}^{2n+1}$ as a principal $U(1)$-bundle is that $U(1)$ acts freely and transitively on the fibers of $\mathbb{S}^{2n+1}$:

   Since the fibers of $\mathbb{CP}^{n+1}$ are complex unit circles (as we saw in example 2 after Proposition 3.1.2), thus $\mathbb{S}^1$ acts transitively on these fibers.

4. $\mathbb{S}^{4n+3}$ as a principal $SU(2)$-bundle over $\mathbb{HP}^n$:

   Consider $\mathbb{H} = \mathbb{C} \times \mathbb{C}$ to be the quaternions; we can thus consider $\mathbb{H}^{n+1}$ (where $\mathbb{H}$ is the set of quaternions) as $\mathbb{R}^{4(n+1)}$. A given $\alpha \in \mathbb{S}^{4n+3} \subset \mathbb{H}^{n+1}$ is given by $\alpha = (\alpha_1, \ldots, \alpha_{n+1})$ where $\alpha_i = a_i\mathbb{1} + b_i + c_i + d_i\mathbb{k}$ $(a_i, b_i, c_i, d_i \in \mathbb{R})$ and $\sum_i a_i^2 + b_i^2 + c_i^2 + d_i^2 = 1$.

   By considering $\mathbb{H}$ as the vector space $\mathbb{C}\times\mathbb{C}$ over $\mathbb{C}$, we can think of $\alpha_i$ as a tuple $(\gamma, \delta) \in \mathbb{C}\times\mathbb{C}$. We thereby define an action of $SU(2)$ on $\mathbb{H}^{n+1}$ by matrix multiplication of the elements $\alpha_i$: $A \cdot \alpha = (A\alpha_1, \ldots, A\alpha_{n+1})$

   By identifying $SU(2)$ with $\int$, the unit quaternions, we see that an action of $SU(2)$ on amounts to multiplying a point on $\mathbb{S}^{4n+3} \hookrightarrow \mathbb{H}^{n+1}$ by a unit quaternion, which does not alter the distance of the point from the origin (so the point $A\alpha$ is on $\mathbb{S}^{4n+3}$).

   The action is trivially smooth, since multiplication in $\mathbb{C} \times \mathbb{C}$ is smooth; the action is proper because $SU(2)$ is compact. Suppose that $q \in \int$ gives us $q \cdot \alpha = \alpha$; we then have that $q\alpha_i = \alpha_i$,

$\forall i \Rightarrow q = 1$, thus the action is free. We therefore have that the smooth manifold $\mathbb{S}^{4n+3}/SU(2)$ or, equivalently, $\mathbb{S}^{4n+3}/\int$, is well defined and we thereby have a quotient map $\pi : \mathbb{S}^{4n+3} \to \mathbb{S}^{4n+3}/SU(2)$. We now observe that the fibers of this quotient map are $\pi^{-1}(p) = \{qp : q \in SU(2) \cong \int\}$.

We also have a quotient map $\pi_1 : \mathbb{H}^{n+1} \to \mathbb{HP}^n$ that sends $\alpha \mapsto [\alpha]$ where $[\alpha]$ denotes the line in $\mathbb{H}^{n+1}$ passing through the origin and $\alpha$. This maps clearly restricts to $\mathbb{S}^{4n+3}$ and is clearly surjective on this restriction. The fibers of $\pi_1$ are $\pi_1^{-1}(p) = \{qp : q \in \mathbb{H}\}$, so clearly these maps are constant on each others fibers; we therefore have that $\mathbb{S}^{4n+3}/SU(2)$ is diffeomorphic to $\mathbb{HP}^n$.

The final three examples above are known as **generalized Hopf fibrations**; the case $n = 1$ was discovered by Heinz Hopf in the first half of the 20th century served as an intriguing motivation for the study of fiber bundles. $\mathbb{RP}^n$, $\mathbb{CP}^n$, and $\mathbb{HP}^n$ are of great interest in many fields of mathematics and physics, and thus their structure as principal bundles offers a different way to shed light on their structure and properties.

# 3 Lie Algebras

## 3.1 Lie Derivatives

Let $V, W$ be smooth vector fields on $M$ with flows $\theta(t), \gamma(t)$, respectively. The Lie derivative, $\mathcal{L}_V W$, is a type of directional derivative for vector fields; it is important to note that this is *one* type of a directional derivative, and we will encounter another type when we define a connection.

$\mathcal{L}_V W$ provides information about how the vector field $W$ changes along the flow of $V$. The subscript is measured with respect to the argument, similarly to an orthogonal projection, $proj_{\vec{b}}(\vec{a})$, where the input $\vec{a}$ is projected *onto* $\vec{b}$. Let $\theta_t$ be the flow of V; at a point $p$, we define:

$$(\mathcal{L}_V W)_p = \frac{d}{dt}\bigg|_{t=0} d(\theta_{-t})_{\theta_t(p)}(W_{\theta_t(p)}) = \lim_{t \to 0} \frac{d(\theta_{-t})_{\theta_t(p)}(W_{\theta_t(p)}) - W_p}{t}$$

This operation yields a new vector field, whose magnitude and direction tell us how $W$ changes along the flow of $V$.

### Lie Brackets

Those familiar with quantum mechanics will recognize $[V, W]$ as the commutator: $[V, W] = VW - WV$. Bringing up the bracket is not so random:

**Proposition 3.1.1.** $(\mathcal{L}_V W)_p = \lim_{t \to 0} \frac{d(\theta_{-t})_{\theta_t(p)}(W_{\theta_t(p)}) - W_p}{t} = [V, W]_p$

A proof of this can be found in [5].

## 3.2 Introduction of Lie Algebras

**Definition.** *A **Lie Algebra** is a vector space along with a bilinear product* $[\cdot, \cdot]$ *with the following properties:*

1. $[V, W] = -[W, V]$
2. $[\cdot, \cdot]$ *is bilinear over* $\mathbb{R}$
3. $[X_1, [X_2, X_3]] + [X_3, [X_1, X_2]] + [X_2, [X_3, X_1]] = 0$ *(this is known as the* ***Jacobi identity****)*

Indeed, one can define Lie algebras *independently* of Lie groups, and the study of Lie algebras by themselves is rich. Our purpose will solely be to use Lie algebras to uncover structures encrypted in Lie groups. We will hereafter refer to arbitrary Lie algebras as $\mathfrak{g}$ and $\mathfrak{h}$; we begin with some basic definitions:

**Definition.**

1. $\mathfrak{h} \subseteq \mathfrak{g}$ *is a* ***Lie subalgebra*** *if it is a vector subspace and is closed under brackets*
2. $\varphi : \mathfrak{h} \to \mathfrak{g}$ *is a* ***Lie algebra homomorphism*** *if* $\varphi([V, W]) = [\varphi(V), \varphi(W)]$.

By definition, any vector space that is closed under the Lie bracket is a Lie algebra, but we will hereafter refer to as the **Lie algebra, $\mathfrak{g}$, of the Lie group** $G$ as the vector space of left invariant vector fields on $G$. This is a Lie subalgebra of the infinite-dimensional Lie algebra $\mathfrak{X}(M)$; we will hereafter assume that $\mathfrak{g}$ is the Lie algebra of some Lie group $G$.

Consider two left-invariant vector fields, $V, W$, on a Lie group $G$. It is intuitive that if $V_e = W_e$ (their evaluation at the identity element), then $V = W$.

**Proposition 3.2.1.** *The function* $\varepsilon : \mathfrak{g} \to T_eG$ *defined by* $\varepsilon(X) = X_e$ *is a vector space isomorphism; therefore,* $dim(\mathfrak{g}) = dim(T_eG) = dim(G)$*. If* $\mathfrak{h} \subset \mathfrak{g}$ *is a Lie subalgebra, then we can identify* $Lie(H) \cong \mathfrak{h}$ *with left invariant vector fields on* $G$ *that are tangent to* $H$ *at the identity.*

The proof of this corollary can be found in [2]. This allows us to easily determine left invariant vector fields on Lie groups. We simply have to identify the tangent space of the manifold, $G$, at the identity. Let's look at the case when $n = 2$:

The main point here is obvious once we recall that any element $A \in SU(2)$ is of the form:

$$A = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}$$

where $\alpha\bar{\alpha} + \beta\bar{\beta} = 1$. If we let $\alpha = x + iy; \beta = z + iw$, we then have that $x^2 + y^2 + z^2 + w^2 = 1$ and the next proposition is clear.

**Proposition 3.2.2.** $SU(2)$ *is topologically* $\mathbb{S}^3$.

The structure of these groups as smooth manifolds and notions such as their tangent spaces are clearer through examples such as the above. The beauty lies in the fact that differential geometric constructions apply to Lie groups; indeed we will encounter geodesics and curvature on Lie groups.

The next relation, between Lie groups and their algebras, to be introduced is the notion of an induced Lie algebra homomorphism.

**Proposition 3.2.3.** *If* $\varphi : G \to H$ *is a Lie group homomorphism then:*

1. *For each* $X \in \mathfrak{g}$*, there is a unique* $Y \in \mathfrak{h}$ *such that* $\varphi_* X = Y$

2. *In fact,* $\varphi_* : \mathfrak{g} \to \mathfrak{h}$ *is a Lie algebra homomorphism, called the* ***induced Lie algebra homomorphism***
3. *The induced homomorphism of* $\mathbb{1}_G$ *is* $\mathbb{1}_{\mathfrak{g}}$
4. *If* $\varphi : G \to H$ *and* $\phi : H \to M$ *are Lie group homomorphisms, then* $(\phi \circ \varphi)_* = (\phi)_* \circ (\varphi)_* : \mathfrak{g} \to \mathfrak{m}$

This proposition is proven in [2]. Before we compute some standard Lie algebras of Lie groups, we address a number of propositions that can be proven with the tool of induced Lie algebra homomorphisms:

**Proposition 3.2.4.** *If* $G$ *is abelian, then* $\mathfrak{g}$ *is abelian.*

*Proof.* Note that if $G$ is abelian, then the inversion map is a Lie group homomorphism: $i(gh) = (gh)^{-1} = (hg)^{-1} = g^{-1}h^{-1} = i(g)i(h)$ and its smoothness is given since $G$ is a Lie group, thus it is a Lie group homomorphism. From Identity B.2, we know that $i_* = di_e : T_eG \to T_eG$ is given by $di_*(X) = -X$. We then have $i_*$ as an induced Lie algebra homomorphism, and we compute: $[X,Y] = [-X,-Y] = [i_*X, i_*Y] = i_*[X,Y] = -[X,Y] = [Y,X] \Rightarrow \forall X,Y \in \mathfrak{g}, [X,Y] = 0$ thus $XY = YX$

□

**Proposition 3.2.5.** *Suppose* $\varphi$ *is a Lie group homomorphism, then we have that* $ker(\varphi_*) = Lie(ker(\varphi))$

*Proof.* $(\supseteq)$

Let $X \in Lie(ker(\varphi))$, we need to show that $X \in ker(\varphi_*)$. We have that $Lie(ker(\varphi))$ is a subalgebra of $\mathfrak{g}$ and thus $\iota_*X \in \mathfrak{g}$ (where $\iota$ is topological inclusion), then $\varphi_*\iota_*X = 0 \Rightarrow X \in ker(\varphi_*)$

$(\subseteq)$

Let $X \in ker(\varphi_*) \Rightarrow \varphi_*X = 0 \Rightarrow d\varphi_e X_e = 0 \Rightarrow X_e \in ker(d\varphi_e) \subseteq T_e(ker(\varphi)) = Lie(ker(\varphi))$

□

There is one particular Lie group homomorphism whose induced Lie algebra homomorphism is of particular importance: the determinant and the trace.

**Proposition 3.2.6.** *Consider* $det : GL(n,F) \to F$ *as a Lie group homomorphism:* $det_* = tr : \mathfrak{gl}(n,F) \to F$

In order to prove this proposition, we need to prove a few things:

**Proposition 3.2.7.**

1. *for any* $M \in M_n(\mathbb{R})$, $\frac{d}{dt}\Big|_{t=0} det(\mathbb{1}_n + tA) = tr(A) = \sum_i A^i_i$
2. *Let* $X \in GL(n,\mathbb{R})$ *and* $B \in T_xGL_n(\mathbb{R}) \cong M_n(\mathbb{R})$. $d(det)_X B = det(X)tr(X^{-1}B)$

*Proof.*

1.
$$det(A) = \sum_{\sigma \in S_n} (sgn(\sigma) A_1^{\sigma(1)} \dots A_n^{\sigma(n)}) \Rightarrow$$
$$det(\mathbb{1}_n + tA) = \sum_{\sigma \in S_n} (sgn(\sigma)(\mathbb{1}_n + tA)_1^{\sigma(1)} \dots (\mathbb{1}_n + tA)_n^{\sigma(n)}) = 1 + t(A_1^1 + A_2^2 + \dots + A_n^n) + O(t^2)$$

   Thus, we have:
$$det(\mathbb{1}_n + tA) = 1 + ttr(A) + O(n^2) \Rightarrow \left.\frac{d}{dt}\right|_{t=0} det(\mathbb{1}_n + tA) = tr(A)$$

2. Recall,
$$det(AB) = det(A)det(B) \Rightarrow det(X + tB) = det(X)det(\mathbb{1} + tX^{-1}B)$$

   Thus we have:
$$\left.\frac{d}{dt}\right|_{t=0} det(X + tB) = det(X) \left.\frac{d}{dt}\right|_{t=0} det(\mathbb{1}_n + tX^{-1}B) = det(X)tr(X^{-1}B)$$

   Let $\gamma(t) = \mathbb{1}_n + tB \Rightarrow \gamma'(0) = B$ and $X \in GL(n, \mathbb{R})$. Thus, by Theorem C.6,
$$d(det)_x B = \left.\frac{d}{dt}\right|_{t=0} (det \circ \gamma)(t) = \left.\frac{d}{dt}\right|_{t=0} det(\mathbb{1}_n + tB) = det(X)tr(X^{-1}B)$$

□

We can now prove Proposition 3.2.7.

*Proof.* First we will evaluate $d(det)$:
$det(A) = \sum_\sigma sgn(\sigma) A_1^{\sigma(1)} \dots A_n^{\sigma(n)}$. We also know that
$$det(\mathbb{1} + tA) = \sum_{\sigma \in S_n} sgn(\sigma)(\mathbb{1} + tA)_1^{\sigma(1)} \dots (\mathbb{1} + tA)_n^{\sigma(n)}$$
$$= 1 + t(A_1^1 + \dots + A_n^n) + O(t^2) = 1 + ttr(A) + O(t^2) \Rightarrow \frac{d}{dt}|_{t=0} det(\mathbb{1} + tA) = tr(A)$$

Now suppose $X \in GL(n, \mathbb{R})$ and $B \in T_X(GL(n, \mathbb{R})) \cong M_n(\mathbb{R})$. We claim that $d(det)_X B = det(X)tr(X^{-1}B)$.

Recall that $det(AB) = det(A)det(B) \Rightarrow det(X + tB) = det(X)det(\mathbb{1} + tX^{-1}B)$; thus
$$\frac{d}{dt}|_{t=0} det(X + tB) = det(X) \frac{d}{dt}|_{t=0} det(\mathbb{1} + tX^{-1}B = det(X)tr(X^{-1}B)$$

Let $\gamma(t) = \mathbb{1} + tB$ (note that $\gamma'(0) = B$), we have (by proposition A.7), $d(det)_X B = \frac{d}{dt}|_{t=0}(det \circ \gamma)(t) = \frac{d}{dt}|_{t=0} det(\mathbb{1} + tB)$.

Alas we can conclude that $det_*(X) = d(det)_{\mathbb{1}}(X) = det(\mathbb{1})tr(\mathbb{1}^{-1}X) = tr(X)$. □

**Proposition 3.2.8.** *(note that the following are all defined in Appendix A)*

1. $Lie(GL(n,F)) \cong \mathfrak{gl}(n,F)$
2. $Lie(SL(n,F)) \cong \mathfrak{sl}(n,F)$
3. $Lie(SO(n)) \cong \mathfrak{o}(n)$
4. $Lie(U(n)) \cong \mathfrak{u}(n)$
5. $Lie(SU(n)) \cong \mathfrak{su}(n)$

*Proof.*

1. This is standard and is included in [3] on pages 193-195.
2. $Lie(SL(n,F))$. We have that $det : GL(n,F) \to F^{\times}$ is a Lie group homomorphism with $ker(det) = SL(n,F)$. Thus $Lie(ker(det)) = ker(det_*) = ker(tr) \Rightarrow Lie(SL(n,F)) \cong \{A \in \mathfrak{gl}(n,F) | tr(A) = 0\}$. This Lie algebra will hereafter be referred to as $\mathfrak{gl}(n,F)$.
3. $Lie(SO(n))$. $SO(n) = \{A \in GL(n,\mathbb{R}) | A \in O(n) \cap SL(n,\mathbb{R})\}$ where $O(n) = \{A \in GL(n,\mathbb{R}) | A^T A = \mathbb{1}\}$. Consider the map $F(A) = A^T A$. We have that $Lie(SO(n)) \cong T_{\mathbb{1}} SO(n) = ker(dF)_{\mathbb{1}}$. Let $\gamma(t) = \mathbb{1} + tA \Rightarrow dF_{\mathbb{1}}(A) = \frac{d}{dt}|_{t=0} F \circ \gamma(t) = \frac{d}{dt}|_{t=0} (\mathbb{1} + tA)^T (\mathbb{1} + tA) = A^T + A$ Thus $A \in ker(dF)$ iff $A^T + A = 0$ thus $\mathfrak{so}(n) = \{A \in \mathfrak{gl}(n,\mathbb{R}) | A^T = -A\}$ i.e. the space of skew symmetric matrices.
4. $Lie(U(n))$. Consider the map $F(A) = A^* A$ where $A^* = \bar{A}^T$. Similarly $Lie(U(n)) \cong T_{\mathbb{1}} U(n) = ker(dF)_{\mathbb{1}}$, again we let $\gamma(t) = \mathbb{1} + tA \Rightarrow dF_{\mathbb{1}}(A) = \frac{d}{dt}|_{t=0} (F \circ \gamma(t)) = \frac{d}{dt}|_{t=0} (\mathbb{1} + tA)^* (\mathbb{1} + tA) = A^* + A$ thus $\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n,F) | A^* = -A = 0\}$
5. $Lie(SU(n))$. $SU(n) = SL_n \cap U(n)$; thus using the maps in 2 and 4, it is easy to confirm that $Lie(SU(n)) \cong Lie(SL_n) \cap Lie(U(n)) = \mathfrak{sl}(n \cap \mathfrak{u}(n) = \mathfrak{su}(n)$

□

**Proposition 3.2.9.** *We have isomorphisms $\mathfrak{su}(n) \cong \mathfrak{o}(n) \cong \mathbb{R}^3$ where $\mathbb{R}^3$ is considered as a Lie algebra with the cross product.*

*Proof.* $\mathfrak{su}_2 = \{S \in \mathfrak{gl}_2(\mathbb{C}) | S^* + S = 0 \text{ and } tr(S) = 0\}$, thus $S \in \mathfrak{su}_2 \Rightarrow \exists \alpha, \beta, \gamma \in \mathbb{C}$ such that $S = \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix}$. We must have

$$S^* + S = \begin{bmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & -\bar{\alpha} \end{bmatrix} + \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix} = \begin{bmatrix} 2Re(\alpha) & \bar{\gamma} + \beta \\ \bar{\beta} + \gamma & -2Re(\alpha) \end{bmatrix} = 0$$

Therefore $Re(\alpha) = 0$. Let $\beta = a + ib, \gamma = c + id$ We have $\bar{\gamma} + \beta = c - id + a + ib = (a+c) + i(b-d) = 0 \Leftrightarrow b = d$ and $a = -c$, note that this implies $\gamma + \bar{\beta} = 0$ (similarly $\gamma + \bar{\beta} = 0$ would give us the same constraints). Thus $S = \begin{bmatrix} ix & a+ib \\ -a+ib & -ix \end{bmatrix} \Rightarrow$ the Pauli spin matrices form a basis for $\mathfrak{su}_2$

Thus $\mathfrak{su}_2$ is spanned by the Pauli spin matrices. We can thus construct a map $\varphi(a,b,c) = \frac{1}{2} \begin{bmatrix} ia & -b+ic \\ b+ic & -ia \end{bmatrix}$. We have that

$$\varphi((a,b,c) \times (d,f,g)) = \varphi(bg - cf, cd - ag, af - bd) = \frac{1}{2} \begin{bmatrix} i(bg - cf) & ag - cd + i(af - bd) \\ cd - ag + i(af - bd) & -i(bg - cd) \end{bmatrix}$$

On the other hand,

$$[\varphi(a,b,c),\varphi(d,f,g)] = \frac{1}{4}\begin{bmatrix} 2i(bg-cf) & 2(ag-cd)+2i(af-bd) \\ 2(cd-ag)+2i(af-bd) & -2i(bg-cf) \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} i(bg-cf) & (ag-cd)+i(af-bd) \\ (cd-ag)+i(af-bd) & -i(bg-cf) \end{bmatrix}$$

Thus it is a Lie algebra isomorphism, and is very clearly a bijection.
Now, for $O \in \mathfrak{o}_3$, we have:

$$O^T + O = 0 \Rightarrow O = \begin{bmatrix} \alpha & a & d \\ \beta & b & e \\ \gamma & c & f \end{bmatrix}$$

and

$$O^T + O = \begin{bmatrix} 2\alpha & \beta+a & \gamma+d \\ \beta+a & 2b & c+e \\ \gamma+d & e+c & 2c \end{bmatrix} = 0 \Rightarrow \alpha=b=c=0, a=-\beta, d=-\gamma, c=-e \Rightarrow O = \begin{bmatrix} 0 & -\beta & -\gamma \\ \beta & 0 & -c \\ \gamma & c & 0 \end{bmatrix}$$

We thus construct a matrix $\varphi : \mathbb{R}^3 \to \mathfrak{o}_3$ by $\varphi(a,b,c) = \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix}$; thus we easily see that the generators of rotations from quantum mechanics form a basis for $\mathfrak{o}_3$. We check that it is a Lie algebra homomorphism:

$$\varphi((a,b,c)\times(d,f,g)) = \varphi(bg-cf, cd-ag, af-bd) = \begin{bmatrix} 0 & cf-bg & ag-cd \\ bg-cf & 0 & bd-af \\ cd-ag & af-bd & 0 \end{bmatrix}$$

On the other hand

$$[\varphi(a,b,c),\varphi(d,f,g)] = \begin{bmatrix} 0 & cf-bg & ag-cd \\ bg-cf & 0 & bd-fa \\ cd-ag & af-bd & 0 \end{bmatrix}$$

Thus it is a Lie algebra homomorphism, and obviously bijective.

□

When I took quantum mechanics, we approached angular momentum by learning about quantized angular momentum (both orbital and intrinsic spin), deriving spherical harmonics, and creating ensembles of quantum systems in various states. When I did the computation above, the mathematical similarity between spin and orbital momentum (such as, for example, the fact that their operators have identical commutation relations) “clicked” since their Lie algebras are isomorphic as algebras.

## 3.3 The Exponential Map

In this section, we give numerous results pertaining to a crucial topic in the study of the relationships between Lie groups and their Lie algebras: the exponential map. The exponential map arises as a generalization of the classic function $f(x) = e^x$ as it arises in the study of matrices. Many of the theorems stated here are standard, and thus stated without proof, and we use those results to prove numerous corollaries.

### 3.3.1 One Parameter Subgroups and the Exponential Map

First, recall that a maximal integral curve is a curve, $\gamma(t) : \mathbb{R} \to M$ such that $\gamma'(t) = X_p \in TM_p$ that is defined for the maximum possible interval $t \in (-\varepsilon, \varepsilon)$. If the vector field $X$ is **complete**, then $t$ is defined for all $\mathbb{R}$. We thus define:

**Definition.** *A **one-parameter subgroup** of a Lie group $G$ is a Lie group homomorphism $\gamma : \mathbb{R} \to G$ (note, these are not groups themselves, they are homomorphisms).*

**Proposition 3.3.1.** *A group homomorphism $\gamma : \mathbb{R} \to G$, i.e. the one-parameter subgroups of $G$, are the maximal integral curves of $X$ starting at $e \in G$, where $X \in \mathfrak{g}$*

*We thus say that $\gamma$ is the one-parameter subgroup generated by $X$.*

**Proposition 3.3.2.** *The one-parameter subgroups of $GL(n, \mathbb{R})$ generated by $A \in \mathfrak{gl}(n\mathbb{R})$ are $\gamma(t) = e^{tA}$*

**Definition.** *if $\gamma$ is the one-parameter subgroup generated by $X \in \mathfrak{g}$; we then define the exponential map as $exp : \mathfrak{g} \to G$ by $exp(X) = \gamma(1)$.*

**Proposition 3.3.3.** *The one-parameter subgroup, of a Lie group $G$, generated by $X$ are given by the exponential map: $\gamma(t) = exp(tX)$.*

**Proposition 3.3.4.** *The exponential map of $GL(V)$ is given by $exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$; in particular, the exponential map $exp : \mathfrak{gl}(n\mathbb{R}) \to GL(n, \mathbb{R})$ is given by $exp(X) = e^A$*

The four propositions above can be found [2].

## Integral Curves in $GL(3, \mathbb{R})$

1. Let $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$

   We have that $e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$. Since $A$ is diagonal, we easily see that $A^k = \begin{bmatrix} a^k & 0 & 0 \\ 0 & b^k & 0 \\ 0 & 0 & c^k \end{bmatrix}$.

   We now have

$$e^{tA} = \begin{bmatrix} \sum \frac{t^k}{k!} a^k & 0 & 0 \\ 0 & \sum \frac{t^k}{k!} b^k & 0 \\ 0 & 0 & \sum \frac{t^k}{k!} c^k \end{bmatrix} = \begin{bmatrix} e^{ta} & 0 & 0 \\ 0 & e^{tb} & 0 \\ 0 & 0 & e^{tc} \end{bmatrix} = \gamma(t)$$

2. Let $B = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$

   Note that $B^2 = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $B^3 = 0$. We thus have

$$e^{tB} = \mathbb{1} + tB + \frac{t^2}{2}B^2 = \begin{bmatrix} 1 & ta & t(b + \frac{t}{2}ac) \\ 0 & 1 & tc \\ 0 & 0 & 1 \end{bmatrix} = \gamma(t)$$

3. Let $C = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

   Some tedious matrix arithmetic would show that $C^{2k} = \begin{bmatrix} (-1)^k & 0 & 0 \\ 0 & (-1)^k & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and

   $C^{2k+1} = \begin{bmatrix} 0 & (-1)^k & 0 \\ (-1)^{k+1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

   Therefore

$$\sum_{k=0}^{\infty} \frac{1}{k!}C^k = \sum_{k=0}^{\infty} \frac{1}{k!}C^{2k} + \sum_{k=0}^{\infty} \frac{1}{k!}C^{2k+1} = \begin{bmatrix} \sum \frac{t^{2k}}{(2k)!}(-1)^k & 0 & 0 \\ 0 & \sum \frac{t^{2k}}{(2k)!}(-1)^k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & \sum \frac{t^{2k+1}}{(2k+1)!}(-1)^k & 0 \\ -\sum \frac{t^{2k+1}}{(2k+1)!}(-1)^k & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   This is precisely the curve in $SO(3)$ that describes rotations about the $\hat{z}$ axis. This curve was "generated" by $C \in \mathfrak{so}_3$, thus at this point it should be clearer what is meant by Lie algebras representing "infinitesimal transformations" and their role in quantum mechanics.

Here, we mention two brief asides:

1. We can check all of our solutions above by finding $\gamma'(0)$. Thus, we can find any infinitesimal transformation of a curve $\gamma(t)$ by simply taking $\gamma'(0)$.
2. Recall how the solution to the Schrodinger Equation in quantum mechanics, $-i\partial_t\psi = \hat{H}\psi$, is of the form $\psi(t) = e^{-i\hat{H}t/\hbar}\psi(0)$. We can reinterpret this well-known equation and solution in quantum mechanics in the following way: quantum systems evolve along the integral curves in some Lie group, where $-i\hat{H}/\hbar$ is in the Lie algebra of said Lie group. In quantum, people typically replace $e^{-i\hat{H}/\hbar}\left|\psi\right\rangle$ with $e^{-i\lambda/\hbar}\left|\psi\right\rangle$ where $\lambda$ is the eigenvalue of $\left|\psi\right\rangle$. We see through our first example above that this substitution is valid since we can diagonalize $\hat{H}$ over $\mathbb{C}$.

**Proposition 3.3.5.** *We have the following properties for the exponential map, where $\mathfrak{g} = Lie(G)$ and $X \in \mathfrak{g}$ is arbitrary:*

1. *$exp : \mathfrak{g} \to G$ is smooth*
2. *For any $s, t \in \mathbb{R}$, $exp((s+t)X) = exp(sX)exp(tX)$*
3. *If $n \in \mathbb{Z}$, then $(exp(X))^n = exp(nX)$*
4. *The map $dexp_0 : T_0\mathfrak{g} \to T_eG$ is the identity map*
5. *The exponential map restricts to a diffeomorphism from some neighborhood $U \in \mathfrak{g}$ containing $0$ to some neighborhood $V \in G$ containing $e$*
6. *If $H$ is another Lie group with $\mathfrak{h} = Lie(H)$ and $\varphi : G \to H$ is a Lie group homomorphism, then the following diagram commutes:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\varphi_*"] \arrow[d, "exp"'] & \mathfrak{h} \arrow[d, "exp"] \\ G \arrow[r, "\varphi"'] & H \end{tikzcd}$$

7. *The flow $\theta$ of $X$ is given by $\theta_t = R_{exp(tX)}$*

A proof of the above can be found in [2]. As a remark, we will now see that the exponential map is not necessarily surjective! We prove this in two steps:

1. Suppose $A \in SL(n, \mathbb{R})$ is of the form $A = e^B$ where $B \in \mathfrak{sl}(n, \mathbb{R})$. There exists a matrix $C$ such that $C^2 = A$.
2. Let $A = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & -2 \end{pmatrix}$; there is no matrix $B \in \mathfrak{sl}(n, \mathbb{R})$ such that $A = e^B$.

*Proof.*

1. Note that if $B \in \mathfrak{sl}(n, \mathbb{R})$ then we define $D = \frac{1}{2}B \in \mathfrak{sl}(n, \mathbb{R})$, thus $e^D \in SL(n, \mathbb{R})$, so we define $C = e^D$. We have that $C^2 = (e^D)^2 = e^{2D} = e^B = A$. Note, this equality is made valid by Proposition 3.3.5(3) above because it ensures that $e^De^D = e^{2D}$.
2. Take $A = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -2 \end{bmatrix}$,

   if we can show that there is no $C$ with $C^2 = A$, then we have shown that the exponential map is not surjective (namely that there is no $B$ such that $A = e^B$, this is essentially the contrapositive of the above claim: if there does not exists $C^2$ then there does not exist $B$).

   Let $C = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in SL_2(\mathbb{R})$ such that $C^2 = \begin{bmatrix} a^2 + cb & ac + cd \\ ab + bd & cb + d^2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -2 \end{bmatrix}$. We therefore have:

   (a) $ab = -bd$

(b) $a^2 + cb = -\frac{1}{2}$
(c) $cb + d^2 = -2$
(d) $ad - bc = 1$

The last equation gives us that $bc = ad - 1$, which, when combined with the first equation, gives: $a^2 + ad = \frac{1}{2}$. We also have that $a = -d$ so $a^2 - a^2 = 0 \neq \frac{1}{2}$. If $b = 0$ then $a^2 + cb = a^2 = -\frac{1}{2} \Rightarrow a \notin \mathbb{R}$

□

**Corollary 3.3.5.1.** $det(e^A) = e^{tr(A)}$

*Proof.* We know that $\det : GL(n, F) \to F^\times$ is a Lie group homomorphism and $det_* = tr$, thus we have that the following diagram commutes, and the corollary is immediate:

$$\begin{tikzcd}
\mathfrak{gl}(n\mathbb{R}) \arrow[r, "det_* = tr"] \arrow[d, "exp"'] & \mathbb{R} \arrow[d, "exp"] \\
GL(n, \mathbb{R}) \arrow[r, "det"'] & \mathbb{R}^*
\end{tikzcd}$$

□

**Proposition 3.3.6.** *Let $X, Y \in \mathfrak{g}$; we thus have that $exp(x)exp(y) =$
$exp(\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{r_i+s_i>0} \frac{[X^{r_1}Y^{s_1}X^{r_2}Y^{s_2}\cdots X^{r_n}Y^{s_n}]}{\sum_{i=1}^{n}(r_i+s_i)\cdot \Pi_{i=1}^{n} r_i!s_i!}) = exp(X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] +$
$[Y, [Y, X]]) - \frac{1}{24}([Y, [X, [X, Y]]]) + \dots)$, where each term in the sum is a nested Lie bracket.*

This combinatorial formula was introduced and proven by Eugene Dynkin in 1947 [4].

**Corollary 3.3.6.1.** *$exp(X + Y) = exp(X)exp(Y)$, $\forall X, Y \in \mathfrak{g} \Leftrightarrow G_e$ is abelian.*

*Proof.* ($\Rightarrow$)

First, we will show that $\varphi : \mathbb{R} \to G$, given by $\varphi(t) = exp(tX)exp(tY)$ is a one parameter subgroup, i.e. we need to show that $\varphi(t + s) = \varphi(t)\varphi(s)$

$$\varphi(t+s) = exp((t+s)X)exp((t+s)Y) = exp((t+s)(X+Y)) = exp(t(X+Y) + s(X+Y))$$

$$= exp(t(X+Y))exp(s(X+Y)) = exp(tX)exp(tY)exp(sX)exp(sY) = \varphi(t)\varphi(s)$$

Thus $\varphi : \mathbb{R} \to G$ is a Lie group homomorphism, thereby a one parameter subgroup. Note that $\varphi(0) = e \Rightarrow \varphi(\mathbb{R}) \subseteq G$ is a neighborhood containing $e$. Thus $\varphi(\mathbb{R})$ generates $G_e$. If we can show that $\varphi(\mathbb{R})$ is abelian, then $G_e$ is abelian. Note that $exp(X + Y) = exp(X)exp(Y)$ and $exp(X + Y) = exp(Y + X) \Rightarrow exp(X)exp(Y) = exp(Y)exp(X)$, thus $\varphi(\mathbb{R})$ is abelian

($\Leftarrow$)

Let $X, Y \in G_e \Rightarrow [X, Y] = 0$. By Proposition 3.3.6, we have that $exp(X)exp(Y) = exp(X+Y)$ because that entire mess of a nested Lie bracket disappears.

□

We saw, in Proposition 3.3.3, that the exponential map of matrix groups is simply the exponential itself; namely, $exp(A) = e^A$ where $A \in \mathfrak{gl}_\mathfrak{n}(F)$. However, the exponential map is not *defined* to be this, rather this example is simply a special case from the general formula $exp(X) = \gamma(1)$ because $e^A$ is the one-parameter subgroup generated by the Lie algebra element $A$. However, it should be noted that the exponential map need not be the map given by $x \mapsto e^x$; we study $\mathbb{R}^n$ and $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ as examples.

We have $exp : \mathfrak{g} \to G$ such that for any $X \in \mathfrak{g}$, $exp(X) = \gamma(1)$ where $\gamma(t)$ is the integral curve of the (left-invariant) vector field $X \in \mathfrak{g}$.

1. We begin by finding the lie algebra of $\mathbb{R}^n$. $Lie(\mathbb{R}^n)$ is just the tangent space at the multiplicative identity, $(1, \ldots, 1) \Rightarrow Lie(\mathbb{R}^n) = \mathbb{R}^n$. Thus let $X \in \mathbb{R}^n$, we have that $\frac{d}{dt}|_{t=0}\gamma(t) = X \Rightarrow \gamma(t) = tX \Rightarrow exp(X) = \gamma(1) = X$, therefore we have that $exp = \mathbb{1}$
2. Now, for the $n$-torus: we identify $\mathbb{T}^n$ as the direct product $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$. Let $\mathfrak{t} = Lie(\mathbb{T}^n)$, we have that $\mathfrak{t} = T_e(\mathbb{T}^n) = T_e\mathbb{S}^1 \oplus \cdots \oplus T_e\mathbb{S}^1$. Clearly $T_e\mathbb{S}^1 = \mathbb{R}$, so $\mathfrak{t} = \oplus_{i=1}^n \mathbb{R} = \mathbb{R}^n$. Thus we will want $\gamma'(0) = X = (x_1, \ldots, x_n) \in \mathbb{R}^n$. So $\gamma(t) \in \mathbb{T}^n \Rightarrow \gamma(t) = (\alpha e^{ix_1 t}, \ldots, \alpha e^{ix_n t})$ for some unknown constant $\alpha$. Since $\gamma'(0) = X$ we have that $\gamma'(0) = \alpha(ix_1, \ldots, ix_n) = (x_1, \ldots, x_n) \Leftrightarrow \alpha = -i$. Thus we have $\gamma(t) = -i(e^{ix_1 t}, \ldots, e^{ix_n t}) \Rightarrow exp(X) = \gamma(1) = -i(e^{ix_1}, \ldots, e^{ix_n})$.

## 3.4 Infinitesimal Generators of Group Actions

Suppose $G$ acts on a smooth manifold, $M$, from the right ([3] shows that there is an analogous construction for left actions, thus we do not lose generality be considering right actions). We thus have the orbit map $\theta : M \times G \to M$, or $\theta^{(p)} : G \to M$. Consider $X$ and its one-parameter subgroup, $exp(tX)$. We can thus consider the set: $\{q \in M \, : \, q = g \cdot p \text{ where } exp(0) = p \text{ and } g \in exp(tX)\}$. Essentially, we are carving out a a curve in $G$, which is a Lie subgroup of $G$; the set just defined is the orbit of $p \in M$ under the action of this curve. Thus, the set written above is the image of this curve acting on a point $p \in M$. We then define $\hat{X}_p = \frac{d}{dt}\Big|_{t=0} p \cdot exp(tX)$. If you picture this curve $\gamma(t)$ on $G$, then picture the curve in $M$ that is given from the action $p \cdot exp(tX)$. The initial velocity vector of this curve is $\hat{X}_p$.

**Definition.** *The map $\hat{\theta} : \mathfrak{g} \to \mathfrak{X}(M)$ is called the* ***infinitesimal generator of*** *$\theta$. The infinitesimal generator allows us to compare vectors on $G$ and their integral curves to the generators of the flow on $M$, where the flow is an orbit from the action.*

**Proposition 3.4.1.** *The infinitesimal generator of a Lie group action $\theta$ is itself a Lie algebra homomorphism.*

**Definition.** *Any Lie algebra homomorphism $\hat{\theta} : \mathfrak{g} \to \mathfrak{X}(M)$ is called a* ***$\mathfrak{g}$-action on*** *$M$. We say that $\hat{\theta}$ is complete if $\hat{\theta}(X)$ is complete*

**Proposition 3.4.2.** *Let $G$ is a simply connected Lie group and $\mathfrak{g}$ is its Lie algebra; further, suppose $\hat{\theta} : \mathfrak{g} \to \mathfrak{X}(M)$ is a complete $\mathfrak{g}$-action on $M$; then there is a unique smooth right $G$-action on $M$ whose infinitesimal generator is $\hat{\theta}$.*

The two propositions above are proven in [2]. Recall how, in quantum mechanics, people consistently reference "infinitesimal generators of rotations". Indeed, these infinitesimal generators are not themselves rotations; rather, they are elements of a Lie algebra whose one-parameter subgroup gives rotations. We will see an example shortly.

**Proposition 3.4.3.** *The infinitesimal generator of right multiplication in $G$ is given by is trivially given by the inclusion $\mathfrak{g} \hookrightarrow \mathfrak{X}(M)$*

*Proof.* $\theta : G \times G \to G$ is given by $\theta^{(h)}(g) = hg = R_g(h)$

Let $X \in \mathfrak{g}$ and let $\gamma(t) = R_{exp(tX)}$ is the flow of $X$ starting at $p$. Thus $\hat{\theta} : \mathfrak{g} \to \mathfrak{X}(M)$ is given by $\hat{\theta}(X_p) = \frac{d}{dt}\big|_{t=0} p \cdot expt(tX) = \frac{d}{dt}\big|_{t=0} pexp(tX) = \frac{d}{dt}\big|_{t=0} R_{exp(tX)}(p) = \frac{d}{dt}\big|_{t=0}\gamma(t) = \gamma'(0) = X_p \Rightarrow$ $\hat{\theta} : \mathfrak{g} \hookrightarrow \mathfrak{X}(M)$.

□

### 3.4.1 The Adjoint Representation

Recall that a representation of a Lie group is a homomorphism $\phi : G \to GL(V)$. We have something similar for Lie algebras:

**Definition.**

1. *A representation of a Lie algebra, $\mathfrak{g}$, is a Lie algebra homomorphism $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$ where $V$ is some vector space; thus $\mathfrak{gl}(V) = \{A : V \to V \,:\, A \text{ is linear }\}$.*
2. *Just like in the group sense, if $\varphi$ is injective then it is said to be **faithful**.*

In group theory, one studies a group action by $G$ acting on itself by conjugation; namely, $C : G \times G \to G$ by $C(g, h) = ghg^{-1}$ or, equivalently, $C_g : G \to G$ by $C_g(h) = ghg^{-1}$.

**Proposition 3.4.4.** *$C_g : G \to G$ is a Lie group homomorphism*

*Proof.* Clearly it is smooth, since it can be written as the composition of multiplication and inversion, which are smooth since $G$ is a Lie group. We need only check that $C_g(ab) = C_g(a)C_g(b)$. $C_g(ab) = gabg^{-1} = gag^{-1}gbg^{-1} = C_g(a)C_g(b)$.

□

Since $C_g$ is a Lie group homomorphism, we can consider its pushforward as an induced Lie algebra homomorphism:

**Definition.** *The **Adjoint Representation** of a Lie group is the induced Lie algebra homomorphism from the conjugation map: $Ad(g) : \mathfrak{g} \to \mathfrak{g}$ is given by $Ad(g) = (C_g)_*$.*

A proof that this is a representation can be found in [2]; instead of regurgitating this proof, we will give an analogous result:

**Proposition 3.4.5.** *Define a map $ad(X) : \mathfrak{g} \to \mathfrak{g}$ by $ad(X)Y = [X, Y]$. For any $\mathfrak{g}$, the map $ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a Lie algebra representation, which we call the **adjoint representation of** $\mathfrak{g}$*

*Proof.* We need to prove that $ad$ is a homomorphism of Lie algebras; namely that $ad([X, Y]) = [ad(X), ad(Y)]$.

$ad([X, Y])Z = [[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]]$ by the Jacobi identity.

On the other hand, we have that $[ad(X), ad(Y)]Z = (ad(X)ad(Y))Z - (ad(Y)ad(X))Z = ad(X)[Y, Z] - ad(Y)[X, Z] = [X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z] = ad([X, Y])Z$ □

**Proposition 3.4.6.** *ad is the induced Lie algebra homomorphism from Ad (i.e. $(Ad)_* = ad$)*

The proof of 3.4.6 can be found in [2].

**Corollary 3.4.6.1.** *Let $G$ be a connected Lie group and $\mathfrak{g} = Lie(G)$*

1. *For any $X, Y \in \mathfrak{g}$, $[X, Y] = 0 \Leftrightarrow exp(tX)exp(sY) = exp(sY)exp(tX)$, $\forall s, t \in \mathbb{R}$*
2. *$G$ abelian $\Leftrightarrow$ $\mathfrak{g}$ abelian*

*Proof.* 1. $(\Rightarrow)$

By Proposition 3.3.5(6), we have that the following diagram commutes:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "Ad(g)"] \arrow[d, "exp"'] & \mathfrak{g} \arrow[d, "exp"] \\ G \arrow[r, "Cg"'] & G \end{tikzcd}$$

Thus $exp(Ad(exp(tX)sY)) = C_{exp(tX)}(exp(sY)) = exp(tX)exp(sY)exp(-tX)$. By proposition 3.2.5(6), we have that the following diagram commutes:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "ad"] \arrow[d, "exp"'] & \mathfrak{gl}(n\mathfrak{g}) \arrow[d, "exp"] \\ G \arrow[r, "Ad"'] & GL(n, \mathfrak{g}) \end{tikzcd}$$

This diagram tells us that $Ad(exp(tX)) = exp(ad(tX))$, thus $exp(Ad(exp(tX))sY) = exp(exp(ad(tX))sY)$ But

$$ad(X) = [X, \cdot] = 0 \Rightarrow$$

$$ad(tX) = 0 \Rightarrow exp(Ad(exp(tX))sY) = exp(exp(ad(tX))sY) = exp(exp(0)sY) = exp(sY)$$

On the other hand,

$$exp(Ad(exp(tX))sY) = exp(tX)exp(sY)exp(-tX) \Rightarrow exp(tX)exp(sY) = exp(sY)exp(tX)$$

$(\Leftarrow)$

This is clear from the Proposition 3.3.6, we have $exp(tX)exp(tY) = exp(t(X+Y)+t^2 Z_1(t)) = exp(t(Y+X)+t^2 Z_2(t)) = exp(tY)exp(tX) \Leftrightarrow Z_1(t) = Z_2(t) = 0 \Leftrightarrow [X, Y] = 0$

2. $(\Rightarrow)$

This is the result of Proposition 3.2.4.

($\Leftarrow$)

suppose $\mathfrak{g}$ is abelian $\Rightarrow \forall X, Y \in \mathfrak{g}$, $exp(tX)exp(sY) = exp(sY)exp(tX)$. By Proposition 3.3.5(5), $\exists$ open neighborhoods $V$ of $0 \in \mathfrak{g}$ and $W$ of $e \in G$ such that $exp|_V : V \to W$ is a diffeomorphism. Clearly $W$ generates $G$ since $W \subseteq G$ is open and $G$ is connected. Thus elements in the image of $exp$ commute in $G$ (since the generators are in $W$ and those commute). In other words, the image of $exp$ is abelian, so if we can show that its image is a subgroup of $G$ containing $e$, then it contains an open neighborhood of $e$, thus $exp(V)$ is generated by said neighborhood...but so is $G$, therefore $exp(V) = G$

First, we need to show that $exp$ is a Lie group homomorphism, namely $exp(X + Y) = exp(X)exp(Y)$ (we already know that $exp$ is smooth). By Proposition 3.3.6, we know that $exp(tX)exp(tY) = exp(t(X + Y))$ (since $[X, Y] = 0$), in particular $exp(X)exp(Y) = exp(X + Y)$. Thus the image $exp(V)$ is a Lie subgroup of $G$ containing the identity; there is a neighborhood of $e$ that generates $exp(V)$, but this must also generate $G$, thus $exp(V) = G$ is abelian.

□

**Proposition 3.4.7.** *Let $G$ be a connected Lie group, $Lie(G) = \mathfrak{g}$, $Ad : G \to GL(\mathfrak{g})$. Then $ker(Ad) = Z(G) = \{g \in G : \forall h \in G, gh = hg\}$, the **center of** $G$*

*Proof.* Recall that $Ad : G \to GL(\mathfrak{g})$ is given by $Ad(g) = (C_g)_*$. We have $((C_g)_*X)_e = dC_{(g,e)}(0, X_e) = \frac{d}{dt}\Big|_{t=0} C(g, exp(tX)) = \frac{d}{dt}\Big|_{t=0} gexp(tX)g^{-1} = gXg^{-1}$

($\subseteq$)

Suppose $g \in G$, $X \in \mathfrak{g}$ such that $g \in ker(Ad)$. $Ad(g) \in GL(\mathfrak{g}) \Rightarrow Ad(g) \equiv \mathbb{1}$, thus $Ad(g)X = X$. Since $exp$ restricts to a diffeomorphism on a neighborhood $U$ of $e \in G$ and $0 \in \mathfrak{g}$, we have that $Ad(g)X = X$ on $U$, but $U$ generates $G_e = G$ since $G$ is connected. Therefore, we have $Ad(g)X = gXg^{-1} \Rightarrow gX = xG \Rightarrow ker(Ad) \subseteq Z(G)$

($\supseteq$)

Suppose $g \in Z(G) \Rightarrow gh = hg\ \forall h \in G$. In other words $C_g(h) = h \Rightarrow C_g \equiv \mathbb{1} \Rightarrow (C_g)_* \equiv \mathbb{1} \Rightarrow g \in ker(Ad)$.

□

Here we will give an explicit example of the adjoint representation and we will show that it is not faithful. The group of interest is $GL(n, \mathbb{R})$:

$((C_g)_*X)_e = \frac{d}{dt}\Big|_{t=0} g \cdot exp(tX) \cdot g^{-1} = \frac{d}{dt}\Big|_{t=0} ge^{tX}g^{-1} = gXg^{-1}$. It's faithful if it's injective, and it's injective if $ker(Ad) = \{\mathbb{1}\} \Leftrightarrow Z(GL(n, \mathbb{R})) = \{\mathbb{1}\}$. However, any scalar multiple of the identity matrix commutes with any other matrix $\Rightarrow Z(GL(n, \mathbb{R})) \neq \{\mathbb{1}\} \Rightarrow$ not faithful.

**Proposition 3.4.8.** $Z(\mathfrak{g}) = \{X \in \mathfrak{g} : [X, Y] = 0 \forall Y \in \mathfrak{g}\}$. *Suppose $G$ is connected, we then have that $Z(Lie(G)) = Lie(Z(G))$*

*Proof.* Note that $ad(X) = [X, \cdot]$ has already been shown to be a Lie algebra homomorphism (again, $ad = (Ad)_*$). If $Z(\mathfrak{g}) = ker(ad)$ then by Proposition 4.2.6, we have $Z(Lie(G)) = Lie(Z(G))$ so we will now prove that $Z(\mathfrak{g}) = ker(ad)$:

($\subseteq$)
Suppose $X \in Z(\mathfrak{g}) \Rightarrow XY = YX \forall Y \in \mathfrak{g} \Rightarrow [X,Y] = 0 \Rightarrow X \in ker(ad)$
($\supseteq$)
Suppose $X \in ker(ad) \Rightarrow [X,Y] = 0 \forall Y \in \mathfrak{g} \Rightarrow X \in Z(\mathfrak{g})$, thus $ker(ad) = Z(\mathfrak{g})$ □

# 4 Riemannian Differential Geometry

Many texts on Differential Geometry begin with defining a *connection*. I believe that defining a connection without the proper motivation makes connections seem more esoteric than necessary, thus we will begin by discussing the notion of *parallel transport*.

## 4.1 Parallel Transport

Say we have some Riemannian manifold $M$ and a curve $\gamma : I \to M$ where $I$ is an interval in $\mathbb{R}$. Let $t_0, t_1 \in I$ and $V \in \mathfrak{X}(M)$ where $V_0 = V(\gamma(t_0)), V_1 = V(\gamma(t_1))$; we define an operator $P_{t_0 t_1} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$ and require that $P_{t_0 t_1} V_0 = V_1$. In other words, $P$ is an operator that takes a vector in some tangent space of $M$ and transports it to other tangent spaces in a particular way; this particular way is *defined* such that if you are considering a so-called "flat" manifold (such as $\mathbb{R}^n$) and you return to the original tangent space, you have the same vector that you started with. In order to describe such transportation, we need to come up with a proper notion of how to differentiate vector fields along curves, which requires the notion of *covariant differentiation*, which itself requires the notion of a *connection*. Before defining a connection, I will make one more attempt to give a visual description of what exactly we're after:

Say we have a curve $\gamma$ and a (smooth) vector field $V$. In order to describe how $V$ changes along $\gamma$ (thereby describing some notion of differentiating $V$ along $\gamma$), we can describe how $V_p$ changes with respect to $\gamma'_p$ as we change $p$. Rather than the Lie derivative, in which we measure how each vector field changes along the entire flow of another vector field, here we are comparing the relationship (angles, magnitudes) of $V_p$ and $\gamma'_p$ as we move along the curve $\gamma$.

## 4.2 Connections

Let $\pi : E \to M$ be a vector bundle over $M$ and $\Gamma(E)$ be the space of smooth sections of $E$ (if $E = TM$ then a local smooth section $\sigma : M \to TM$ is precisely a local vector field on $M$). We define:

**Definition.** *A* ***linear connection*** *is an operator* $\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$*, written* $\nabla_X Y$ *such that:*

1. *Given* $f, g \in C^\infty(M)$*,* $\nabla_{fX_1 + gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y$
2. *Given* $a, b \in \mathbb{R}$*,* $\nabla_X (aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2$
3. *Given* $f \in C^\infty(M)$*,* $\nabla_X (fY) = f\nabla_X Y + (Xf)Y$

**Definition.** *The* ***Christoffel Symbols of*** $\nabla$ *are functions* $\Gamma^k_{ij}$ *that define the evaluation of a linear connection over a frame in* $TM$*. In other words, if* $\{E_i\}$ *is a local frame in* $TM$ *(i.e. a frame on* $M$*), then we define* $\Gamma^k_{ij}$ *by* $\nabla_{E_i} E_j = \Gamma^k_{ij} E_k$.

We can also "work backwards" in a sense: locally, connections are in one-to-one correspondence with Christoffel symbols, so we could recover a connection by defining these Christoffel symbols. It's true that these Christoffel symbols are local quantities, but connections are uniquely determined by their local data, which we will see shortly.

**Proposition 4.2.1.** $\nabla_X Y\big|_p$ *only depends on* $X$ *and* $Y$ *in an arbitrarily small neighborhood of* $p$. *In fact,* $\nabla_X Y\big|_p$ *depends only on* $Y$ *in an arbitrarily small neighborhood of* $p$ *and* $X$ *at* $p$.

**Proposition 4.2.2.** *Let* $\nabla$ *be a linear connection,* $X, Y \in \mathfrak{X}(U)$ *(where* $U \subseteq M$*), and* $\{E_i\}$ *be a local frame for* $U$ *such that* $X = X^i E_i$ *and* $Y = Y^j E_j$*, then we have* $\nabla_X Y = (XY^k + X^i Y^j \Gamma^k_{ij})E_k$

In fact, we can make an analogous definition for tensor fields:

**Proposition 4.2.3.** *Let* $\nabla$ *be a linear connection on* $M$. *There is a unique connection in each tensor bundle* $T^k_l M$*, also denoted* $\nabla$ *such that:*

1. *On* $TM$*,* $\nabla$ *is the ordinary linear connection*
2. *On* $T^0 M$*,* $\nabla$ *is given by* $\nabla_X f = Xf$
3. $\nabla$ *obeys* $\nabla_X(F \otimes G) = (\nabla_X F) \otimes G + F \otimes (\nabla_X G)$
4. $\nabla_X(trY) = tr(\nabla_X Y)$ *where* $tr$ *is the trace*
5. *If* $\omega$ *is a covector field and* $Y$ *a vector field, then* $\nabla_X\langle\omega, Y\rangle = \langle\nabla_X\omega, Y\rangle + \langle\omega, \nabla_X Y\rangle$
6. *For any* $F \in \mathfrak{T}^k_l(M)$*, vector fields* $Y_i$ *and* 1*-forms* $\omega^j$*, (where* $\mathfrak{T}^k_l(M)$ *denotes the space of smooth sections in* $T^k_k(M)$*), we have:*
   $(\nabla_X F)(\omega^1, \ldots, \omega^l, Y_1, \ldots, Y_k) = X(F(\omega^1, \ldots, \omega^l, Y_1, \ldots, Y_k)) -$
   $\sum_{j=1}^l F(\omega^1, \ldots, \nabla_X\omega^j, \ldots, \omega^l, Y_1, \ldots, Y_k) - \sum_{i=1}^k F(\omega^1, \ldots, \omega^l, Y_1, \ldots, \nabla_X Y_i, \ldots, Y_k)$

The proofs of the above four propositions can be found in [6].

**Corollary 4.2.3.1.** *If* $\nabla$ *is a linear connection,* $\omega$ *a* 1*-form, and* $X$ *a vector field, then we have that* $\nabla_X\omega = (X^i\partial_i\omega_k - X^i\omega_j\Gamma^j_{ik})dx^k$ *is the coordinate expression for* $\nabla_X\omega$

*Proof.* Let $\omega = \omega_k dx^k \Rightarrow \nabla_X(\omega_k dx^k) = X(\omega_k dx^k) - \omega_k\nabla_x(dx^k)$. Let $X = X^i\partial_i$ and we have: $\nabla_X\omega = X^i\partial_i(\omega_k dx^k) - \omega_k X^i \nabla_{\partial_i}(dx^k) = X^i(\partial_i\omega_k)dx^k - \omega_k X^i\nabla_{\partial_i}(\partial_k dx^k) = X^i(\partial_i\omega_k) - \omega_k X^i\Gamma^j_{ik}dx^k = (X^i\partial_i\omega_k - \omega_k X^i\Gamma^j_{ik})dx^k$

□

**Definition.** *Given a smooth section* $F$ *of the tensor bundle* $T^k_l(M)$*, we can define the* ***total covariant derivative of*** $F$***,*** $\nabla F$ *as:*

$$\nabla F(\omega^1, \ldots, \omega^l, Y_1, \ldots, Y_k, X) = \nabla_X F(\omega^1, \ldots, \omega^l, Y_1, \ldots, Y_k)$$

*If* $f \in C^\infty(M)$*, we define the* ***covariant Hessian of*** $f$ *to be the* 2*-tensor* $\nabla^2 f$.

**Proposition 4.2.4.** *Let* $f \in C^\infty(M)$*; we have:*

1. $\nabla f = df$
2. $\nabla^2 f(X, Y) = Y(Xf) - (\nabla_Y X)f$

*Proof.*

1. Consider $\nabla f$; we have $\langle \nabla f, X\rangle = \nabla_X f = Xf = \langle df, X\rangle \Rightarrow df = \nabla f$
2. By Proposition 4.2.4(6), we have $\nabla^2 f(XY) = \nabla_Y(\nabla_X f) = \nabla_Y(Xf) = Y(Xf) - f(\nabla_Y X)$

$\square$

**Definition.** *The following make use of the **torsion tensor**: $\tau(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$, where $[X,Y]$ is the Lie bracket. If $\tau \equiv 0$, we say that $\nabla$ is **torsion free**.*

**Proposition 4.2.5.**

1. *$\tau$ is a $\binom{2}{1}$ tensor field, entitled the torsion tensor of $\nabla$*
2. *$\nabla$ is said to be symmetric if $\nabla$ is torsion free. $\nabla$ is symmetric if and only if its Christoffel symbols are symmetric with respect to any coordinate frame, i.e. $\Gamma^k_{ij} = \Gamma^k_{ji}$*
3. *$\nabla$ is symmetric if and only if $\nabla^2 f$ is a symmetric 2-tensor field (where $f \in C^\infty(M)$*

*Proof.*

1. We need to show that $\tau$ is linear over smooth functions. Let $f \in C^\infty(M)$:
   (a) $\tau(fX,Y) = \nabla_{fX}Y - \nabla_Y(fX) - [fX,Y] = f\nabla_x Y - f\nabla_y X - (Yf)X + (Yf)X + f[Y,X] = f\nabla_X Y - f\nabla_Y X - f[X,Y] = f\tau(X,Y)$
   (b) $\tau(X,fY) = \nabla_X(fY) - \nabla_{fY}X - [X,fY] = XfY - f\nabla_x Y - f\nabla_Y X - XfY - f[X,Y] = f\nabla_X Y - f\nabla_Y X - f[X,Y] = f\tau(X,Y)$
2. $(\Rightarrow)$
   Assume $\nabla$ is symmetric $\Rightarrow \tau = 0 \Rightarrow \nabla_X Y = \nabla_Y X + [X,Y]$. We have that $\nabla_X Y = XY + X^i Y^j \Gamma^k_{ij}\partial_k$; on the other hand, $\nabla_Y X + [X,Y] = YX + Y^i X^j \Gamma^k_{ij}\partial_k = Y^i X^j \Gamma^k_{ij}\partial_k = Y^j X^i \Gamma^k_{ji}\partial_k \Leftrightarrow \Gamma^k_{ij} = \Gamma^k_{ji}$
   $(\Leftarrow)$
   Suppose $\Gamma^k_{ij} = \Gamma^k_{ji} \Leftrightarrow X^i Y^j \Gamma^k_{ij}\partial_k = Y^j X^i \Gamma^k_{ji}\partial_k = Y^i X^j \Gamma^k_{ij}\partial_k + XY - YX - XY + YX \Rightarrow XY + X^i Y^j \Gamma^k_{ij}\partial_k = \nabla_X Y = YX + Y^i X^j \Gamma^k_{ij}\partial_k + XY - YX = YX + Y^i X^j \Gamma^k_{ij}\partial_k + [X,Y] = \nabla_Y X + [X,Y] \Rightarrow \tau = 0$
3. $\nabla$ is symmetric $\Leftrightarrow \nabla_X Y = \nabla_Y X + [X,Y] \Leftrightarrow \nabla_X Yu = \nabla_Y Xu + [X,Y]u = \nabla_Y Xu + XYu - YXu \Leftrightarrow \nabla_Y Xu = \nabla_X Yu + YXu - XYu$. Thus, we have $-YXu + \nabla_Y Xu = -XYu + \nabla_X Yu \Leftrightarrow YXu - \nabla_Y Xu = XYu - \nabla_X Yu \Leftrightarrow \nabla^2 u(XY) = \nabla^2 u(Y,X)$

$\square$

**Proposition 4.2.6.** *Let $\{E_i\}$ be a local frame with dual coframe $\{\varphi^i\}$*

1. *There is a unique matrix of (real valued) 1-forms $\omega^j_i$ such that $\nabla_X E_i = \omega^j_i(X)E_j$*
2. *$d\varphi^j = \varphi^i \wedge \omega^j_i + \tau^j$ where $\{\tau^1, \ldots, \tau^n\}$ are the torsion 1-forms defined by $\tau(X,Y) = \tau^j(X,Y)E_j$; this is known as Cartan's First Structure Equation.*

*Proof.*

1. Let $X = X^j E_j \Rightarrow \nabla_X E_i = X^j \nabla_{E_j} E_i = X^k \Gamma^j_{ki} \Rightarrow$ we define $\omega^j_i = \Gamma^j_{ki}\varphi^k$. Quite obviously, $\omega^j_i$ is uniquely determined by the connection. We see that defining a connection amounts to defining a matrix (with no restrictions on its properties). We will reference this fact when discussing connections on principal bundles.

2. First, let's deduce these $\tau^j(X,Y)$ terms. We have $\tau(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$; let $X = X^i E_i$ and $Y = Y^j E_j$ to arrive at the fact that $\nabla_X Y = \nabla_{X^i E_i} Y^j E_j = X^i(E_i(Y^j E_j) + Y^j \nabla_{E_i} E_j)$ and $\nabla_Y X = Y^j(E_j(X^i E_i) + X^i \nabla_{E_j} E_i)$, we also recall that $[X,Y]$ expands as $[X,Y] = X^i E_i(Y^j E_j) - Y^j E_j(X^i E_i)$.

   Thus $\tau(X,Y) = X^i E_i(Y^j E_j) + X^i Y^j \nabla_{E_i} E_j - Y^j E_j(X^i E_i) - Y^j X^i \nabla_{E_j} E_i - X^i E_i(Y^j E_j) + Y^j E_j(X^i E_i) = (X^i Y^j \Gamma^k_{ij} - Y^j X^i \Gamma^k_{ji}) E_k = (X^j Y^i \Gamma^k_{ji} - Y^j X^i \Gamma^k_{ji}) E_k$. Therefore we have that $\tau(X,Y)^k = (X^j Y^i \Gamma^k_{ji} - Y^j X^i \Gamma^k_{ji})$.

   Now, let's evaluate the right hand side of Cartan's first structure equation: $\varphi^i \wedge \omega^k_i(X,Y) = \varphi^i(X)\omega^k_i(Y) - \varphi^i(Y)\omega^k_i(X) = X^i \Gamma^k_{ji} Y^j - Y^i \Gamma^k_{ji} X^j \Rightarrow \varphi^i \wedge \omega^j_i + \tau^i = X^i \Gamma^k_{ji} Y^j - Y^i \Gamma^k_{ji} X^j + X^i Y^j \Gamma^k_{ij} - Y^j X^i \Gamma^k_{ji}$.

   We will now need the coordinate independent evaluation of the exterior derivative of a one form; namely, if $\varphi$ is a one form, then we have $d\varphi(X,Y) = X(\varphi(Y)) - Y(\varphi(X)) - \varphi([X,Y]) = X\Gamma^k_{ji} Y^j - Y \Gamma^k_{ji} X^j - \Gamma^k_{ji}(XY)^j + \Gamma_{ji}(YX)^j$. We have that $(XY)^j = X^i E_i Y^j E_j \Rightarrow \Gamma^k_{ji}(XY)^j = \Gamma^k_{ji} X^i Y^j$ and $\Gamma^k_{ji}(YX)^j = \Gamma^k_{ji} X^j Y^i$. Therefore, we have $d\varphi^k(X,Y) = X^i \Gamma^k_{ji} Y^j - Y^i \Gamma^k_{ji} X^j - \Gamma^k_{ji} X^i Y^j + \Gamma^k_{ji} X^j Y^i = \varphi^i \wedge \omega^j_i + \tau^i$

□

The use of connection 1-forms will provide a much more convenient framework for describing the differential geometry of principal bundles and is, thus, how we will define connections and curvature on principal bundles.

## 4.3 Covariant Derivatives and Geodesics

With the connection defined, we can now get closer to our goal of rigorously defining parallel transport. The next step is to describe covariant differentiation:

**Proposition 4.3.1.** *Let $\nabla$ be a linear connection on $M$ and $\gamma : I \to M$ a curve. $\nabla$ determines a unique operator $D_t : \Gamma(T\gamma) \to \Gamma(T\gamma)$ (where $\Gamma(T\gamma)$ is a shorthand I'm using for a section over a curve $\gamma \subset M$) that satisfies:*

1. *If $a, b \in \mathbb{R}$, then $D_t(aV + bW) = aD_tV + bD_tW$*
2. *If $f \in C^\infty(M)$, then $D_t(fV) = \dot{f}V + fD_tV$*
3. *If $V$ is extendible, then for any extension $\tilde{V}$ of $V$, we have $D_t(V(t)) = \nabla_{\dot{\gamma}(t)}\tilde{V}$*
4. *In a local coordinate frame, we have that $D_tV(t_0) = \dot{V}^j(t_0)\partial_j + V^j(t_0)\nabla_{\dot{\gamma}(t_0)}\partial_j = (\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0)))\partial_k$*

**Definition.** *A **geodesic** is a curve $\gamma$ such that $D_t\dot{\gamma}(t) = 0$, meaning that the curve $\gamma$ is not accelerating on the manifold. If $\gamma'(0) = V$ and $\gamma(0) = p$ we say that $\gamma$ is the geodesic of $V$ starting at $p$.*

**Proposition 4.3.2.** *If $M$ is a manifold equipped with a linear connection, then for any $p \in M$, $V \in T_pM$, and $t_0 \in \mathbb{R}$, there is a geodesic $\gamma : I \to M$ of $V$ starting at $p$.*

The proofs of the above two propositions can be found in [6].

## 4.4 Parallel Transport Made Rigorous

We can now rigorously define what was originally meant by the notion of parallel transport.

**Definition.** *A vector field $V$ along a curve $\gamma$ is said to be* ***parallel along*** *$\gamma$ if $D_tV = 0$; namely, "comparing" the vector of $\gamma'_p$, and $V_p$ looks the same at any $p$. A vector field $V$ is parallel if and only if $\nabla V = 0$; in this sense, $V$ is parallel on $M$ if it is parallel along any curve in $M$.*

**Proposition 4.4.1.** *Given a curve $\gamma$ on $M$ and a vector $V_0 \in T_{\gamma(t_0)}M$, there exists a unique parallel vector field $V$ along $\gamma$ such that $V(t_0) = V_0$*

A proof of this can be found in [6]. In fact, we can recover our notion of covariant differentiation along a curve $\gamma$ from the notion of parallel transport. Recall from above that parallel translation defines the operator $P_{t_0t_1} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$ by defining $P_{t_0t_1}V_0 = V(t_1)$.

**Proposition 4.4.2.** *If $\nabla$ is a linear connection on $M$, we have that*

$$D_tV(t_0) = \lim_{t\to t_0} \frac{P^{-1}_{t_0t}V(t) - V(t_0)}{t - t_0}$$

*Proof.* Let $\{E^i\}$ be a parallel frame along $\gamma$, namely a frame such that $\nabla_{\dot\gamma(t)}E_i = 0$. Thus $D_tV(t_0) = \dot V^j(t_0)E_j + V^j(t_0)\nabla_{\dot\gamma(t_0)}E_j$; but $\{E^j\}$ is parallel indeed yields that $\nabla_{\dot\gamma(t_0)}E_j = 0 \Rightarrow D_tV = \dot V^j(t_0)E_j$. Thus $P^{-1}_{t_0t}(V(t)) = V(t_1) \Rightarrow \lim_{t\to t_0} \frac{P^{-1}_{t_0t}V(t) - V(t_0)}{t - t_0} = \lim_{t\to t_0} \frac{V(t_1) - V(t_0)}{t_1 - t_0} = \frac{d}{dt}\big|_{t_0} V = \dot V^j(t_0)E_j$.

Therefore, we have $D_tV(t_0) = \lim_{t\to t_0} \frac{P^{-1}_{t_0t}V(t) - V(t_0)}{t - t_0}$

□

### The Euclidean Connection: An Example

An example of a connection is the Euclidean connection, which is defined to be $\bar\nabla_X(Y^j\partial_j) = (XY^j)\partial_j$. In this example, we will compute the Christoffel symbols and geodesics of $\bar\nabla$:

1. Christoffel Symbols:

   Let $E_i = \partial_i$ be a local frame for $TM$. We have that $\bar\nabla_{\partial_i}E_j = \Gamma^k_{ij}$. Indeed, $\bar\nabla_{\partial_i}\partial_j = (\partial_i\partial_j)\partial_k = 0$. Trivially, the Christoffel symbols are symmetric in $i$ and $j$, so $\bar\nabla$ is torsion free.

2. Geodesics:

   $D_t\dot\gamma = (\ddot\gamma^k + \dot\gamma^j\dot\gamma^i\Gamma^k_{ij})\partial_k = \ddot\gamma^k\partial_k = 0 \Rightarrow \dot\gamma = c \in \mathbb{R} \Rightarrow \gamma(t) = ct + b$; thus the geodesics of $\bar\nabla$ are the straight lines in $\mathbb{R}^n$.

## 4.5 The Levi-Civita Connection

**Definition.** *We say that a linear connection, $\nabla$, is* ***compatible with the Riemannian metric*** *$g$ if, for all vector fields, $X, Y, Z$, we have $\nabla_X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$*

**Proposition 4.5.1.** *If $\nabla$ is a linear connetion, then the following are equivalent on a Riemannian manifold:*

1. *$\nabla$ is compatible with $g$*
2. *$\nabla g \equiv 0$ (where this $\nabla g$ denotes the total covariant derivative of $g$)*
3. *If $V, W$ are vector fields along a curve $\gamma$, then $\frac{d}{dt}\langle V, W\rangle = \langle D_t V, W\rangle + \langle V, D_t W\rangle$*
4. *If $V, W$ are parallel vector fields along a curve $\gamma$, then $\langle V, W\rangle$ is constant*

*Proof.* First, we will show $1 \Leftrightarrow 2$:

$(\Rightarrow)$

Assume $\nabla$ is compatible with $g \Rightarrow \nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$. $\nabla g(Y, Z, X) = (\nabla_X g)(Y, Z) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z)$. We have that $g(Y, Z) \in C^\infty(M) \Rightarrow \nabla_X g(Y, Z) = X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \Rightarrow \nabla g \equiv 0$

$(\Leftarrow)$

$\nabla g \equiv 0 \Rightarrow \forall X, Y, Z \in \mathfrak{X}(M)$, we have that $\nabla g(Y, Z, X) = (\nabla_X g)(Y, Z) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0$. $X(g(Y, Z)) = \nabla_X(g(Y, Z)) \Rightarrow \nabla_X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \Rightarrow \nabla$ is compatible with $g$.

Next, we will show $1 \Leftrightarrow 3$:

$(\Rightarrow)$

$\nabla_X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$. Let $\gamma = t \Rightarrow \dot{\gamma} = \frac{d}{dt}$ and consider $X = \dot{\gamma}$. We then have that $\frac{d}{dt}\langle Y, Z\rangle = \langle \nabla_{\frac{d}{dt}} Y, Z\rangle + \langle Y, \nabla_{\frac{d}{dt}} Z\rangle = \langle \nabla_{\dot{\gamma}} Y, Z\rangle + \langle Y, \nabla_{\dot{\gamma}} Z\rangle = \langle D_t Y, Z\rangle + \langle Y, D_t Z\rangle$

$(\Leftarrow)$

Let $X = \frac{d}{dt}$. We have that $\frac{d}{dt}\langle Y, Z\rangle = X\langle Y, Z\rangle = \nabla_X\langle Y, Z\rangle$. $\langle D_t Y, Z\rangle + \langle Y, D_t Z\rangle = \langle \nabla_{\dot{\gamma}} Y, Z\rangle + \langle Y, \nabla_{\dot{\gamma}} Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$

Next, we will show $3 \Leftrightarrow 4$:

$(\Rightarrow)$

$\frac{d}{dt}\langle V, W\rangle = \langle D_t V, W\rangle + \langle V, D_t W\rangle$; if $V$ and $W$ are parallel, then $D_t V = D_t W = 0 \Rightarrow \frac{d}{dt}\langle V, W\rangle = 0 \Rightarrow \langle V, W\rangle = c \in \mathbb{R}$

$(\Leftarrow)$

$\frac{d}{dt}\langle V, W\rangle = 0 = \langle D_t V, W\rangle + \langle V, D_t W\rangle$, since $V$ and $W$ are parallel, each term on the right is zero so the equality holds.

$\square$

**Definition.** *We say that a linear connection $\nabla$ is* ***symmetric*** *if its torsion tensor is zero; in other words, if $\nabla_X Y - \nabla_Y X \equiv [X, Y]$*

**Proposition 4.5.2.** *On a Riemannian manifold $(M, g)$, there is a unique linear connection, entitled the* ***Levi-Civita Connection****, $\nabla$ that is compatible with $g$ and symmetric.*

The above proposition is sometimes referred to as *the Fundamental Lemma of Riemannian Geometry*. In a proof of this proposition, such as in [6] (on page 69), one uncovers the following paramount formula:

**Corollary 4.5.2.1.** $2\langle \nabla_X Y, Z\rangle = X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle Y, [X, Z]\rangle - \langle Z, [Y, X]\rangle + \langle X, [Z, Y]\rangle$

The two results above are proven in [6]. Consequently, the Levi-Civita Connection is given by the Christoffel symbols $\Gamma^m_{ij} = \frac{1}{2} g^{ml}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij})$ where $g^{ml}$ is the inverse of $g_{ml}$.

## 4.6 Curvature

Here, I do not go into great detail about the construction and identities of different types of curvature, rather I give their definitions and some basic results:

**Definition.** *The* ***Riemann Curvature Endomorphism*** *is defined as* $R : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ *by* $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$. *In coordinates,* $R$ *has the form* $R = R^l_{ijk} dx^i \otimes dx^j \otimes dx^k \otimes \partial_l$

We then attain the **Riemann curvature tensor** by lowering the last index of $R$: $Rm = R_{ijkl} dx^i \otimes dx^j \otimes dx^k \otimes dx^l$ where $R_{ijkl} = g_{lm} R^m_{ijk}$; it is evaluated as $Rm(X, Y, Z, W) = \langle R(X, Y)Z, W\rangle$

**Definition.** *Let* $\nabla$ *be the Levi-Civita Connection and* $\omega^j_i$ *be the connection* 1*-forms already introduced (with respect to a local frame* $\{E_i\}$*). We define the* ***curvature*** 2***-forms*** *as* $\Omega^j_i = d\omega + \omega \wedge \omega$

**Proposition 4.6.1.** $\Omega^j_i = \frac{1}{2} R^j_{kli} \varphi^k \wedge \varphi^l$; *this is known as Cartan's Second Structure Equation.*

*Proof.* Recall $\omega^j_i = \Gamma^j_{ki} \varphi^k$, locally. We then have, where we recall that we can relabel all the dummy indices: $\Omega^j_i = d(\Gamma^j_{ki} \varphi^k) + (\Gamma^j_{\ell m} \varphi^\ell) \wedge (\Gamma^\ell_{ki} \varphi^k) = \varphi^k \wedge \varphi^\ell (\partial_\ell \Gamma^j_{ki} + \Gamma^j_{\ell m} \Gamma^\ell_{ki}) = \frac{1}{2} R^j_{k\ell i} \varphi^k \wedge \varphi^\ell$ ☐

## 4.7 Special Case: Lie Groups

**Proposition 4.7.1.** *If* $G$ *is a Lie group and* $\mathfrak{g}$ *its Lie algebra, we say that a Riemannian metric* $g$ *on* $G$ *is left-invariant if* $L_p^* g = g$, $\forall p \in G$ *(there is an analogous definition for right-invariant, and, thus, bi-invariant). We have that* $g$ *is left-invariant if and only if* $g_{ij} := g(X_i, X_j)$ *with respect to any left-invariant fram* $\{X_i\}$ *are constants*

*Proof.* ($\Rightarrow$) Suppose $g$ is left-invariant, we then have that $g_{ij}\big|_q = L_p^* g_{ij}\big|_q = g_{ij}\big|_{q \cdot p}$. The point of the proof is that a group acts transitively on itself by left multiplication: let $p = q^{-1}\alpha$ for any arbitrary $\alpha \in G$, we then have that $g_{ij}\big|_q = g_{ij}\big|_\alpha$ for any $\alpha \in G$. The converse is the same argument, but in reverse. ☐

**Corollary 4.7.1.1.** *The map* $g \mapsto g\big|_{T_e G}$ *gives a bijection between left-invariant metrics on* $G$ *and inner products on* $\mathfrak{g}$

*Proof.* This statement is analogous to saying that left-invariant vector fields on a Lie group $G$ are given via a bijective correspondence with vectors in $T_e G$. If $g$ is a left invariant metric, then it is uniquely determined by its value, as an inner product, at $e \in G$. ☐

**Proposition 4.7.2.** *Suppose $G$ is a compact, connected Lie group with a left-invariant metric $g$, and suppose $dV$ is the Riemannian volume element off $g$. $dV$ is bi-invariant*

*Proof.* First, we will show that $R_p^* dV$ is left-invariant. Let $\{X_i\}$ be a left invariant frame taken from $\mathfrak{g}$. We then have that $L_p^* R_p^* dV(X_1, \dots, X_n) = R_p^* dV(L_p X_1, \dots, L_p X_n)$. Therefore, we have that $R_p^* dV = \varphi(p) dV$ for some $\varphi(p) \in \mathbb{R}^+$. If $\varphi : G \to \mathbb{R}^+$ is a Lie group homomorphism, then its image is a compact subgroup of $\mathbb{R}^+$:

$\varphi(pq) = R_p^* R_q^* dV = R_p^* \varphi(q) dV = \varphi(p)\varphi(q) dV$. Thus $\varphi(G) \subseteq \mathbb{R}^+$ is a compact subgroup, which means that it is closed and bounded. Suppose $\exists \alpha \in \varphi(G)$, we then have that $\alpha^n \in \varphi(G)$ for any $n \in \mathbb{Z}$; therefore, $\alpha \neq 1$ contradicts the fact that $\varphi(G)$ is compact because it is not bounded. We thus have that $\alpha = 1 \Rightarrow R_p^* dV = dV$. Thus $dV$ is right-invariant, and $L_p^* R_p^* dV = L_p^* dV = dV$ gives us that $dV$ is left invariant. □

**Proposition 4.7.3.** *If $G$ is a Lie group and $p \in G$, recall the inner automorphism $C_p(q) = pqp^{-1}$ and the adjoint representation $Ad : G \times \mathfrak{g} \to \mathfrak{g}$ by $Ad_p \coloneqq (C_p)_*$ An inner product on $\mathfrak{g}$ induces a bi-invariant metric on $G$ if and only if it is invariant under the adjoint representation*

*Proof.* Suppose $\langle \cdot, \cdot \rangle$ is a bi-invariant metric. We have that $C_p = L_p R_{p^{-1}} \Rightarrow (C_p)^* = (L_p)^* (R_{p^{-1}})^*$. We thus have that $L_p^* \langle X, Y \rangle = \langle X, Y \rangle$ and $R_p^* \langle X, Y \rangle = \langle X, Y \rangle \Rightarrow (L_p)^* (R_{p^{-1}})^* \langle X, Y \rangle = \langle X, Y \rangle \Leftrightarrow Ad_p \langle X, Y \rangle = \langle X, Y \rangle$

□

**Definition.** *Suppose $G$ is a Lie group with Lie algebra $\mathfrak{g}$ with basis $(X_1, \dots, X_n)$. We define the* ***structure constants*** *$c_{ij}^k$ by $[X_i, X_j] = \sum_k c_{ij}^k X_k$.*

**Proposition 4.7.4.** *The Christoffel symbols of the Levi-Civita Connection on $G$ is given by $\Gamma_{ij}^m = \frac{1}{2} g^{mk} \sum_l (c_{kj}^l g_{il} - c_{ik}^l g_{jl} - c_{ji}^l g_{kl})$*

*Proof.* From Corollary 5.5.2.1, we begin with $\langle \nabla_X Y, Z \rangle = \frac{1}{2}(X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle Y, [X, Z]\rangle - \langle Z, [Y, X]\rangle + \langle X, [Z, Y]\rangle)$. Assuming that $X, Y, Z \in \mathfrak{g}$, we have that $\langle X, Y \rangle$ is constant, so $Z\langle X, Y\rangle = X\langle Y, Z\rangle = Y\langle Z, X\rangle = 0 \Rightarrow 2\langle \nabla_X Y, Z\rangle = \langle X, [Z, Y]\rangle - \langle Y, [X, Z]\rangle - \langle Z, [Y, X]\rangle$

Let $X = X_i, Y = X_j, Z = X_k$, we then have that $2\langle \nabla_{X_i} X_j, X_k \rangle = \langle X_i, [X_k, X_j]\rangle - \langle X_j, [X_i, X_k]\rangle - \langle X_k, [X_j, X_i]\rangle = \langle X_i, \sum_l c_{kj}^l X_l\rangle - \langle X_j, \sum_l c_{ik}^l X_l\rangle - \langle X_k, \sum_l c_{ji}^l X_l\rangle = \sum_l (c_{kj}^l \langle X_i, X_l\rangle - c_{ik}^l \langle X_j, X_l\rangle - c_{ji}^l \langle X_k, X_l\rangle = \sum_l (c_{kj}^l g_{il} - c_{ik}^l g_{jl} - c_{ji}^l g_{kl}) = 2\langle \nabla_{X_i} X_j, X_k\rangle = 2\langle \Gamma_{ij}^m X_m, X_k\rangle = 2\Gamma_{ij}^m g_{mk} \Rightarrow$
$\Gamma_{ij}^m = \frac{1}{2} g^{mk} \sum_l (c_{kj}^l g_{il} - c_{ik}^l g_{jl} - c_{ji}^l g_{kl})$

□

**Proposition 4.7.5.** *Let $g$ be a bi-invariant metric on $G$.*

1. *For any $X, Y, Z \in \mathfrak{g}$, we have that $\langle [X, Y], Z \rangle = -\langle Y, [X, Z]\rangle$*
2. *For any left invariant vector fields $X, Y$, we have $\nabla_X Y = \frac{1}{2}[X, Y]$*
3. *The geodesics of $g$ starting at $e$ are the one-parameter subgroups.*

*Proof.*

1. Let $\gamma(t) = exp(tX)$. $Ad_{\gamma(t)} = (R_{\gamma(t)})_*(L_{\gamma(t)})_* = (exp(-tX))_*(L_{\gamma(t)})_*$.

   Thus $\frac{d}{dt}\Big|_{t=0} \langle Ad_{\gamma(t)}Y, Ad_{\gamma(t)}Z\rangle = \langle \frac{d}{dt}\Big|_{t=0} Ad_{\gamma(t)}Y, Ad_{\gamma(t)}Z\rangle + \langle Ad_{\gamma(t)}Y, \frac{d}{dt}\Big|_{t=0} Ad_{\gamma(t)}Z\rangle$

   We have that $\frac{d}{dt}\Big|_{t=0} Ad_{\gamma(t)}Y = \frac{d}{dt}\Big|_{t=0}(Ad(exp(tX))Y) = (Ad_*)(exp(tX)Y) = [X, Y]$. Thus $\frac{d}{dt}\Big|_{t=0} \langle Ad_{\gamma(t)}Y, Ad_{\gamma(t)}Z\rangle = \langle [X, Y], Z\rangle + \langle Y, [X, Z]\rangle = 0$ where the equality to 0 follows from the fact that the derivative of bi-invariant metrics is 0. Alas, we have $\langle [X, Y], Z\rangle = -\langle Y, [X, Z]\rangle$

2. $\langle \nabla_X Y, Z\rangle = \frac{1}{2}(\langle X, [Z, Y]\rangle - \langle Y, [X, Z]\rangle - \langle, Z, [Y, X]\rangle) = \frac{1}{2}(\langle [X, Y], Z\rangle + \langle [X, Y], Z\rangle + \langle X, [Z, Y]\rangle)$.
   $\langle X, [Z, Y]\rangle = -\langle X, [Y, Z]\rangle = \langle [Y, X], Z\rangle = -\langle [X, Y], Z\rangle \Rightarrow \nabla_X Y = \frac{1}{2}[X, Y]$

3. $D_t(\dot\gamma) = D_t(\frac{d}{dt}\Big|_{t=0} exp(tX)) = \nabla_{\dot\gamma}(\dot\gamma) = \frac{1}{2}[\dot\gamma, \dot\gamma] = 0$

□

**Proposition 4.7.6.** *If $G$ is a Lie group and $g$ is a bi-invariant metric, then $R(X, Y)Z = \frac{1}{4}[Z, [X, Y]]$ where $X, Y, Z \in \mathfrak{g}$*

*Proof.* $R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z = \frac{1}{2}(\nabla_X[Y, Z] - \nabla_Y[X, Z] - [[X, Y], Z]) = \frac{1}{2}([X, [Y, Z]] - [Y, [X, Z]] - 2[[X, Y], Z])$.

The Jacobi identity gives us that $[Y, [X, Z]] = [X, [Y, Z]] + [Z, [X, Y]] \Rightarrow R(X, Y)Z = \frac{1}{4}([X, [Y, Z]] - [X, [Y, Z]] - [Z, [X, Y]] - 2[[X, Y], Z]) = \frac{1}{4}(-[Z, [X, Y]] + 2[Z, [X, Y]]) = \frac{1}{4}[Z, [X, Y]]$ □

# 5 Differential Geometry of Principal Bundles

In this section we will define prerequisites necessary to fully describe connections on principal bundles, and their relations to connections and curvature in the Riemannian sense. We will begin by defining an associated vector bundle (and give an example that shows its relevance to Riemannian geometry) and then we will proceed to define the, quite abstract, notion of a connection 1-form on a principal bundle $P$.

## 5.1 Associated Vector bundles

Say we have a principal $G$-bundle called $P$ and a vector space $V$. Furthermore, say we have a representation $\rho : G \to GL(V)$. We can define a relation on $P \times V$ in the following way:

$(p, v) \sim (q, w)$ if there is a $g \in G$ such that $(q, w) = (pg^{-1}, \rho(g)v)$. We check that it is an equivalence relation:

1. Reflexive: $(p, v) \sim (p, v)$ if we choose $g = e$
2. Symmetric: $(p, v) \sim (q, w)$ means there is a $g \in G$ such that $(q, w) = (pg^{-1}, \rho(g)w) \Rightarrow qg = p$ and $\rho(g)^{-1}v = w$ (note $\rho(g)^{-1}$ is not $\rho^{-1}(g)$ and $\rho(g) \in GL(V)$ gives us that $\rho(g)^{-1}$ exists). Thus $(p, v) = (qg, \rho(g)^{-1}w)$ so $(q, w) \sim (p, v)$

3. Transitive: $(p,v) \sim (q,w) \sim (x,y)$ gives that $(q,w) = (pg^{-1}, \rho(g)v)$ and $(x,y) = (qh^{-1}, \rho(h)w)$ so composing these maps gives us $(p,v) \sim (x,y)$

In the section on cocycle constraints, we proved that the space formed by taking the direct product of a smooth manifold and a vector space, then quotient out by an equivalence class, is a vector bundle, which we called the **associated vector bundle**.

### 5.1.1 The Tangent Bundle

We have already seen that $F(TM)$, the set of all frames on a manifold $M$ (whose tangent space is $V$), is a principal $GL(n,\mathbb{R})$-bundle. We will now denote $F(TM)$ as $F(TM)$.

Let $\rho : GL(k,\mathbb{R}) \to GL(V)$ be the defining representation and consider $F(TM) \times V/\sim$, where $(p,v) \sim (pg^{-1}, \rho(g)v)$; we will write $F(TM) \times V/\sim$ as $F(TM) \times_\rho V$. We then have $(p,v) \in F(TM) \times V$ gets sent to $[p,v] \in F(TM) \times_\rho V$.

**Proposition 5.1.1.** *$F(TM) \times_\rho \mathbb{R}^n \cong TM$ (where $dim(M) = n$)*

*Proof.* First, consider $[p\big|_x, v] \in F(TM)\big|_x \times_\rho \mathbb{R}^n$. We have that $p\big|_x = \{v_1\big|_x, \dots, v_n\big|_x\} \in T_xM$, a frame at $x$ (i.e. a basis); moreover, we have $v = (q_1, \dots, q_n) \in \mathbb{R}^n$.

Define a homomorphism $\varphi : F(TM)\big|_x \times_\rho \mathbb{R}^n \to T_xM$ by
$\varphi([p\big|_x, v]) = q_1 v_1\big|_x + \cdots + q_n v_n\big|_x \in T_xM$. The bijectivity of $\varphi$ follows clearly from the fact that $\{v_1\big|_x, \dots, v_n\big|_x\}$ is linearly independent and forms a basis for $T_xM$.

We therefore have that $F(TM)\big|_x \times_\rho \mathbb{R}^n \cong T_xM \Rightarrow \coprod_{x\in M} F(TM)\big|_x \times_\rho \mathbb{R}^n \cong \coprod_{x\in M} T_xM \Rightarrow F(TM) \times_\rho \mathbb{R}^n \cong TM$ $\square$

## 5.2 Pullback Vector bundles

We will not go terribly in depth into this topic, rather we will give enough definitions to make sense of the appearance of Pullback bundles in the upcoming section.

Suppose we have a vector bundle $E \xrightarrow{\pi_E} M$ and a continuous, surjective map $\pi : P \to M$ (of course, if $P$ is a principal $G$-bundle over $M$, then the projection map is satisfactory, as the employed notation suggests).

**Definition.** *The **pullback bundle** $\pi^*E$ is defined as $\pi^*E = \{(p,q) \in P \times E : \pi(p) = \pi_M(q)\}$, in other words the pullback bundle consists of pairs of points that project to the same place on $M$.*

**Proposition 5.2.1.** *$\pi^*E$ is a vector bundle over $P$ with projection $\pi' : \pi^*TM \to P$ by $\pi'(p,q) = p$. We also have $\pi'' : \pi^*E \to E$ by $\pi''(p,q) = q$, which gives the commutative diagram:*

$$\begin{tikzcd} \pi^*E \arrow[r, "\pi''"] \arrow[d, "\pi'"'] & E \arrow[d, "\pi_M"] \\ P \arrow[r, "\pi"'] & M \end{tikzcd}$$

The above proposition is proven in [3]. This concept will be employed when $E = TM$, in which case we then have that the pullback bundle $\pi^*TM \subset P \times TM$.

## 5.3 The Connection 1-form

Recall that in the Riemannian sense, we had a connection as a map $\nabla : \Gamma(TM)\times\Gamma(TM) \to \Gamma(TM)$. By the introduction of the connection 1-form, $\omega \in \Omega^1(M;TM)$, we are able to reformulate a connection as a map $D : \Gamma(TM) \to \Gamma(TM \otimes \Omega^1(M;TM))$, such that $D(v) = \omega_i^j(v)E_k$ where $\{E_k\}$ forms a local frame for $TM$. This is the starting point for defining connections on principal bundles; we will phrase the entire theory in terms of 1-forms, following closely the introduction to the topic given by Taubes in [3], and so we begin:

Let $P \xrightarrow{\pi} M$ be a principal $G$-bundle with local triviliazations $\varphi_i$ corresponding to an open cover $\{U_i\}$ of $M$. We have the following sequence:

$$0 \hookrightarrow ker(\pi_*) \hookrightarrow TP \to \pi^*TM \to 0$$

It is most illuminating to explicitly note that $\pi^*TM \subset P \times TM$ and rewrite the sequence as:

$$0 \hookrightarrow ker(\pi_*) \hookrightarrow TP \to P \times TM \to 0$$

The map $TP \mapsto P \times TM$ is given by $(v\big|_p) \in TP \mapsto (\pi_P(p), \pi_*(v))$, where we are thinking of $TP \xrightarrow{\pi_P} P$ as a vector bundle.

The set $ker(\pi_*)$ denotes the set of tangent vectors in $TP$ that are sent to the zero section in $TM$. In this sense, these vectors have no component that is tangent to $TM$, as such they are often said to be "tangent to the fibers of $TM$". More often, they are referred to as:

**Definition.** *The* ***vertical subbundle of*** $TP$ *is* $ker(\pi_*)$. *This is visually intuitive if we think of fibers in* $TP$ *over* $TM$ *as vertical. Given* $T_qP$, *we can write* $T_qP = V_qP \oplus H_qP$ *where* $V_qP$ *is the vertical space on* $P$ *and* $H_qP$ *is a subspace complement.* $HP = \sqcup_{q\in P} H_qP$ *is called the* ***horizontal subbundle of*** $TP$.

We note three significant characteristics about our sequence:

1. The sequence is exact, i.e. the image of each map is the kernel of the following map:
   (a) $0 \hookrightarrow ker(\pi_*) \hookrightarrow TP$. $0 \in TP$ corresponds to the 0 section in $TP$, which is, of course, the 0 section in $ker(\pi_*)$ since $ker(\pi_*) \subset TP$
   (b) $ker(\pi_*) \hookrightarrow TP \to P \times TM$. The image of the inclusion of $ker(\pi_*)$ into $TP$ is, of course, $ker(\pi_*)$, which is sent to the zero section in $TM$. Being vertical, all the vectors in $ker(\pi_*)$ are sent to 0 in $P$
   (c) $TP \to P \times TM \to 0$ is trivial
2. The action of $G$ on $P$ lifts to an action of $G$ on each of the bundles in the sequence:
   (a) First, we define our action of $G$ on $P$ by $\theta : G \times P \to P$ by $(g, p) \mapsto pg^{-1}$ (or, analogously, $\theta_g : P \to P$ by $\theta_g(p) = pg^{-1}$). We have that $\theta$ lifts to an action on $TP$ by pushforward $\theta_*$.
   (b) Does $\theta_*$ act on $ker(\pi_*)$? In other words, does the action of $\theta_*$ as described above send elements of $ker(\pi_*)$ to $ker(\pi_*)$, or does the action somehow take elements of $ker(\pi_*)$ outside of $ker(\pi_*)$? The trick here is to recall that the projection $\pi : P \to M$ is invariant under the action of $G$ on $P$, so we have that $\pi \circ \theta = \pi$. Therefore,

$(\theta_*)\big|_{ker(\pi_*)} : ker(\pi_*) \to ker(\pi_*)$ and so the action's well-defined-ness trivially follows from its well-defined-ness on $TP$.

(c) Alas, here is the reason why it is most useful to realize $\pi^*TM \subset P \times TM$: we lift $\theta$ onto $P \times TM$ via its action on $P$. Let $\theta_g : P \times TM \to P \times TM$ by $(p, v) \mapsto (pg^{-1}, v)$

3. All the maps in the sequence are equivariant with respect to the given action

**Proposition 5.3.1.** $ker(\pi_*) \cong P\big|_q \times \mathfrak{g}$

*Proof.* First, define a map $\psi : P \times \mathfrak{g} \to ker(\pi_*)$ by $\psi(q, X) = \hat{\theta}(X) = \frac{d}{dt}\Big|_{t=0} q \cdot exp(tX)$. We, of course, must confirm that $\frac{d}{dt}\Big|_{t=0} q \cdot exp(tX) \in ker(\pi_*)$:

$\pi_*(\psi(q, X)) = \frac{d}{dt}\Big|_{t=0}(\pi(p \cdot exp(tX)))$. By the invariance of $\pi$ by the action of $G$, we have that $\frac{d}{dt}\Big|_{t=0}(\pi(p \cdot exp(tX))) = \frac{d}{dt}\Big|_{t=0}\pi(p) = 0$, thus $\psi(q, X) \in ker(\pi_*)$

The fact that $\psi$ is a homomorphism is straight-forward, so we will show that it is injective by showing that its kernel is trivial:

$ker(\psi\Big|_q) = \{X \in \mathfrak{g} — \psi(q, X) =\}$. Note that if $X \neq 0$, then $\frac{d}{dt}\Big|_{t=0}(p \cdot exp(tX)) \neq 0$, thus the contrapositive of this statement gives us that $\frac{d}{dt}\Big|_{t=0}(p \cdot exp(tX)) = 0 \Rightarrow X = 0 \Rightarrow ker(\psi) = \{0\}$

For surjectivity, note that for any $X \in ker(\pi_*)$, we can attain its infinitesimal generator by letting $G$ act on the fibers of $P$ □

### 5.3.1 Connections on a Principal Bundles

We will give one general definition of a connection and then refine it until we arrive at the definition of the connection 1-form:

**Definition.** *A connection on $P$ is a map $A : TP \to ker(\pi_*)$ such that:*

1. $A\big|_{ker(\pi_*)} \equiv \mathbb{1}$
2. $A((\theta_g)_*v) = (\theta_g)_*A(v)$, *i.e. $A$ is equivariant with respect to the action of $G$ on $P$.*

By employing the isomorphism $\psi : P\big|_q \times \mathfrak{g} \to ker(\pi_*)$, we refine our definition of a connection as a map $A : TP \to \mathfrak{g}$ such that:

1. $A(\psi(q, X)) = X$ (i.e. if $X$ is the infinitesimal generator of a curve on $P$, then the connection returns the infinitesimal generator)
2. If $g \in G$ and $v \in TP$, then $A((\theta_g)_*v) = Ad(g)A(v)$

Connections on principal bundles are discussed in terms of a connection 1-form, which is an element $\omega \in \Omega^1(P, \mathfrak{g}) \cong C^\infty(P, T^*P \otimes \mathfrak{g})$. The properties above translate to:

1. $Ad(g)(R_g^*\omega) = \omega$ where $Ad(g)(X) = \frac{d}{dt}\Big|_{t=0} g \cdot exp(tX) \cdot g^{-1}$
2. if $\hat{X} \in \mathfrak{g}$ is the infinitesimal generator of $X \in \Gamma(TM)$, then $\omega(X) = \hat{X}$

The isomorphism $\psi$, which illuminates that the vertical subbundle at a point $q \in TP$ is isomorphic to $\mathfrak{g}$, helps make sense of a 1-form evaluated on $P$ that takes values in $\mathfrak{g}$.

### 5.3.2 Relation between Riemannian Connections and Principal Connections on the Frame Bundle

Consider $F(M)$, the frame bundle, as a principal $GL(n, \mathbb{R})$-bundle. A connection 1-form, $\omega$, on $F(M)$ takes an element of $F(M)$, which is a collection on tangent vectors at a point on $M$, and gives back an element of $\mathfrak{gl}(n, \mathbb{R})$.

Recall in our discussion of the Christoffel symbols that we can define a connection by locally defining Christoffel symbols. In Proposition 4.2.6(1) we showed that a connection on a manifold, in the Riemannian sense, boils down to a unique matrix of 1-forms, and in the proof of 4.2.6(1) we showed that they are uniquely defined by the Christoffel symbols.

Let $\sigma : U \subset M \to \pi^{-1}(U) \subset F(M)$ be a local section. By pulling back $\omega$ via $\sigma$, i.e. taking $\sigma^*\omega$, we recover a unique matrix of 1-forms. By considering $TM$ as an associated bundle to $F(M)$, we can recover the Riemannian sense of a connection by taking $\sigma^*\omega$ and pushing it forward by $\pi' : TM \to M$. We can also send $\omega$ into $TM$ via the isomorphism $F(M) \times \mathbb{R}^n / \sim \mapsto TM$.

It is in the sense described in the previous paragraph that connections on principal bundles describe Riemannian connections. However, it's not entirely true that principal $G$-bundle connections *simplify* to the Riemannian case; it is much more accurate to claim that the two exist in correspondence. Via the formulations of:

1. $TM$ as an associated vector bundle to $F(M)$

2. Pulling back $\omega \in \Omega^1(F(M), \mathfrak{gl}(n, \mathbb{R}))$ to $M$ via a local section $\sigma : U \subset M \to \pi^{-1}(U) \subset F(M)$

we can clearly think of connections on $F(M)$ as "interchangeable" or "in correspondence with" Riemannian connections.

### 5.3.3 The Curvature 2-Form

**Definition.** *We define the* ***curvature 2-form of a connection*** $\omega$ *to be* $\Omega(X, Y) = d\omega(X, Y) + [\omega(X), \omega(Y)]$

**Proposition 5.3.2.** *In local coordinates,* $\Omega^i = d\omega^i + \frac{1}{2}c^i_{jk}\omega^j \wedge \omega^k$ *where* $c^i_{jk}$ *are the structure constants defined in section 4.7.*

*Proof.* This is easily proven via explicit computation; we begin by defining $\{e_i\}$ to be a basis for $\mathfrak{g}$ and writing $\omega = \omega^i e_i$. Thus, we have:

$[\omega(X), \omega(Y)] = \omega(X)\omega(Y) - \omega(Y)\omega(X) = \omega^i(X)e_i\omega^j(Y)e_j - \omega^i(Y)e_i\omega^j(X)e_j = \omega^i(X)\omega^j(Y)(e_ie_j - e_je_i) = \omega^i(X)\omega^j(Y)[e_i, e_j] = \omega^i(X)\omega^j(Y)c^k_{ij}e_k = \frac{1}{2}\omega^i \wedge \omega^j c^k_{ij}e_k(X, Y)$ (where for the final equality and the introduction of the $\frac{1}{2}$ coefficient, we recalll that $\omega^i \wedge \omega^j = \frac{1}{2}(\omega^i \otimes \omega^j - \omega^j \otimes \omega^i)$ □

Recall that one of Maxwell's equations, in its more general form, is $dF = 0$. In fact, this equation was not derived via Yang-Mills:

**Proposition 5.3.3.** $dF = 0$*, where* $F$ *is the curvature 2-form*

The above is proven in [7]. Yang-Mills Theory can be thought of as giving us the *second* condition $\star d \star F = 0$ or $\star d \star F = J$ (whether $\star d \star F$ equals 0 or $J$ will be discussed in section 6.2).

It *is not* immediately obvious (at least when I first encountered it) that $\Omega$ above has anything to due with curvature. The idea is the following:

Let $\sigma : U \subset M \to P$ be a section, then we can pullback the connection on $P$ into $U$. We already know, from Proposition 4.6.1, that the Riemannian sense of the curvature 2-form is deeply related to the Riemann curvature endomorphism. By relating principal connections on $F(TM)$ with their Riemannian counterparts, and then analyzing the curvature 2-forms of each of those types of connections, we can relate our notion of curvature on a principal bundle to the more familiar and intuitive notion of curvature in the Riemannian sense. Of course, principal bundles extend far beyond the concept of a frame bundle and thus interpretations of "curvature" in more abstract settings can yield fascinating descriptions of geometry.

Let $\sigma$ be a local section as above and define $A = \sigma^*\omega$ and $F = \sigma^*\Omega$ as the *local representatives of the connection and curvature forms.* We have that $F^i = dA^i + \frac{1}{2}C^i_{jk}A^j \wedge A^k$. In fact, this equation will arise again shortly: in considering electromagnetism as a $U(1)$ gauge, we consider the electromagnetic potential, $A$, as the connection 1-form on the principal $U(1)$-bundle, and the curvature 2-form is therein identified as the Faraday tensor.

Saying that electromagnetism is a $U(1)$ gauge means that we can take $A \mapsto A + df$ where $df$ is a function valued in the Lie algebra of $U(1)$, which is simply $\mathbb{R}(\cong \mathfrak{u}(1))$, and the physical quantity, $F$, is unaffected.

### 5.3.4 A First Look at Quantum Chromodynamics

As has been noted right above, if we consider our connection $A$ to be the electromagnetic four potential, then we recover the curvature 2-form as the Faraday tensor familiar to electromagnetism. As I have alluded to numerous times, our constructions in this thesis can be easily applied to attain dynamical equations for the Standard Model of Particles, so here I will scratch the surface on some constructions in the strong interaction, also known as *Quantum Chromodynamics (QCD).* The strong interaction describes the strong force, which is what holds elementary particles together; for example, while the electron is known to be an elementary particle, the proton and the neutron are not. The proton consists of two up quarks and one down quark, whereas the neutron consists of one up quark and two down quarks. The strong force is the force that binds these quarks together, thus it is dominant at short distances (overpowering both the electromagnetic force and the electroweak force), but falls off quickly with distance.

We have $\Omega^i = d\omega^i + \frac{1}{2}C^i_{jk}\omega^j \wedge \omega^k$ where $\omega$ is the connection on our principal $G$-bundle. In QCD, the gauges come from $SU(3)$, so $\omega$ is a connection on a principal $SU(3)$-bundle. Much like in electromagnetism, the connection is denoted by $\mathcal{A}$.

The connection $\mathcal{A}$ is entitled the **Gluon field** and is analogous to the electromagnetic potential (analogous in a physical sense). The curvature 2-form is denoted as $\mathcal{G}$ and is entitled the textitGluon field strength tensor, and is locally written as $\mathcal{G}^i_{jk} = \partial_j\mathcal{A}^i_k - \partial_k\mathcal{A}^i_j + gf^{ibc}\mathcal{A}^b_j\mathcal{A}^c_k$, where $g$ is called a **coupling constant** and $f^{ibc}$ are the structure constants of $SU(3)$.

# 6 Yang-Mills

## 6.1 The Exterior Covariant Derivative

Before we give the Yang-Mills equations, we need to describe one last type of operator, the *exterior covariant derivative*. The idea is the following:

We already have the exterior derivative, an operator $d : \Omega^p(M) \to \Omega^{p+1}(M)$, where $\Omega^p(M)$ consists of real-valued $p$-forms on $M$; however, our connection 1-form $\omega$ is a $\mathfrak{g}$-valued form on $P$. Whereas $d : \Omega^p(M) \to \Omega^{p+1}(M)$ makes sense because we know how to take partial derivatives of real-valued functions, we run into more trouble for our $\omega$: how exactly can we take an exterior derivative of something that looks like $\omega = f_i dx^i$, where $f_i$ is a $\mathfrak{g}$-valued smooth function and $x^i$ are local coordinates for $P$? The answer is to combine the notion of an exterior derivative with our already established notion of a covariant derivative:

Let $\eta$ be an arbitrary $E$-valued $p$-form, i.e. $\eta \in \Omega^p(E; M)$, where $\Omega^p(E; M) = \{$ $p$-forms on $M$ that return values in a fiber bundle $E\}$.

**Definition.** *The* ***exterior covariant derivative*** *of an $E$-valued $p$-form $\eta = s_J \otimes dx^J$ (where $dx^J$ denotes the wedge product $dx^{j_1} \wedge \cdots \wedge dx^{j_p}$, $x^i$ are local coordinates on $M$, and $s_J$ is a local section of $E$) is an operator $d_\omega : \Omega^p(E; M) \to \Omega^{p+1}(E; M)$ defined by $d_\omega(\eta) = \omega_k s_J \otimes dx^k \wedge dx^J$ where $\omega$ is a connection 1-form on $E$ and $\omega_k$ is shorthand for $\omega_{\partial_k}$. More generally, we can write:*

$$d_\omega \eta = \omega_k \otimes dx^k \wedge \eta$$

## 6.2 The Yang-Mills Equations

The Yang-Mills Equations came about as a result of researching mathematical formalisms of Lagrangian physics. The question arises from the study of the quantity $F \wedge \star F$, called the **Yang-Mills Action**.

In electromagnetism, we have $F = B + E \wedge dt$. We already computed $\star F$ in section 1.3, so we can clearly see that:

$$F \wedge \star F = (|E|^2 + |B|^2) dV_\eta$$

where $\eta$ is the Minkowski metric and $|E|^2, |B|^2$ are the magnitudes of the electric and magnetic forms, respectively, with respect to the Minkowski metric. Before delving into the physics below, it should be noted that all of this can be found in any introduction to electromagnetism textbook; [8] is beautifully optimal.

One may recall the Poynting theorem in electromagnetism:

$$\partial_t u = -\vec{\nabla} \cdot \vec{S} - \vec{J}_f \cdot \vec{E}$$

Where $\vec{J}_f$ is the current density free charges in a given region. This describes *Poynting vector*, $\vec{S}$, which describes the outward flux of energy from a radiating electromagnetic system. $u$ is entitled the density total energy stored in electromagnetic fields over volume and it is defined as $u = \frac{1}{8\pi}(E^2 + B^2)$ in Gaussian units.

Therefore, our quantity $F \wedge \star F$ corresponds to an energy density of sorts (of course, this is not immediately obvious because $F$ describes the curvature on some principal $G$-bundle, again $G = U(1)$ for electromagnetism). When we integrate over the entire space, we get

$$\int_P F \wedge \star F = \text{the total energy}$$

It turns out that the critical points of this functional, which gives the restrictions on $F$ that optimizes the total energy, are:

1. $d_\omega F = 0$
2. $\star d_\omega \star F = 0$

This should seem mildly unexpected, as we might expect to recover the second equation as $\star d \star F = J$ where $J$ is the so-called generalized current. If we modify our equation, $\int_P F \wedge \star F$ by $\int_P F \wedge \star F + \star J \wedge A$ (where $A$ is the connection, and $J$ is a 1-form representing something analogous to a current or charge), then we recover the second equation as $\star d_\omega \star F = J$.

## 6.3 Maxwell's Equations and the Yang-Mills Equations

We already did this in Section 1.3, so we really just need to make one remark:

Recall that we have $F_{ij} = dA_i + A_i \wedge A_j$. We note that since the connections $A_i$ are $\mathfrak{u}(1)$-valued and $\mathfrak{u}(1)$ is abelian (this can be seen in the obvious way that $U(1)$ is abelian, so its Lie algebra is, or the fact that $\mathfrak{u}(1) = \mathbb{R}$ is abelian), the wedge product disappears! Why exactly does it disappear? Recall that we have $A \wedge A = \frac{1}{2}(A_i \otimes A_j - A_j \otimes A_i)$ and since $\mathfrak{u}(1)$ is abelian, we have that $A_i \otimes A_j = A_i \otimes A_j$. Thus we have $F = dA \Rightarrow F_{ij} = \partial_i A_j - \partial_j A_i$. This curvature 2 form is the Faraday 2-form in terms of the connection 1-form on a principal $U(1)$-bundle, which is the electromagnetic potential 1-form, $A = \phi dt + A_x dx + A_y dy + A_z dz$. In flat spacetime, the principal $U(1)$-bundle is trivial, i.e. $\mathbb{R}^4 \times U(1)$, over $\mathbb{R}^4$, where $\mathbb{R}^4$ is equipped with the Minkowski metric.

## 6.4 A Second Look at Quantum Chromodynamics

The process above for Maxwell's Equations is analogous to the procedure for describing the laws of QCD. Our curvature form remains as $\mathcal{G}_{ij} = d\mathcal{A}_i + \mathcal{A}_i \wedge \mathcal{A}_j$, without the wedge product disappearing since $SU(3)$ is not abelian. Just like we did in section 1.3, we could derive equations of motion for the strong interaction:

We have already defined the connection on a principal $SU(3)$-bundle to be the so-called Gluon Field $\mathcal{A}$, though we do not give more details, so Yang-Mills would give the equations $d_\mathcal{A}\mathcal{G} = 0$ and $\star d_\mathcal{A} \star \mathcal{G} = J$. However, the individual components of the quantities in this equation are physical, and sadly there is not enough time for me to learn them for this thesis.

Nearly all modern particle physics is done through the analysis of Lagrangians and their quantization via the path-integral formulation. For the sake of completion, the QCD Lagrangian is listed below:

$$\mathcal{L}_{QCD} = \bar{\psi}_i(i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \tfrac{1}{4}\mathcal{G}^a_{\mu\nu}\mathcal{G}_a^{\mu\nu}$$

where $\psi_i$ is the quark field, $\gamma^\mu$ are the Dirac matrices, $m$ is the quark mass, and $D_\mu$ is the gauge covariant derivative $D_\mu := \partial_\mu - ig\mathcal{A}_\mu^\alpha \lambda_\alpha$ where $\gamma_\alpha$ are the Gell-Mann matrices (which form the "standard basis" for $\mathfrak{su}(3)$, they are analogous to the Pauli spin matrices for $\mathfrak{su}(2)$). We will conclude this thesis with a quick explanation of the gauge covariant derivative, and a bare-bones explanation of the Aharonov-Bohm Effect.

## 6.5 Final Remarks

### 6.5.1 The Gauge Covariant Derivative

The idea about a gauge is that we can take the connection on our principal $G$-bundle, and add a $\mathfrak{g}$-valued function without altering the curvature, $F$. Quantum electrodynamics is a $U(1)$-gauge, meaning we can add a $\mathfrak{u}(1)$-valued function to $A$ without altering $F$:

Suppose we add $df$, where $df$ is $\mathbb{R}$ valued, to $A$. We would have $\phi dt + A_x dx + A_y dy + A_z dz + \partial_t f dt + \partial_x f dx + \partial_y f dy + \partial_z f dz = (\phi + \partial_t f)dt + (A_x + \partial_x f)dx + (A_y + \partial_y)dy + (A_z + \partial_z)dz$.

Quantum mechanics is governed the Schrodinger Equation: $i\partial_t |\psi\rangle = \hat{H}|\phi\rangle$ where the Hamiltonian, $\hat{H}$, has the form $\hat{H} = \frac{1}{2m}(-i\nabla - eA_m)^2 + e\phi$ (where $A_m$ is just the magnetic portion of the potential 1-form). We see that by adding a gauge to $A$, we get (setting constants to 1):

$$\hat{H} \mapsto \frac{1}{2}(-i\nabla - eA_m + \nabla f)^2 + \phi + \partial_t f = \frac{1}{2}(-i(\nabla + ieA_m) + \nabla f)^2 + \phi + \partial_t f$$

This term $\nabla + ieA_m$ is called the **gauge covariant derivative** (in this case, it is valid if there is no external field). It is essentially what happens to the partial derivative terms inside the original Hamiltonian when we apply a gauge.

### 6.5.2 The Aharonov-Bohm Effect

The mathematics behind the physics of the Aharonov-Bohm Effect corresponds to the *holonomy* (which will not be defined here, but is described in chapter 13 of [3]) of a flat connection on a principal $U(1)$-bundle. By flat connection, we mean one with no curvature; i.e., a connection $A \in \Omega^1(P; \mathfrak{u}(1))$ such that $F = dA = 0$. Since $F = B + E \wedge dt$ in electromagnetism, we have that the $B$ and $E$ forms are both zero.

An early experiment by Werner Ehrenberg and Raymond Siday, and later reobserved by Aharanov and Bohm, showed one of the quirks of quantum physics. The experiment showed that if one confines an electric particle to a region with no electric or magnetic fields, the particle will still undergo some physical process, as if it is reacting to some sort of electric or magnetic field. This effect is a manifestation of the change in phase of the wavefunction; in other words:

If our system is originally in the state $|\psi\rangle$, and we alter our potential via a gauge, such as $A + df$, then the wavefunction will be transformed such that $|\psi\rangle \mapsto e^{-if}|\psi\rangle$.

A more rigorous treatment of the Aharonov-Bohm Effect can be found in [8]. I find it rather incredible that mathematicians have been able to elegantly describe this physics via flat connections on principal $U(1)$-bundles (and their holonomy); whereas the abstract definitions of principal $G$-bundles and the relevant differential geometry might seem esoteric, they actually allow us to tap into some notion of mathematical truth that is intrinsic in physics.

# References


[1] Misner, Charles W., Kip S. Thorne, and John Archibald Wheeler. Gravitation. San Francisco: W.H. Freeman, 1973.

[2] Lee, John. Introduction to Smooth Manifolds. 2nd ed. New York: Springer, 2013.

[3] Taubes, Clifford. Differential Geometry: Bundles, Connections, Metrics and Curvature. Oxford: Oxford U, 2012.

[4] "Baker-Campbell-Hausdorff Series." Baker-Campbell-Hausdorff Series – from Wolfram MathWorld. N.p., n.d. Web. 23 Apr. 2017.

[5] F., Torres Del Castillo G. Differentiable Manifolds: A Theoretical Physics Approach. New York: Springer, 2012.

[6] Lee, John. Riemannian Manifolds: An Introduction to Curvature. New York: Springer, 1997.

[7] Morrison, Scott. "Connections on Principal Fibre Bundles." (3 Sept. 2007): 1-18. Web.

[8] Frankel, Theodore. The Geometry of Physics: An Introduction. 3rd ed. Cambridge: Cambridge UP, 2012.


# Appendix A Definitions

$dV_g$ is the Riemannian volume form; it is defined such that $\int_M dV_g$ gives the "volume" (length if one dimensional, area if two dimensional, etc.) of the manifold $M$. It is defined by $dV_g = \sqrt{|det(g)|}dx_1 \wedge \cdots \wedge dx_n$ where $g$ is the Riemannian metric of $M$ and $x_i$ are (smooth) coordinates.

$GL(n, F)$ is the set of all invertible matrices acting on $F^n$, namely the matrices with nonzero determinant; it is referred to as the general linear group.

$SL(n, F)$ is the set of all invertible matrices acting on $F^n$ with unit determinant; it is referred to as the special linear group.

$O(n)$ is the set of all matrices $A$ in $GL(n, \mathbb{R})$ such that $A^T = A^{-1}$; it is referred to as the orthogonal group.

$SO(n)$ is the subset of $O(n)$ consisting of matrices with unit determinant; it is referred to as the special orthogonal group.

$U(n)$ is the set of all matrices $A$ in $GL(n, \mathbb{C})$ such that $A^\dagger = A^{-1}$ where $A^\dagger = \bar{A}^T$ is the hermitian conjugate; it is referred to as the unitary group.

$SU(n)$ is the subset of $U(n)$ consisting of matrices with unit determinant; it is referred to as the special unitary group.

**Remark.** *The following are all referred to as "the Lie algebra of..."*

$\mathfrak{gl}(n, F)$ is the set of all matrices acting on $F^n$; it is referred to as the Lie algebra of the general linear group.

$\mathfrak{sl}(n, F)$ is the set of all matrices in $\mathfrak{gl}(n, F)$ with zero trace.

$\mathfrak{o}(n)$ is the set of all matrices in $\mathfrak{gl}(n, \mathbb{R})$ that are anti-symmetric, i.e., $A^T + A = 0$.

$\mathfrak{u}(n)$ is the set of all matrices in $\mathfrak{gl}(n, \mathbb{C})$ that are anti-hermitian, i.e. $A^* + A = 0$.

$\mathbb{RP}^n$ is the set of all real lines passing through 0 in $\mathbb{R}^{n+1}$; $\mathbb{CP}^n$ and $\mathbb{HP}^n$ are defined analogously.

# Appendix B Some Identities

**Identity B.1.** *Let $m : G \times G \to G$ be Lie multiplication and identify $T_{(e,e)}(G \times G)$ as $T_eG \oplus T_eG$. $dm_{(e,e)} : T_{(e,e)}(G \times G) \to T_eG$ is thus given by $dm_{(e,e)}(g, q) = g + q$*

*Proof.* First, note that $dm_{(e,e)}(g, q) = dm_{(e,e)}(g, 0) + dm_{(e,e)}(0, q) = dm_eg + dm_eq$

$dm_eg$ comes from $d(m(e, g))$, but $m(e, g) = \mathbb{1}_g \Rightarrow dm_e = \mathbb{1}_g \Rightarrow dm_eg = g$ and $dm_eq = q$

Thus, $dm_{(e,e)}(g, q) = g + q$ □

**Identity B.2.** *Let $i : G \to G$ be Lie inversion. $di_e : T_eG \to T_eG$ is given by $di_e(g) = -g$*

*Proof.* $m(g, i(g)) = g(i(g)) = e \Rightarrow dm(g, i(g)) = 0$. We thus have, thanks to part 1., that $dm(g, i(g)) = \mathbb{1}_gg + di_eg = g + di_eg = 0 \Rightarrow di_eg = -g$ □

# Appendix C Assumed Theorems

All of the propositions below (and their proofs) can be found in [3].

**Theorem C.1.** *Suppose $M$ and $N$ are smooth manifolds and $F : M \to N$ is a map. $F$ is a local diffeomorphism $\Leftrightarrow$ it is both a smooth immersion and a smooth submersion.*

**Theorem C.2.** *If $\Gamma$ is a discrete Lie group acting continuously and freely on a manifold $M$, then the action is proper if and only if the following conditions hold:*

1. *Every $p \in M$ has a neighborhood $U$ such that for each $g \in \Gamma$, $(g \cdot U) \cap U =$ unless $g = e$*
2. *If $p, q \in M$ are not in the same orbit of the $\Gamma$-action, then there is a neighbhorhood $V$ of $p$ and $W$ of $q$ such that $(g \cdot V) \cap W =$ for all $g \in \Gamma$*

**Theorem C.3.** *If $M$ is a connected smooth manifold and $\Gamma$ is a discrete Lie group acting smoothly, freely, and properly on $M$, then the orbit space $M/\Gamma$ is a smooth manifold with a unique structure such that $\pi : M \to M/\Gamma$ is a smooth normal covering map.*

**Theorem C.4.** *If $\pi_1 : M \to N_1$ and $\pi_2 : M \to N_2$ are both quotient maps of topological spaces that are constant on each others fibers, then there is a unique homeomorphism $\Phi : N_1 \to N_2$ such that $\Phi \circ \pi_1 = \pi_2$. If $M, N_1$, and $N_2$ are smooth manifolds, then $\Phi$ is a diffeomorphism.*

**Theorem C.5.** *Let $M$ be a smooth manifold; suppose that for each $p \in M$ we are given a real vector space $E_p$ of some fixed dimension $k$. Let $E = \coprod_{p \in M} E_p$; let $\pi : E \to M$ be the map that* *takes $E_p \mapsto p$. If the following three conditions are satisfied, then $E$ has a unique smooth structure making it into a smooth manifold and a smooth rank $k$ vector bundle over $M$, with $\pi$ as projection and $\{(U_i, \Phi_i)\}$ as smooth local trivializations.*

1. *An open cover $\{U_i\}_{i \in I}$ of $M$*
2. *For each $i \in I$, a bijective $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k$ whose restriction to each $E_p$ is a vector space isomorphism from $E_p$ to $\{p\} \times \mathbb{R}^k$*
3. *For each $i, j \in I$ with $U_i \cap U_j \neq \emptyset$, a smooth map $\tau_{ij} : U_i \cap U_j \to GL_k(\mathbb{R})$ such that the map $\Phi_i \circ \Phi_j^{-1} : (U_i \cap U_j) \times \mathbb{R}^k \to (U_i \cap U_j) \times \mathbb{R}^k$ has the form $\Phi_i \circ \Phi_j^{-1}(p, v) = (p, \tau_{ij}(p)v)$*

**Theorem C.6.** *If $F : M \to N$ is a smooth map, $p \in M$, $v \in T_pM$, then $dF_p(v) = (F \circ \gamma)'(0)$ for any smooth $\gamma : J \subset \mathbb{R} \to M$ such that $0 \in J$, $\gamma(0) = p$, and $\gamma'(0) = v$*