Geometry of gauge theories

Turk, Marija

UNIVERSITY OF RIJEKA
FACULTY OF PHYSICS

Marija Turk

GEOMETRY OF GAUGE THEORIES

Master's Thesis

Rijeka, 2024

University of Rijeka - Faculty of physics

Master's degree - Particle Physics and Astrophysics

Marija Turk

GEOMETRY OF GAUGE THEORIES

Master's Thesis

Supervisor: dr. sc. natur. Mateo Paulišić

Rijeka, 2024

Abstract

The goal of this thesis was to mathematically explain concepts from gauge theories. The most emphasis was put on explaining gauge transformations, covariant derivatives, and curvature. The main tools we used for such description were *principal fiber bundles* and *connections*. As there exists a lot of background geometry, preparations had to be made before the introduction of those two. Those preparations included an understanding of fiber bundles and Lie groups. Fiber bundles allow for the description of spaces that are not trivial, i.e. those that can be described only locally. Lie groups, on the other hand, were used as a tool for the description of the symmetries we were dealing with. On principal fiber bundles connections were defined as a formal way of comparison that is not well defined in Euclidean geometry. This helped us understand gauge transformations. However, for a full description of covariant derivatives, associated bundles need to be mentioned as they serve as a tool for defining *parallel transport map* in the context of a vector bundle in which we know how to compare points. This is enough for a formal definition of a covariant derivative, which can be found in both particle physics and general relativity. Curvature relied on no more new concepts as understanding covariant derivative and exterior covariant differentiation was enough for its description. Examples of curvature in gauge theories were given alongside terminology used in these fields as naming conventions differ.

Keywords: Gauge theory, Fiber bundles, Principal fiber bundles, Associated bundles, Lie groups

Contents

§1 Introduction

One of the biggest issues gauge theories have to compensate for is that of *comparison of quantities defined on different points in spacetime.* From the point of view of someone who has only known *Eucledian* geometry, this issue might seem laughable, however, as we will soon learn, this is no trivial matter. This issue can formally be solved with the use of *connections*. To illustrate where such issues arise, we will be using examples given in Schwartz [1], and alongside, we will be noting to what exactly a term *connection* refers to in these cases.

Let $\phi(x)$ denote any scalar field invariant under a gauge transformation $\phi(x) \rightsquigarrow e^{i\alpha}\phi(x)$. The meaning of gauge symmetry is that the phase is a matter of convention. Suppose we are interested in knowing how such a field behaves at two specific points x and y that are very far from each other. Locally, the convention one chooses at a point x should be independent of a convention chosen at a point y. However, if those conventions are different, how can we tell if $\phi(x) = \phi(y)$? If those conventions are to be changed to a set of different conventions, we would get

$$\phi(y) - \phi(x) \rightsquigarrow e^{i\alpha(y)}\phi(y) - e^{i\alpha(x)}\phi(x), \tag{1.1}$$

making it, again, impossible to distinguish between those two fields. This is not only a matter of wanting to know if a given field is equal at two different points. If we are not able to determine $|\phi(y) - \phi(x)|$ it is impossible to compute $\partial_\mu\phi(x)$, since derivatives naturally contain a difference. To make such a comparison possible, *covariant derivatives* are introduced in the form of

$$D_\mu\phi(x) = \partial_\mu\phi(x) - iA_\mu\phi(x). \tag{1.2}$$

We can write that

$$D_\mu\phi(x) \rightsquigarrow e^{i\alpha(x)}D_\mu\phi(x). \tag{1.3}$$

A_μ is known as a *gauge field*, and in this context, it represents a *connection.* The gauge field itself also follows a transformation law, known as a *gauge transformation.* In the case of Abelian gauge theories we can write

$$A_\mu \rightsquigarrow A_\mu + \partial_\mu \alpha(x). \tag{1.4}$$

Another case that exemplifies the issue of comparison comes from general relativity. There, Christoffel symbols are employed as a way to compensate for the fact that we are calculating derivatives in a curved spacetime. A *Christoffel connection* in this context then allows for the comparison of fields at two different points, despite their different local coordinate frames.

As we will learn later in this work, the term *connection* does not only involve such specific examples, but is more general as it refers to a specific kind of structure built on a *principal fiber bundle.* A connection in the context of principal fiber bundles is a much more general concept than the gauge field physicists refer to as a connection. As we will see later, the gauge field should be identified with the *pull-back of a connection 1-form to a base manifold.* The main goal of this thesis is to formally explain *connections* and *covariant derivatives*, and only after doing so, compare those more general and formal terms to the various cases of their use in gauge theories.

Other than being a mathematical background for understanding connections and covariant derivatives, exploring the nature of principal fiber bundles and connections will also lead to a general expression of a *gauge transformation*, so alongside connections and covariant derivatives, gauge transformations will also be studied in great detail. As will be shown below, there exists a strong link between connections and gauge potentials, so the imprecise use of this term in physics can be pardoned.

The issue of comparing the points is a consequence of working in a curved space, so gauge theories also frequently include the term *curvature* for a description of such spaces. As covariant derivatives are needed for a formal description of curvature, and preparation for their formal definition will occupy a majority of this thesis, *curvature* will be the last concept this thesis will include.

§2 Bundles

By themselves, tangent vectors, cotangent vectors, and tensors are all defined pointwise. For us to be able to do anything remotely similar to calculus, they should all be defined on the whole manifold. However, this requires more structure - namely, **bundles**, or more specifically, we are going to need **tangent** and **cotangent bundles**.

A *bundle* as a concept is a very simple idea, making it a perfect tool for the introduction of more complex concepts that build on the starting idea. Bundles that are more complex than the aforementioned tangent or cotangent bundle will be introduced. Namely, principal fiber bundles and associated bundles, which will later be used directly in the description of gauge fields.

2.1 Bundle

We can think of bundles as a way of decomposing manifolds into simpler shapes. For example, figure (2.1) depicts a cylinder that can be obtained by gluing a bunch of lengths L onto a circle S^1. In other words, we can think of a cylinder as a product $S^1 \times L$. Such a product is called a **product manifold** [2, 3, 4].

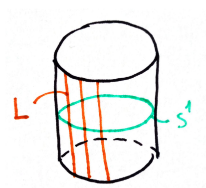


Figure 2.1: Cylinder as a product of a circle S^1 and lengths L.

The projection of a cylinder is a circle S^1. In general, projections of bundles are named a **base spaces** and are labeled B as seen in figure (2.2). The base space for a cylinder is S^1 and the cylinder itself is called a **total**

space. A total space in general is labeled E, and in the case of a cylinder $E = S^1 \times L$.[1] The projection map is labeled π and it projects all the lengths L to single points that make up a base circle. As π projects multiple points into a single point, it is set by definition to be a surjective map. Map π should also be continuous as we should always have points on the base space onto which π projects. The base space B, total space E and the projection map π together form a **bundle**.

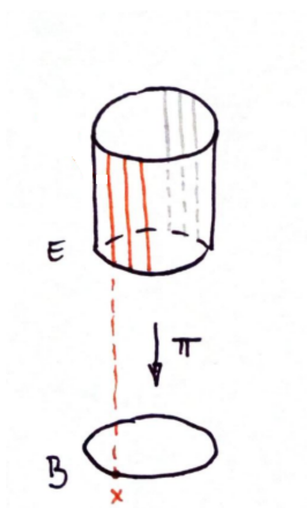


Figure 2.2: Cylinder as a bundle.

Definition 2.1. *Let E and B be topological manifolds. A* ***bundle*** *is a triple (E, B, π), where $\pi : E \to B$ is a continuous and surjective map called a* ***projection map****. E is called a* ***total space*** *and B is called a* ***base space*** *[2, 5, 6].*

Note that, sometimes a bundle is defined in a way that E and B are topological spaces [5, 6]. Another example of a bundle that can easily be visualized is a torus, which is then considered a product $S^1 \times S^1$. Geometrically it be thought of as a circle being attached to each point of a *base* circle.

[1] Here, a total space is a **product manifold**.

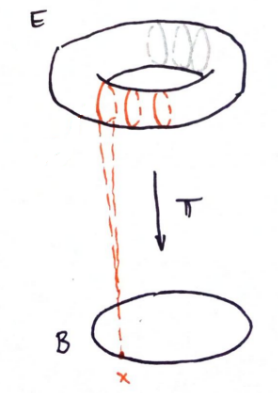


Figure 2.3: Torus as a bundle.

2.2 Fiber bundle

Sometimes it is not possible to describe the entire topological manifold as a product of simpler manifolds. To give an example of such a scenario, we can think of how to physically construct a cylinder. If we start with a strip of paper and glue its ends together we get a cylinder as shown in the figure (2.4).

Figure 2.4: Cylinder from a strip of paper.

We know this shape *can* be described as a product $S^1 \times L$. However, if one were to *turn* one side of our paper, as seen in figure (2.5), and then glue the ends together, one would get a **Möbius strip**.

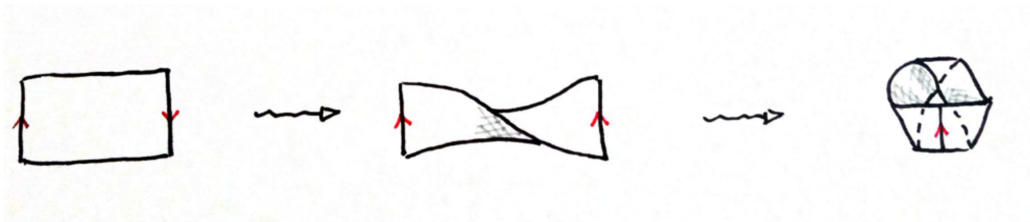


Figure 2.5: Möbius strip from a strip of paper.

A Möbius strip **can not** be described as $S^1 \times L$ - at least, not globally. The twist one gets from turning one side of the paper does not allow us to describe the resulting shape as a product. But, each local section[2] of a Möbius strip **can** be described as a product of a circular arc (of a circle S^1) and the same length L.

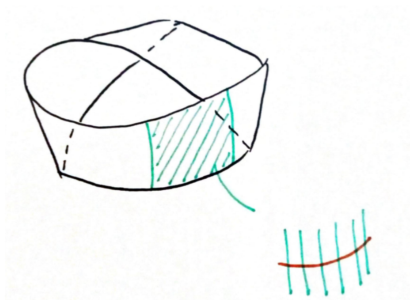


Figure 2.6: Local section of a Möbius strip.

The local section of a Möbius strip, labeled $\pi^{-1}(U_j)$, shown in figure (2.7), can be written as a product $U_j \times L$. As it is represented only locally, we needed a map that takes us from the Möbius strip to a product $U_j \times L$. Such a map is labeled φ_j.

[2]Note that local sections *are* open sets.

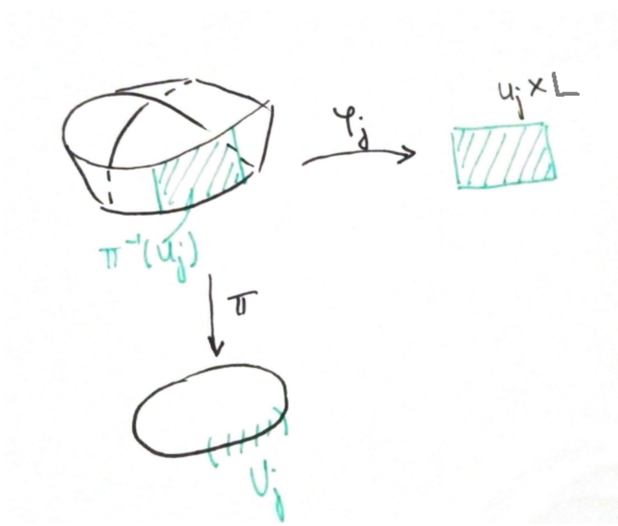


Figure 2.7: Local section of a Möbius strip.

Any bundle whose total space can be described as a *product manifold* is called a **trivial bundle** or a **product bundle**. Local sections of a bundle that can be described as a product manifold are then also considered *locally* trivial bundles [6, 5, 3, 4].[3] For that reason, a map φ_j can be thought of as a map that *trivializes* sections of total space. The pairs $\{(U_j, \varphi_j)\}$ are called **local trivializations** [6, 4]. Some sources [3] call a map $\varphi_j^{-1} : U_j \times F \to \pi^{-1}(U_j)$ a local trivialization.[4]

Figure (2.7) can be represented with a commutative diagram.

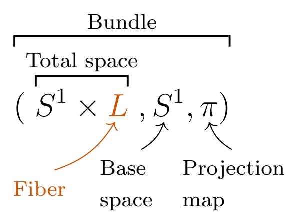
$$\begin{array}{ccc} \pi^{-1}(U_j) & \xleftarrow{\varphi_j^{-1}} & U_j \times F \\ \downarrow \pi & \swarrow \text{Cannonical projection} & \\ U_j & & \end{array} \tag{2.1}$$

[3]Depending on how one thinks about it, they can be considered trivial subbundles as opposed to locally trivial bundles.

[4]When considered in this form, notation is set up in a way that local trivialization is not the inverse. Note that the inverse that appears here is a consequence of considering both terminologies simultaneously.

One should take note of the equation $\pi(\varphi_j^{-1}(x, f)) = x$ that follows from the diagram above as it will make an appearance in the formal definition of a fiber bundle.

If we look at the cylinder that was given as an example of a bundle, we can notice that all things have a name except for lengths L that together with a circle actually make up a cylinder. The lengths L are called **fibers**, and for the cylinder or any trivial bundle, all fibers are the same apart from being defined at different points.



Fibers allow us to see where the twists, like the one presented in the example of a Möbius strip, are located, so it would be good to define them. Let $x \in B$. A **Fiber at the point x** is $\pi^{-1}(x)$ and is denoted F_x. Since one can not distinguish different fibers within a trivial bundle, one can talk about a **typical fiber** - a representative of fibers that are defined at a point. The typical fiber will be denoted F [6, 2, 5, 3, 4].

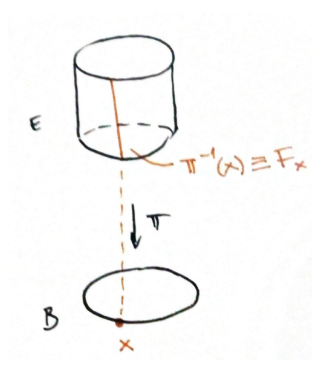


Figure 2.8: Fiber at the point $x \in B$.

To see where the twist is happening we can overlap two sections of a Möbius strip and look at the fiber in that overlap. If one looks at the figure below it is clear that for two sections we get two overlaps. We can construct homeomorphic maps going from fibers at a point to typical fiber to see if typical fiber on different overlaps behaves differently. Such maps are represented in the figure below.

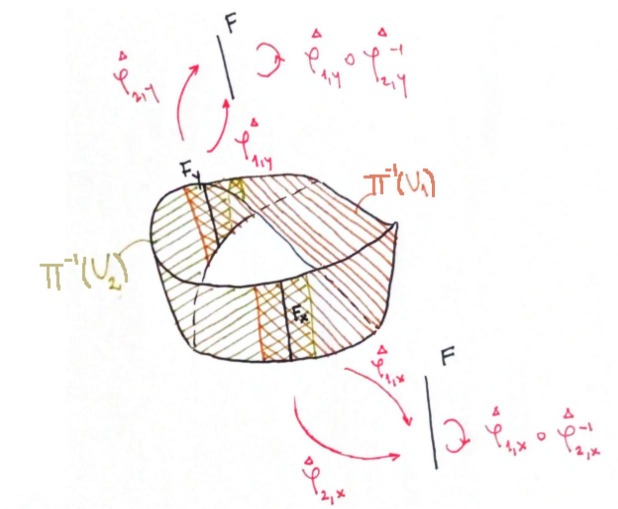


Figure 2.9: Fibers F_x and F_y in the overlap of two local sections $\pi^{-1}(U_1) \cap \pi^{-1}(U_2) \in E$.

As such a picture can get crowded really quickly, a clearer way to represent this sentiment is to imagine individual fibers growing out of a base space. The notation for homeomorphic maps and visualization of those homeomorphisms come from [6, p. 125].

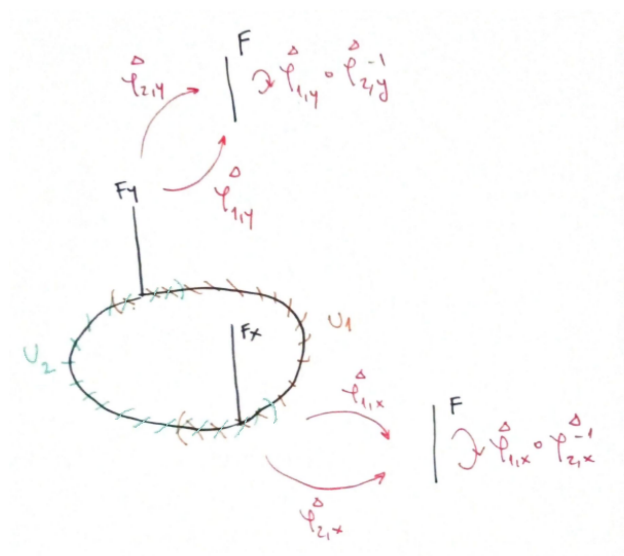


Figure 2.10: Fibers F_x and F_y in the overlap of two local sections $U_1 \cap U_2 \in B$.

Homeomorphisms that go from F_x to F are labeled $\overset{\triangle}{\varphi}_{1,x}$ and $\overset{\triangle}{\varphi}_{2,x}$ and they map in the following manner:

$$\overset{\triangle}{\varphi}_{1,x} : F_x \to F, \text{ where } F_x \in \pi^{-1}(U_1) \tag{2.2}$$

$$\overset{\triangle}{\varphi}_{2,x} : F_x \to F, \text{ where } F_x \in \pi^{-1}(U_2). \tag{2.3}$$

Homeomorphisms that go from F_y to F are labeled $\overset{\triangle}{\varphi}_{1,y}$ and $\overset{\triangle}{\varphi}_{2,y}$ and they map in the following manner:

$$\overset{\triangle}{\varphi}_{1,y} : F_y \to F, \text{ where } F_y \in \pi^{-1}(U_1) \tag{2.4}$$

$$\overset{\triangle}{\varphi}_{2,y} : F_y \to F, \text{ where } F_y \in \pi^{-1}(U_2). \tag{2.5}$$

As such homeomorphisms map from specific sections U_1 and U_2 which are overlapping, they will help us pin down the exact overlap on which the twist is happening. To see how fibers show information about the twist we can look at the specific example.

Example 2.1. *The following example is a modified version of an example given in [6, p. 126]. Let us start with a strip that is* $5a$ *wide and* 2 *long as shown in figure (2.11). The shaded area should overlap to form a Möbius strip. Labels for the width are kept on the Möbius strip so we can keep track of what lands were. The bottom line where the labels for the width are, is also colored, so that we can see what happens to it after forming a Möbius strip.*

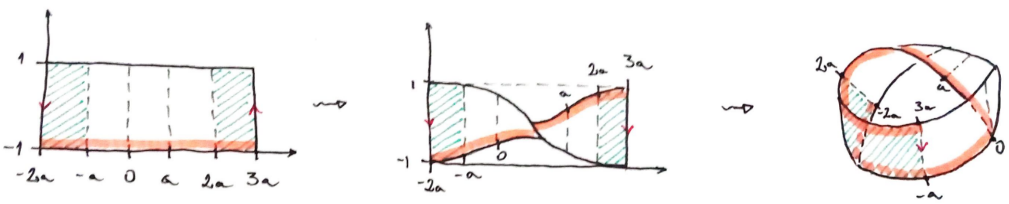


Figure 2.11: Möbius strip.

We can now project the Möbius strip in a way that we see the labeled points in the base space. For the sake of keeping track of where the labeled points land the midpoint between a total space and a base space is added, as seen in figure (2.12). Because of the overlap, we get $-2a = 2a$ *and* $-a = 3a$ *in the base space. Homeomorphisms that map from the open neighborhood* U_1 *at the point* $p \in B$ *will be labeled* $\overset{\triangle}{\varphi}_{1,p}$*, while those that map from* U_2 *at the point* $p \in B$ *will be labeled* $\overset{\triangle}{\varphi}_{2,p}$*.*

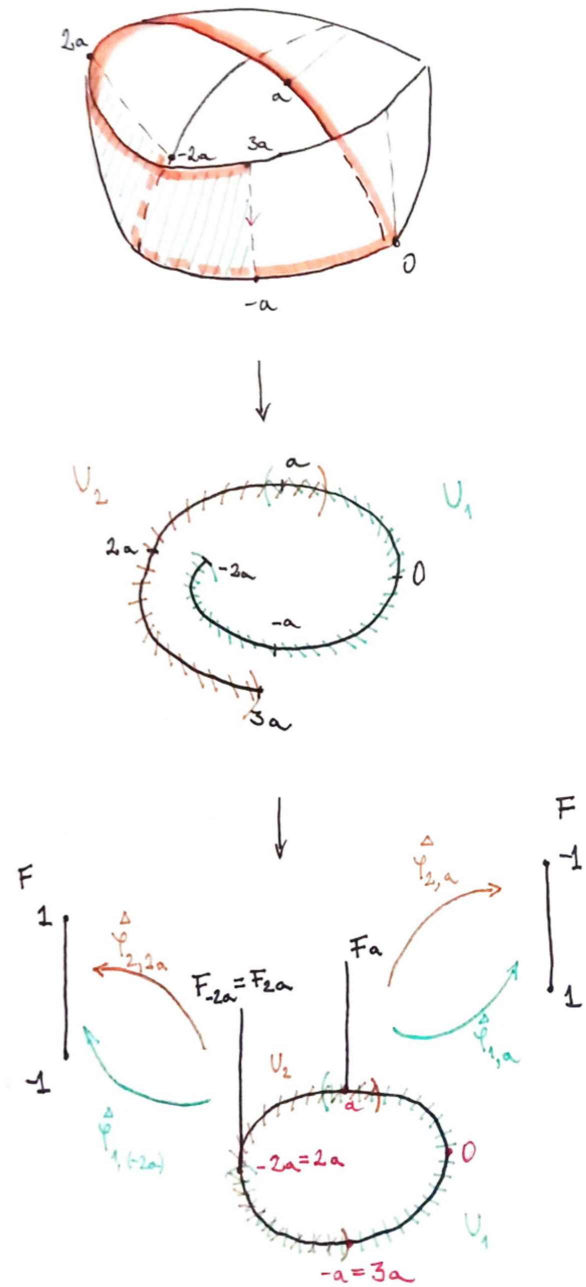


Figure 2.12: Möbius strip.

Figure (2.12) depicts the action of homeomorphic maps. Looking at the mid drawing in the figure (2.11) we can see where each of the points maps to. Looking at the midpoint in the figure (2.12) we can see from which section our points are being mapped. Note that points from $-2a$ to $3a$ only exist in a base space but are shown in all figures for clarity. Looking at the homeomorphisms that act on a fiber $F_{2a} = F_{-2a}$ we can write:

$$\begin{aligned} \overset{\triangle}{\varphi}_{1,(-2a)} :& F_{-2a} \to F \\ -1 &\mapsto -1 \end{aligned} \qquad \textit{and} \qquad \begin{aligned} \overset{\triangle}{\varphi}_{2,2a} :& F_{2a} \to F \\ -1 &\mapsto 1 \end{aligned} \tag{2.6}$$

This means that for the same point $-2a = 2a$ we get two different points in the typical fiber. On the other hand, let us look at the fiber F_a:

$$\begin{aligned} \overset{\triangle}{\varphi}_{1,a} : F_a &\to F \\ 1 &\mapsto 1 \end{aligned} \qquad \textit{and} \qquad \begin{aligned} \overset{\triangle}{\varphi}_{2,a} : F_a &\to F \\ 1 &\mapsto 1 \end{aligned} \tag{2.7}$$

Here, the same point a is mapped to the same point in the typical fiber. The homeomorphic maps $\overset{\triangle}{\varphi}_{1,(-2a)}$ and $\overset{\triangle}{\varphi}_{2,2a}$ tell us that the endpoints switched as they mapped to different values for the same fiber. That is how fibers, or rather homeomorphisms $\overset{\triangle}{\varphi}_{i,p}$ hold information on where the twists are happening.

Other than showing where the twists are happening, such homeomorphisms form a group called a **structural group G** [6, 3]. In the figure (2.10) we can see that:

$$\overset{\triangle}{\varphi}_{1,x} \circ \overset{\triangle}{\varphi}^{-1}{}_{2,x} : F \to F \quad \text{and} \quad \overset{\triangle}{\varphi}_{1,y} \circ \overset{\triangle}{\varphi}^{-1}{}_{2,y} : F \to F. \tag{2.8}$$

The elements $\overset{\triangle}{\varphi}_{1,x} \circ \overset{\triangle}{\varphi}^{-1}{}_{2,x}$ and $\overset{\triangle}{\varphi}_{1,y} \circ \overset{\triangle}{\varphi}^{-1}{}_{2,y}$ from a group. If we now go back to the example (2.1), we can write

$$\overset{\triangle}{\varphi}_{1,(-2a)}(-1) = -1, \quad \overset{\triangle}{\varphi}_{2,2a}(-1) = 1, \quad \overset{\triangle}{\varphi}_{1,a}(1) = 1 \quad \text{and} \quad \overset{\triangle}{\varphi}_{2,a}(1) = 1. \tag{2.9}$$

In the case of a fiber F_a we got $\overset{\triangle}{\varphi}_{1,a}(1) = \overset{\triangle}{\varphi}_{2,a}(1) = 1$. As this holds for any point q from a base space and any value in the fiber i_q, we can write $\overset{\triangle}{\varphi}_{1,q}(i_q) = \overset{\triangle}{\varphi}_{2,q}(i_q)$. As that is the case the element $\overset{\triangle}{\varphi}_{1,q}(i_q) \circ \overset{\triangle}{\varphi^{-1}}_{2,q}(i_q)$ represents the neutral of the group G as it does nothing.[5] For the other two we just have $\overset{\triangle}{\varphi}_{1,p}(i_p) \circ \overset{\triangle}{\varphi^{-1}}_{2,p}(i_p) = g \in G$. Here e and g are elements of a group of all possible permutations containing two elements. Such a group is called a symmetrical group of rank 2 and is labeled S_2 [6].

$$S_2 = \left\{ e = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \; g = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right\}. \tag{2.10}$$

Note that elements e and g from the expression above are written in a notation that does not represent a matrix notation as it has more in common with the table that compares the values. The first row of this "table" represents the starting order, while the 2nd row represents the permutation that occurred as the result of the action of homeomorphisms.

For now, we have explained how to describe manifolds with twists as a bundle. We describe them locally as a trivial bundle and look at their fibers to see where the twisting is happening. Roughly speaking, that is what one thinks of when talking about **fiber bundles**, but a proper formal definition

[5]

$$\overset{\triangle}{\varphi}_{1,q}(i_q) = \overset{\triangle}{\varphi}_{2,q}(i_q) \Longrightarrow \overset{\triangle}{\varphi^{-1}}_{1,q}(i_q) = \overset{\triangle}{\varphi^{-1}}_{2,q}(i_q)$$
$$\Longrightarrow \overset{\triangle}{\varphi}_{1,q}(i_q) \circ \overset{\triangle}{\varphi^{-1}}_{2,q}(i_q) = \underbrace{\overset{\triangle}{\varphi}_{1,q}(i_q) \circ \overset{\triangle}{\varphi^{-1}}_{1,q}(i_q)}_{\substack{\text{This is how neutral} \\ \text{element is defined.}}} = e$$

contains one more thing - **transition functions**. Transition functions are continuous maps such that:

$$\begin{aligned} g_{jk} :& U_j \cap U_k \to G \\ x \mapsto & g_{jk}(x) := \overset{\triangle}{\varphi}_{j,x} \circ \overset{\triangle}{\varphi}^{-1}{}_{k,x}. \end{aligned} \tag{2.11}$$

Transition functions satisfy the relation $g_{ji}(x) = g_{jk}(x)g_{ki}(x)$ [6, 3].

If we were to look at all of what defines a fiber bundle in one figure, it would look something like the following picture.

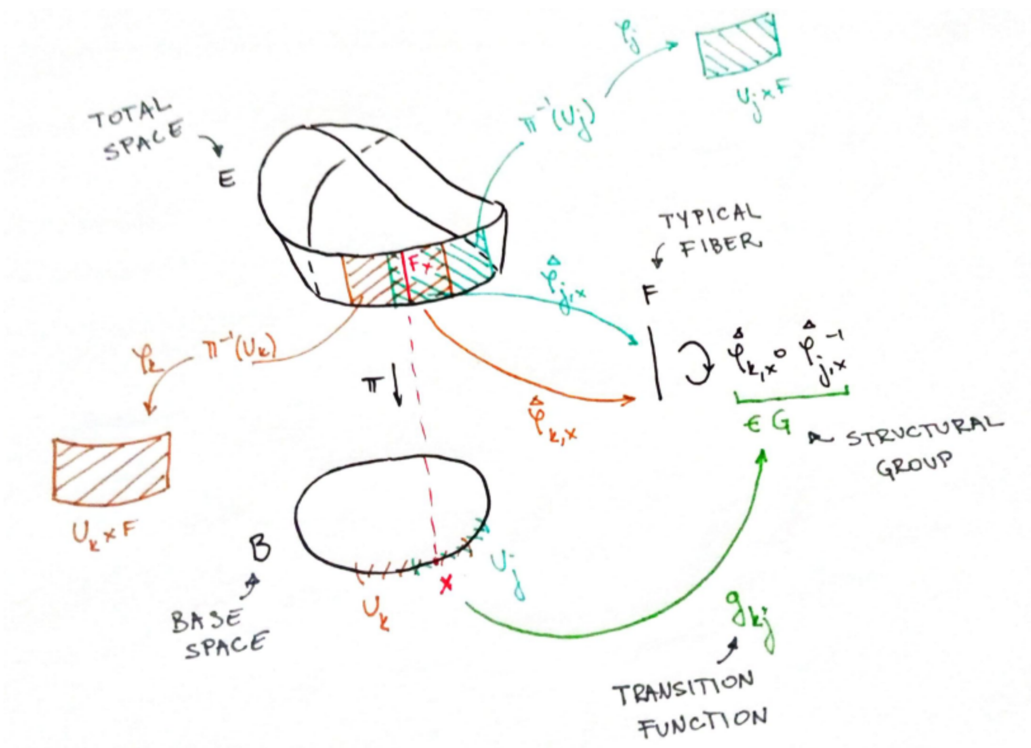


Figure 2.13: Möbius strip as a fiber bundle.

Definition 2.2. *A **fiber bundle** (E, B, π, G) is a bundle (E, B, π) together with:*

(i) a typical fiber F,

(ii) a structural group G of homeomorphisms of F onto itself,

(iii) a family of local trivializations $(\{U_j, \varphi_j\})$ for which all of the open neighborhoods U_j cover the whole manifold M and for which the following holds true: $\pi(\varphi_j^{-1}(x, f)) = x$

(iv) and transition functions $g_{jk} : U_j \cap U_k \to G$.

[6, 3]

Note that sometimes, fiber bundles are defined as bundles whose fibers $\pi^{-1}(x)$ are homeomorphic to the typical fiber F [2, 5]. We labeled such maps $\overset{\triangle}{\varphi}$ and they are not directly in the definition but follow from (ii), as such homeomorphisms are part of a structural group. The property (iii) in such a definition follows from the fact that all fibers are homeomorphic to the same typical fiber F meaning that locally we can still write $(U_i \times F)$. That is what local trivializations also allow us to write. In such cases *local trivializations* are not yet introduced and properties (ii) and (iv) are left out.

2.2.1 Bundle map isomorphism

The following definition and subsequent explanation can be found in [5, 6, 3, 7, 8]. Let (E, M, π) and (E', M', π') be bundles such that all of their manifolds are smooth. Homomorphisms preserve the structure. In the case of such "smooth bundles", the differentiability of those manifolds should be kept while mapping. They should also map fibers to fibers. The "smoothness" of the manifolds is preserved if the maps u and f from the diagram below are set to be smooth, while the fibers are preserved by mapping the same point from one total space onto the same point in the other base manifold. That is achieved by letting the following diagram commute.

$$\begin{array}{ccc} E & \xrightarrow{u} & E' \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi'} \\ M & \xrightarrow[f]{} & M' \end{array}, \tag{2.12}$$

i.e. iff there exist maps u and f such that $\pi' \circ u = f \circ \pi$ then the fibers get mapped into fibers. If smooth maps u and f are also diffeomorphic, then such a homomorphism is an isomorphism, *i.e.*

$$\begin{array}{c} E \\ {\scriptstyle\pi}\big\downarrow \\ M \end{array} \cong_{bundle} \begin{array}{c} E' \\ \big\downarrow{\scriptstyle\pi'} \\ M' \end{array}. \tag{2.13}$$

2.3 Tangent bundle

A specific example of a fiber bundle is a **tangent bundle**. Tangent bundles are important because they allow us to define a tangent space at every point of a base space M. By having tangent spaces at every point of M we have well-defined tangent vectors at every point of M, *i.e.* we can precisely define a **vector field**. If we denote the set of all tangent spaces on a smooth manifold with TM, a tangent bundle can be thought of as a fiber bundle that can locally be described as a trivial bundle whose total space is a product manifold $TM \times M$. $T_pM \in TM$ would be a fiber at $p \in M$ of such a bundle, and M is a total space.

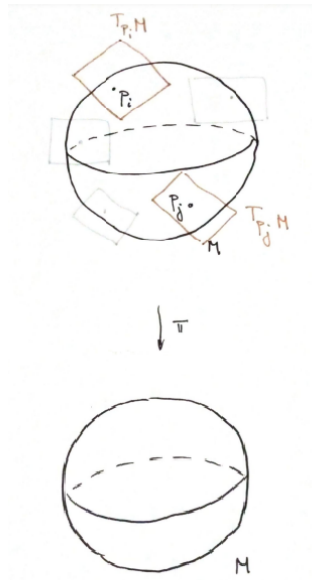


Figure 2.14: Tangent bundle.

Definition 2.3. *Let M be a smooth manifold. Let TM denote the disjoint union $TM := \dot{\bigcup}_{p \in M} T_pM$. A **Tangent bundle** is a set TM together with a projection map*

$$\begin{aligned} \pi : & TM \times M \to M \\ & (x, p) \mapsto p, \end{aligned} \tag{2.14}$$

where $x \in T_pM$ and $p \in M$. The typical fiber of such a bundle is $\mathbb{R}^n$. The structural group of a tangent bundle is $GL(n, \mathbb{R})$. Transition functions then map points $p \in M$ to $GL(n, \mathbb{R})$. Local trivializations $(\{U_j, \varphi_j\})$ describe local sections $\pi^{-1}(U_j)$ of $TM \times M$. [6]

Note that some sources define a tangent bundle as $\pi : TM \to M$ seemingly leaving out the information on total space being a product manifold. However, since each $T_pM \in M$ within itself contains a point $p \in M$ it *can* be thought of as a product of that point p and a tangent space defined at that point [5, 3, 9]. Tu [4, p. 119] explicitly makes this clear by defining a set TM as $TM \coloneqq \dot{\bigcup\limits_{p \in M}} T_pM = \dot{\bigcup\limits_{p \in M}} \{p\} \times T_pM$. We should also note that the sources that left out structure groups, transition maps, and local trivializations from their definitions naturally do not introduce them here either.

2.3.1 Structural group of a tangent bundle

The following explanation relies on [10, 3]. The typical fiber of a tangent bundle is $\mathbb{R}^n$ as it is possible to construct an isomorphism between T_pM and $\mathbb{R}^n$, where $\mathbb{R}^n$ is a real vector space [5]. Structure group G consists of homeomorphisms $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}$. A basis for T_pM are $\left(\frac{\partial}{\partial x^\nu}\right)_p$ and $\left(\frac{\partial}{\partial y^\mu}\right)_p$, depending on a choice of open neighborhoods $\pi^{-1}(U_\nu)$ and $\pi^{-1}(U_\mu) \in TM \times M$. Since the typical fiber of a tangent bundle is isomorphic to the $\mathbb{R}^n$, the typical fiber and $\mathbb{R}^n$ where our charts map to are now the same space. This means that the basis for the typical fiber are then $(x^\nu)_p$ and $(\tilde{x}^\mu)_p$, depending on our choice of charts.

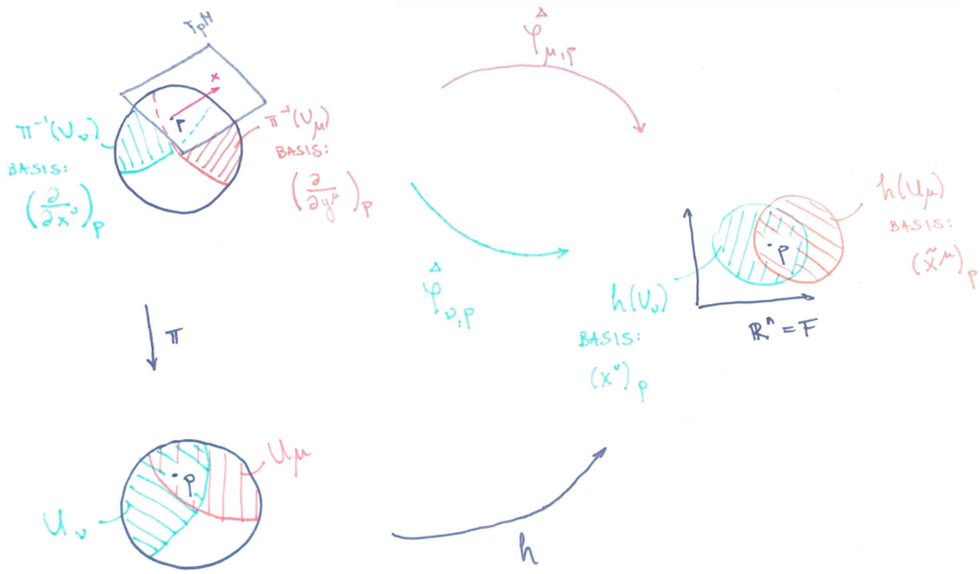


Figure 2.15: Structure group of a tangent bundle.

Vector $x \in T_pM$ can be written as $x = x^\nu \left(\frac{\partial}{\partial x^\nu}\right)_p = \tilde{x}^\mu \left(\frac{\partial}{\partial y^\mu}\right)_p$ depending on a choice of basis. This means that homeomorphisms map in the following fashion.

Let

$$\begin{aligned}&\overset{\triangle}{\varphi}_{\nu,p} : T_pM \to \mathbb{R}^n \\ &\overset{\triangle}{\varphi}_{\nu,p}\left(\left(x^\nu \frac{\partial}{\partial x^\nu}\right)_p\right) = (x^\nu)_p\end{aligned} \quad \text{and} \quad \begin{aligned}&\overset{\triangle}{\varphi}{}^{-1}{}_{\mu,p} : \mathbb{R}^n \to T_pM \\ &\overset{\triangle}{\varphi}{}^{-1}{}_{\mu,p}\left((\tilde{x}^\mu)_p\right) = \left(\tilde{x}^\mu \frac{\partial}{\partial y^\mu}\right)_p\end{aligned} \quad . \tag{2.15}$$

We will calculate $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}{}^{-1}{}_{\mu,p}$, since it belongs to a structural group.

$$\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p} \left((\tilde{x}^{\mu})_p \right) = \overset{\triangle}{\varphi}_{\nu,p} \left(\tilde{x}^{\mu} \frac{\partial x^{\nu}}{\partial y^{\mu}} \frac{\partial}{\partial x^{\nu}} \right)_p = \left(\tilde{x}^{\mu} \frac{\partial x^{\nu}}{\partial y^{\mu}} \right)_p$$
$$\iff \overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p} = \left(\frac{\partial x^{\nu}}{\partial y^{\mu}} \right)_p \tag{2.16}$$

We also know that $x = x^{\nu} \left(\frac{\partial}{\partial x^{\nu}} \right)_p = \tilde{x}^{\mu} \left(\frac{\partial}{\partial y^{\mu}} \right)_p$. This can be rewritten in a way that shows the relation between vector components from both bases.

$$x^{\nu} \left(\frac{\partial}{\partial x^{\nu}} \right)_p = \tilde{x}^{\mu} \left(\frac{\partial x^{\nu}}{\partial y^{\mu}} \frac{\partial}{\partial x^{\nu}} \right)_p$$
$$x^{\nu} = \tilde{x}^{\mu} \left(\frac{\partial x^{\nu}}{\partial y^{\mu}} \right)_p \tag{2.17}$$

Written like this, it is clear that the element of a structural group of a tangent bundle $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}$ represents the transition matrix $\left(\frac{\partial x^{\nu}}{\partial y^{\mu}} \right)_p$ between the two bases. Since the bases we started with are chosen and not given, we can label a transition matrix in a more general manner: $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p} := G^{\nu}_{\mu}$. In general, any such G^{ν}_{μ} is the element of a structural group called a **general linear group** of degree n. Such a group is labeled $GL(n, \mathbb{R})$ and it contains $n \times n$ invertible matrices, which should not be surprising considering that the maps $\overset{\triangle}{\varphi}_{\nu,p}$ and $\overset{\triangle}{\varphi}_{\mu,p}$ are both homeomorphic, and hence invertible by definition. No matter the tangent bundle, $GL(n, \mathbb{R})$ is always its structure group. Here, the kind of "*twisting*" that homeomorphisms in $GL(n, \mathbb{R})$ describe, is not a physical twist of a total space, but rather the change of basis.

2.3.2 Vector field

Now that we have a tangent space at every point of our manifold we can start describing a *vector field*. However, note that in every $T_pM \in TM$ we have more than one vector, pointing in all possible directions. To accurately describe a vector field, we only need one vector from each of the tangent spaces T_pM. To be able to pick out one specific vector, the **cross-section** is introduced.

Because it is geometrically easier to understand, we will first look at the cylinder and its cross-section. The cross-section is a map that maps points from the base space B to the total space E, but, unlike π^{-1} which maps the point $x \in B$ to the whole fiber $\pi^{-1}(x)$, the cross-section σ maps the point $y \in B$ to a single point $\sigma(y) \in \pi^{-1}(y)$. This is represented in figure (2.16).

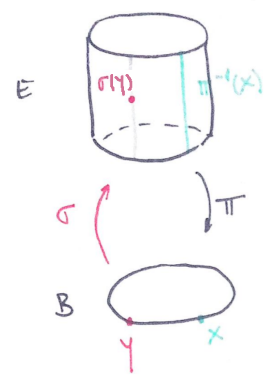


Figure 2.16: Cross-section of a cylinder at a point $y \in B$.

Since we want this point $\sigma(y)$ to project to the same y we started with, by definition we set $\pi(\sigma(y)) = y$, *i.e.* $\pi \circ \sigma = id_B$. If we were to look at a cross-section of a whole base space, we would get one point in each fiber, as shown in figure (2.17).

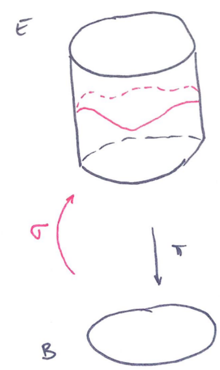


Figure 2.17: Cross-section of a whole base space of a cylinder.

Definition 2.4. *A **cross-section** of a bundle (E, B, π) is a map $\sigma : B \to E$ for which every $\sigma(p)$ lies in the fiber $\pi^{-1}(p)$, i.e. $\pi \circ \sigma = id_B$.*
[6, 5, 4]

Similar to the cylinder, a cross-section of a tangent bundle is not the whole fiber $T_pM \in TM$ but rather one element of T_pM, *i.e.* one vector from T_pM. This sentiment is represented in figure (2.18). The vectors that σ mapped to are colored red, while the rest of the vectors that live in a tangent space are gray. Similar to before $\pi \circ \sigma = id_B$.

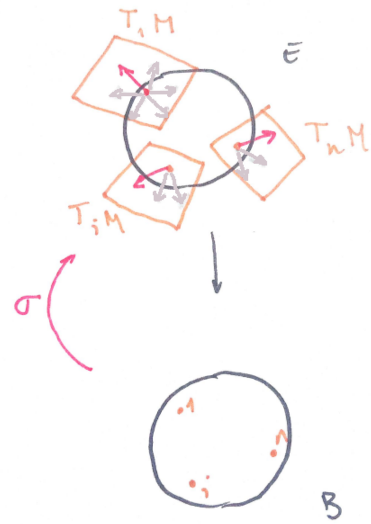


Figure 2.18: Cross-section of a tangent bundle.

Notice how that is *exactly* what we wanted for our definition of a vector field - a well-defined vector for each point $p \in M$. The cross-section of a tangent bundle is then what is considered a **vector field**[6].

Definition 2.5. *Let M be a smooth manifold and let $(TM \times M, M, \pi)$ be its tangent bundle. A* ***vector field* X** *is a smooth section of a tangent bundle such that $\pi \circ \sigma = id_M$. [6, 5, 4, 3, 11]*

Note that both π and σ are smooth maps. Sometimes a vector field is defined without mention of a cross-section. In such definitions, it is simply defined as a *smooth assignment of a vector to each point* $p \in M$ [5].

As σ can map to any vector in fiber, the vector field of a given manifold is in no way unique. The set of all vector fields on a given manifold is denoted by $\Gamma(TM) := \{\sigma : M \to TM \times M \mid \pi \circ \sigma = id_M\}$.

[6] Note that "*field*" in the name does not represent an algebraic field, but is more akin to the ideas of electric or magnetic *fields*, as in something that is defined at each point in a given space.

2.4 Cotangent bundle

Similar to the idea of a tangent bundle, we can give each $p \in M$ its own cotangent space. By doing that we have defined a **cotangent bundle**.

Definition 2.6. *Let M be a smooth manifold. Let $(TM)^*$ denote the disjoint union $(TM)^* := \dot{\bigcup}_{p \in M} (T_pM)^*$. A **Cotangent bundle** is a set $(TM)^*$ together with a projection map*

$$\begin{aligned} \pi : (TM)^* \times M &\to M \\ (\omega, p) &\mapsto p, \end{aligned} \tag{2.18}$$

where $\omega \in (T_pM)^*$ and $p \in M$. The typical fiber of such a bundle is $\mathbb{R}^n$. The structural group of a tangent bundle is $GL(n, \mathbb{R})$. Transition functions then map points $p \in M$ to $GL(n, \mathbb{R})$. Local trivializations $(\{U_j, \varphi_j\})$ describe local sections $\pi^{-1}(U_j)$ of $(TM)^* \times M$. [6, 3]

An alternative definition of a cotangent bundle can be summed up as $\pi : (TM)^* \to M$ [5, 4].

2.4.1 Structural group of a cotangent bundle

The following explanation relies on [10, 3]. As was the case with the tangent bundle, we can look at what homeomorphisms in G are twisting, and how they act. Let $(dx^\nu)_p$ and $(dy^\mu)_p$ be a basis of $(T_pM)^*$, depending on a choice of open neighborhoods $\pi^{-1}(U_\nu)$ and $\pi^{-1}(U_\mu) \in (TM)^* \times M$. The cotangent space is isomorphic to the tangent space, *i.e.* $T_pM \cong (T_pM)^*$ meaning that the typical fiber of a cotangent bundle is also isomorphic to the $\mathbb{R}^n$ in the same manner that tangent bundle was. A typical fiber and $\mathbb{R}^n$ where our charts map to are now the same space. This means that the basis for a typical fiber are then $(\omega_\nu)_p$ and $(\tilde{\omega}_\mu)_p$, depending on our choice of charts.

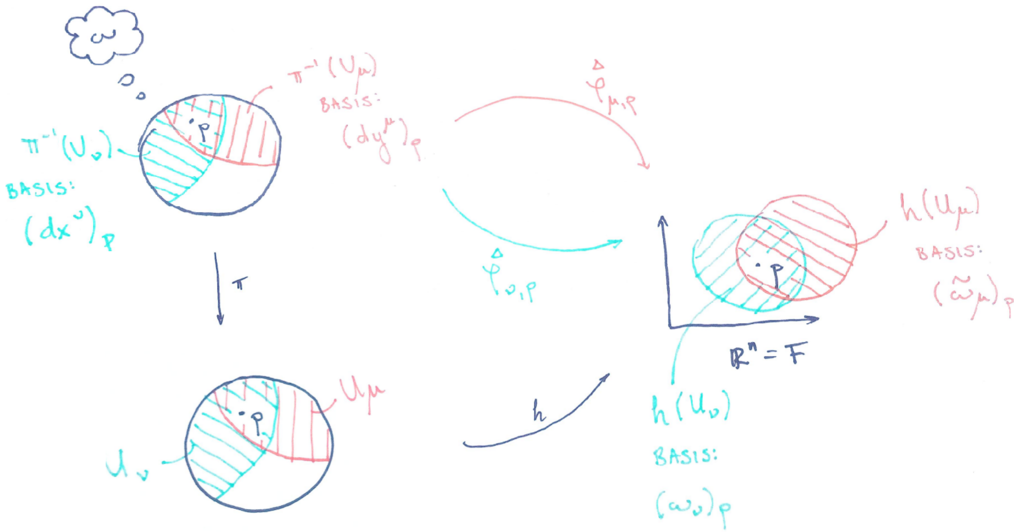


Figure 2.19: Structural group of a cotangent bundle.

The n-form $\omega \in (T_pM)^*$ can be written as $\omega = \omega_\nu (dx^\nu)_p = \tilde{\omega}_\mu (dy^\mu)_p$ depending on a choice of basis. This means that homeomorphisms map in the following fashion:

Let

$$\begin{array}{lcl} \overset{\triangle}{\varphi}_{\nu,p} : (T_pM)^* \to \mathbb{R}^n & & \overset{\triangle}{\varphi}^{-1}{}_{\mu,p} : \mathbb{R}^n \to (T_pM)^* \\ \overset{\triangle}{\varphi}_{\nu,p}\left(\omega_\nu \left(dx^\nu\right)_p\right) = \left(\omega_\nu\right)_p & \text{and} & \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}\left(\left(\tilde{\omega}_\mu\right)_p\right) = \left(\tilde{\omega}_\mu \left(dy^\mu\right)\right)_p . \end{array} . \tag{2.19}$$

We will calculate $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}$, since it belongs to a structural group.

$$\begin{aligned} \overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}\left(\left(\tilde{\omega}_\mu\right)_p\right) &= \overset{\triangle}{\varphi}_{\nu,p}\left(\tilde{\omega}_\mu \frac{\partial y^\mu}{\partial x^\nu} dx^\nu\right)_p = \left(\tilde{\omega}_\mu \frac{\partial y^\mu}{\partial x^\nu}\right)_p \\ \iff \overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p} &= \left(\frac{\partial y^\mu}{\partial x^\nu}\right)_p \end{aligned} \tag{2.20}$$

The equation $\omega = \omega_\nu (dx^\nu)_p = \tilde{\omega}_\mu (dy^\mu)_p$ can be rewritten so we get the relation between components from both bases.

$$\begin{aligned} \omega_\nu \left(dx^\nu\right)_p &= \tilde{\omega}_\mu \left(\frac{\partial y^\mu}{\partial x^\nu} dx^\nu\right)_p \\ \omega_\nu &= \tilde{\omega}_\mu \left(\frac{\partial y^\mu}{\partial x^\nu}\right)_p \end{aligned} \tag{2.21}$$

Similar to before, the transition matrix $\left(\frac{\partial y^\mu}{\partial x^\nu}\right)_p$ is precisely what our homeomorphisms $\overset{\triangle}{\varphi}_{\nu,p} \circ \overset{\triangle}{\varphi}^{-1}{}_{\mu,p}$ are. Same as before, they are part of a $GL(n, \mathbb{R})$ group. Also same as before, the "*twisting*" such homeomorphisms describe is just the change of basis in a cotangent bundle.

2.4.2 Covector field and an n-form field

Analogous to the case of the tangent bundle, a cross-section of a cotangent bundle is a **covector field**. This time a section σ maps a point $p \in M$ to a 1-form $\omega \in (T_pM)^*$.

Definition 2.7. *Let M be a smooth manifold and let $((TM)^* \times M, M, \pi)$ be its cotangent bundle. A* ***covector field****,* ***covariant vector field****,* ***differential 1-form*** *or simply* ***1-form*** $\boldsymbol{\omega}$ *is a smooth section of a cotangent bundle such that $\pi \circ \sigma = id_M$ [6, 5, 4, 12].* [7]

Analogous to how 1-forms or covectors act on vectors, covector fields act on vector fields. Because of that, they can be defined in terms of their action on vector fields. The name *1-form* in such a context makes more sense, as they act identically to 1-forms with the only difference being they act on vector fields as opposed to vectors. Similar to the set of all vector fields, we get a set of all covector fields that will be denoted by $\Gamma(TM)^* \coloneqq \{\sigma : M \to (TM)^* \times M \mid \pi \circ \sigma = id_M\}$.

[7] Note that both π and σ are smooth maps. Analogous to the case of vector fields, sometimes covector field is simply defined as the assignment of a 1-form to each $p \in M$.

At this point, we only have a 1-form defined at every point $p \in M$. We have yet to define an n-form at every $p \in M$. Similar to how we did not need to define a cotangent bundle to define a covector field, an "*n-form field*" or a *differential n-form* can be defined in terms of its action on n vector fields.

Definition 2.8. *Let M be a smooth manifold and let $0 \leq n \leq dim\ M$. A* ***(differential) n-form*** *is the following map*

$$\omega \underbrace{(X_1, .., X_n)}_{\text{Those are now vector fields.}} := sign(\sigma)\omega \underbrace{(X_{\sigma(1)}, .., X_{\sigma(n)})}_{\text{Those are now vector fields.}}. \tag{2.22}$$

[13, 4]

The space of all "n-form fields" defined on M shall be denoted $\Omega^n(M)$.

2.4.3 Exterior derivative operator

One useful notion to mention here is the idea of the **exterior derivative operator**. One can think of the exterior derivative operator as a generalized "*total derivative field*". To justify the wording "*generalized*", we will first introduce the idea of a **gradient operator** to make it clearer what exactly will be generalized upon. To do that, we need a notion of a *0-form* or a 0-form field. By definition, 0-forms are set to be maps $f : M \to \mathbb{R}$. This function might look familiar as that is exactly how $C^\infty(M)$ functions are defined. Gradient operators are not usually introduced separately as they are special cases of the aforementioned exterior derivative operator. However, they will be defined separately here since they will be needed in the later part of this work. The terminology used comes from Schuller's lectures [9, 13].

Definition 2.9. *Let $f \in C^\infty(M)$. Let $\Omega^0(M) \coloneqq C^\infty(M)$ denote the set of all the 0-forms, and $\Omega^1(M) \coloneqq (T_pM)^*$ the set of all the n-forms. Then at every point $p \in M$ we define a linear map*

$$\begin{aligned} d_p : \Omega^0(M) &\xrightarrow{\sim} \Omega^1(M) \\ f &\mapsto d_p f \end{aligned} \tag{2.23}$$

where,

$$(d_p f)X \coloneqq Xf. \tag{2.24}$$

The map d_p is called a ***gradient operator*** *at a point $p \in M$. d_pf, also considered as a map, is then called a* ***gradient*** *of f at p.*
[9, 13]

The generalization that occurs for the case of the exterior derivative operator can be summed up in the map $d : \Omega^n(M) \xrightarrow{\sim} \Omega^{n+1}(M)$.

Definition 2.10. *The* ***exterior derivative operator d*** *is a map*

$$d : \Omega^n(M) \xrightarrow{\sim} \Omega^{n+1}(M) \tag{2.25}$$

$$\begin{aligned} (d\omega)\underbrace{(X_1, \ldots, X_{n+1})}_{\text{These are now vector fields.}} \coloneqq{}& \sum_{i=1}^{n+1} (-1)^{i+1} X_i \left(\omega\left(X_1, \ldots, \cancel{X_i}, \ldots, X_{n+1}\right)\right) \\ &+ \sum_{i<j} (-1)^{i+j} \omega\left([X_i, X_j], X_1, \ldots, \cancel{X_i}, \ldots, \cancel{X_j}, \ldots, X_{n+1}\right), \end{aligned}$$

where $[X_i, X_j]$ is a commutator between two vector fields. It is defined as follows.

$$[X, Y]f = X(Yf) - Y(Xf), \tag{2.26}$$

where X, Y are two vector fields, and $f \in C^\infty(M)$.
[13]

§3 Lie groups and their Lie algebras

Groups represent what is considered a *symmetry* and *Lie groups* describe *continuous symmetries*. One can think of symmetry as a *refusal to change under a certain action* or, put into more technical terms, symmetry represents *invariance under transformations*. If one looks at properties that define a group, they all represent actions that do not change the system. An excellent explanation of such an outlook on groups is given in Schwichtenberg's Physics from Symmetry [14, Chapter 3.1]. As (continuous) symmetries are an integral part of many physical theories, Lie groups make a frequent appearance in the field. The study of fiber bundles and connections is no exception to this.

The main reason for the introduction of Lie groups is their direct correlation with *principal fiber bundles* whose direct use is to be a platform on which one defines a *connection* - a very important term that can informally be summed up as a structure which allows for the connection of fibers defined at different points. From a physical perspective, Yang-Mills fields (or more generally, gauge potentials) take on values directly from Lie groups, *i.e.* one can think of gauge potentials as being *Lie algebra-valued* connection 1-forms. The exact meaning of the last statement will be explained in chapter (6.1).

In the context of Lie groups, it is helpful to think of continuity as the opposite of discreteness. Such an outlook allows one to think of a neighborhood around a given point that consists of infinitely many members. This setup then allows for the existence of infinitely many symmetry transformations. A famous example of such symmetry would be a unit circle with symmetry transformation, or rather a group operation being rotation. Such a group even gets a special name $U(1)$ and is related to an aforementioned $GL(n, \mathbb{R})$ which is also considered a Lie group. If one were to rotate a circle around the origin we would not be able to distinguish between a rotated circle and the starting circle. Because the angle of rotation does not affect that outcome, there are infinitely many rotations one could perform that do not affect the circle in any recognizable way. That is why such a circle is considered to have infinitely many symmetry transformations and why $U(1)$ is considered a Lie group.

As there are more requirements to be met for the case of continuous symmetries as opposed to non-continuous counterparts, Lie groups not only have to be groups, but they should also have smooth maps that allow one to describe such a behavior. Namely, the group operation $\bullet$ and the inverse of each element should be the results of a smooth map. As smooth maps are defined on smooth manifolds, the group set G is also a smooth manifold.

Definition 3.1. *A **Lie group** $(G, \bullet)$ is:*

(i) a group $(G, \bullet)$

(ii) a smooth manifold G together with the following maps being smooth:

(i) $\mu : G \times G \to G$
$(g_1, g_2) \mapsto g_1 \bullet g_2$

(ii) $i : G \to G$
$g \mapsto g^{-1}$.

[15, 5, 4, 3]

It is possible to write these two maps as one map that maps in the following manner $(g_1, g_2) \mapsto g_1 \bullet g_2^{-1}$, so some sources define them as such [6, 12].

An arbitrary Lie group does not necessarily commute, so, in general, there is a difference between $g_1 \bullet g_2$ and $g_2 \bullet g_1$. As for non-commutative groups, one should distinguish between left and right. The product $\bullet$ can be described in two ways - "taking the product of" g_1 from the left, *i.e.* $g_1 \bullet g_2$ and "taking the product of" g_1 from the right, *i.e.* $g_2 \bullet g_1$. As we only need one of these, in the case of Lie algebra the pick is left multiplication and it is called **left translation**. Note that it is of course possible to define a *right translation* in the analogous fashion.

Definition 3.2. *Let $(G, \bullet)$ be a Lie group. Then, for any $g \in G$ we define a* ***left translation*** *with respect to g as a map*

$$\begin{aligned} l_g &: G \to G \\ h &\mapsto l_g(h) := g \bullet h. \end{aligned} \tag{3.1}$$

[5, 12, 3, 15]

The **left translation is a diffeomorphism on** G. From the definition of a Lie group, we know that $\bullet$ and i are both smooth maps. l_g is defined using $\bullet$ so it is by definition a smooth map *i.e.* l_g is a C^∞- diffeomorphism.

The set of all vector fields $\Gamma(TM)$ defined on M together with a Lie bracket is a Lie algebra. The following theorem connects such a set and the set of all left-invariant vector fields.

Theorem 3.1. *Let $L(G)$ denote the set of all left-invariant vector fields on a Lie group G. $(L(G), [\_, \_])$ is a Lie subalgebra of $(\Gamma(TM), [\_, \_])$. [15]*

Proof. To prove this statement one needs to show that the product of two left-invariant vector fields is again a left-invariant vector field, *i.e.* $[\_, \_] : L(G) \times L(G) \to L(G)$.

Let $X, Y \in L(G)$.

$$[X,Y](f \circ l_g) = X(Y(f \circ l_g)) - Y(X(f \circ l_g))$$

$$= X((Yf) \circ l_g) - Y((Xf) \circ l_g) = (X(Yf) - Y(Xf)) \circ l_g$$

As $X, Y \in L(G)$ we know that $X(f \circ l_g) = (Xf) \circ l_g$.

$$= ([X,Y]f) \circ l_g \tag{3.2}$$

The equation $[X,Y](f \circ l_g) = ([X,Y]f) \circ l_g$ is precisely the 3rd equivalent statement of those that define left-invariant vector fields. □

Not only is a $L(G)$ a Lie (sub)algebra but it can be shown that there exists an isomorphism between $L(G)$ and T_eG. This means that the set of all left-invariant vector fields can be regarded as a set of all tangent vectors at the identity of a Lie group G.

Theorem 3.2. $L(G) \underset{\substack{vec.\\space}}{\cong} T_eG$. *[15, 5, 3, 7, 6]*

Proof. The way we are going to prove this statement is by constructing a linear isomorphism j between the spaces. Let $A \in T_eG$.

$$\begin{aligned} j : T_eG &\to L(G) \\ A &\mapsto j(A). \end{aligned} \tag{3.3}$$

As $j(A)$ is supposed to be a left-invariant vector field, we define it as

$$j(A)_g \coloneqq l_{g*}A, \ \ \forall g \in G. \tag{3.4}$$

We need to show that $j(A)$ *is* indeed left-invariant, that it is linear, and that it is bijective.

(i) $j(A)$ is left-invariant.

$$l_{h*}(j(A)_g) = l_{h*}(l_{g*}A) = l_{hg*}A = j(A)_{hg} \tag{3.5}$$

This satisfies the 2nd condition of left-invariant vector fields, *i.e.* $l_{g*}X_h = X_{gh}$. $j(A)$ really is left-invariant.

(ii) $j(A)$ is linear. As $j(A) \coloneqq l_{g*}A$ and l_{g*} is a liner map, so is $j(A)$.

(iii) $j(A)$ is injective.

Let $j(A) = j(B)$. For $\forall g \in G$ it then follows

$$j(A)_g = j(B)_g. \tag{3.6}$$

As this statement holds true for all $g \in G$, it holds true for $g = e$ as well.

$$\begin{aligned} j(A)_e = j(B)_e &\iff l_{e*}A = l_{e*}B \\ &\iff A = B \end{aligned} \tag{3.7}$$

There is a one-to-one correspondence between the domain and a codomain of j.

(iv) $j(A)$ is surjective.

Let $X \in L(G)$. We define $A^X := X_e \in T_eG$.

$$\begin{aligned} j(A^X)_g = l_{g*}A^X = l_{g*}X_e = X_{ge} = X_g \\ \implies j(A^X) = X. \end{aligned} \tag{3.8}$$

This means that arbitrary $A^X \in T_eG$ gets mapped to arbitrary $X \in L(G)$.

$$\implies L(G) \underset{\substack{vec.\\ space}}{\cong} T_eG. \tag{3.9}$$

□

What follows from this theorem is that, geometrically, one can think of a Lie algebra as a tangent to a Lie group. This statement holds true in general, and not just for left-invariant Lie algebras.

Also, because $L(G) \cong T_eG \implies dim(L(G)) = dim(T_eG)$, and from the definition of a manifold we know that $dim(T_eG) = dim(G)$. So, we have $dim(L(G)) = dim(T_eG) = dim(G)$.[8]

[8] Note that although for arbitrary groups the idea of dimension is not well defined, in the case of a Lie group, G is both a group and a manifold. Thus, it makes sense to talk about its dimensionality.

3.1 Exponential maps and one-parameter subgroups

The relationship between Lie groups and their algebras comes in the form of an *exponential map*. However, before the formal definition of such a map, we will first explore the nature of such a map on an example of a Lie group $SO(2)$ and its Lie algebra $\mathfrak{so}(2)$.

Example 3.1. *(Lie groups $O(2)$ and $SO(2)$) [16, 17, 18]*

Let

$$O(2) = \{R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \mid \ RR^T = R^T R = I\} \tag{3.10}$$

Such a set forms a Lie group under matrix multiplication. The group properties are easily checked and both matrix multiplication and transposition are smooth so such a set really is a Lie group.[9] *The whole of $O(2)$ can be divided into two subsets - a subset with condition $\det R = 1$ and a subset with condition $\det R = -1$ and that holds true for all $R \in O(2)$.*[10]

Let

$$S'O(2) \coloneqq \{R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \mid \ RR^T = R^T R = I \text{ and } \det R = -1.\} \quad \text{and} \tag{3.11}$$

$$SO(2) \coloneqq \{R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \mid \ RR^T = R^T R = I \text{ and } \det R = 1.\}. \tag{3.12}$$

We can obtain the form of an arbitrary orthogonal matrix only using the

[9] Arbitrary entry of a product matrix AB can be written as $(AB)_{ij} = \Sigma_k a_{ik} b_{kj}$. The entries of a product matrix are thus considered polynomials. As any arbitrary polynomial is infinitely many times differentiable it is always a smooth map and so, by extension, product matrix entries are smooth. Similar argumentation can be set for a transposition\inversion of a matrix, as from properties of a group, they can always be obtained through some product that we know is smooth.

[10] To check that holds true one only needs to take a determinant of both sides of equation $RR^T = I$. Note that the following holds true for all R: $\det R = \det R^T$.

orthogonality condition.

Let $R \in O(2)$*.*

$$\begin{aligned} RR^T &= \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} r_{11} & r_{21} \\ r_{12} & r_{22} \end{pmatrix} \\ &= \begin{pmatrix} r_{11}^2 + r_{12}^2 & r_{11}r_{21} + r_{12}r_{22} \\ r_{11}a_{21} + r_{12}a_{22} & r_{21}^2 + r_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{3.13}$$

$$\iff r_{11}^2 + r_{12}^2 = r_{21}^2 + r_{22}^2 = 1 \text{ and } r_{11}r_{21} + r_{12}r_{22} = 0. \tag{3.14}$$

$$\begin{aligned} R^T R &= \begin{pmatrix} r_{11} & r_{21} \\ r_{12} & r_{22} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \\ &= \begin{pmatrix} r_{11}^2 + r_{21}^2 & r_{11}r_{12} + r_{21}r_{22} \\ r_{11}r_{12} + r_{21}r_{22} & r_{12}^2 + r_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{3.15}$$

$$\iff r_{11}^2 + r_{21}^2 = r_{12}^2 + r_{22}^2 = 1 \text{ and } r_{11}r_{12} + r_{21}r_{22} = 0. \tag{3.16}$$

From these, we can equate the following

$$r_{11}^2 + r_{12}^2 = r_{11}^2 + r_{21}^2 \Longrightarrow r_{12}^2 = r_{21}^2 \tag{3.17}$$

$$r_{11}^2 + r_{12}^2 = r_{12}^2 + r_{22}^2 \Longrightarrow r_{11}^2 = r_{22}^2 \tag{3.18}$$

As we got a quadratic equation we get two solutions. If we plug the solution $r_{12}^2 = r_{21}^2$ *into 2nd equation of (3.16) (or (3.14)) we get* $r_{11}^2 = -r_{22}^2$*. This means we can write our starting matrix* R *as*

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{12} & -r_{11} \end{pmatrix} \Longrightarrow \det R = -1. \tag{3.19}$$

Similarly, for solution $r_{12}^2 = -r_{21}^2$ *we get* $r_{11}^2 = r_{22}^2$ *and we can write*

$$\begin{pmatrix} r_{11} & r_{12} \\ -r_{12} & r_{11} \end{pmatrix} \Longrightarrow \det R = 1. \tag{3.20}$$

Here, it is clear that the arbitrary matrix form for $R \in O(n)$ *is one of the two matrices we got, each corresponding to either condition* $\det R = 1$ *or* $\det R = -1$. *This means we could write our sets as the following*

$$S'O(2) = \{\textit{All matrices of the form } \begin{pmatrix} r_{11} & r_{12} \\ r_{12} & -r_{11} \end{pmatrix}\} \quad \textit{and} \tag{3.21}$$

$$SO(2) = \{\textit{All matrices of the form } \begin{pmatrix} r_{11} & r_{12} \\ -r_{12} & r_{11} \end{pmatrix}\}. \tag{3.22}$$

One possible representation of such sets is for $r_{11} = \cos\theta$ *and* $r_{12} = -\sin\theta$. *This means that now, any element of* $S'O(2)$ *can be represented with a matrix* $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ *and any element of* $SO(2)$ *can be represented with a matrix* $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. *Because, if one were to act on an arbitrary vector with a matrix* $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$, *such a matrix would mirror that vector across the line at an angle* $\frac{\theta}{2}$, *so such matrices are considered reflection matrices. Similarly, if one were to act on an arbitrary vector with* $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ *such vector would get rotated by the angle* θ *so such matrix is considered rotation matrix.*

Note that the set $S'O(2)$ *is not closed under matrix multiplication and so, it does not form a group.*[11] *On the other hand, the set* $SO(2)$ *does form a*

[11] Let A, B be in such a subset. By definition, their respective determinants are equal to -1. The product determinant is then always equal to $det(AB) = (-1)(-1) = 1$ and, by definition, elements with $\det = 1$ are not contained in such a set.

Lie group under matrix multiplication. As the set $S'O(2)$ does not form a Lie group, we will be looking at only $SO(2)$ from now on.

As $SO(2)$ is both a group and a manifold, we can talk about its shape as a manifold. If one were to act on an arbitrary vector in $\mathbb{R}^2$ with all possible values of $\theta \in \mathbb{R}$, the outputs one would get would trace a circle. That is the reason why the $SO(2)$, considered as a manifold, can be visualized as a circle with the center of a circle being the origin. As $SO(2)$ is a set consisting of all rotation matrices and as one can think of it as a circle, note that all points on such a circle are represented with a 2×2 rotation matrices. That ***does not*** *imply that $SO(2)$ is two-dimensional, as a dimension of a manifold is determined by a dimension of a locally Euclidean tangent space, which is, in the case of a circle, a one-dimensional line or $\mathbb{R}$. The interpretation of matrix columns being vectors does not apply to this case as $SO(2)$ is only a group and not a vector space.*

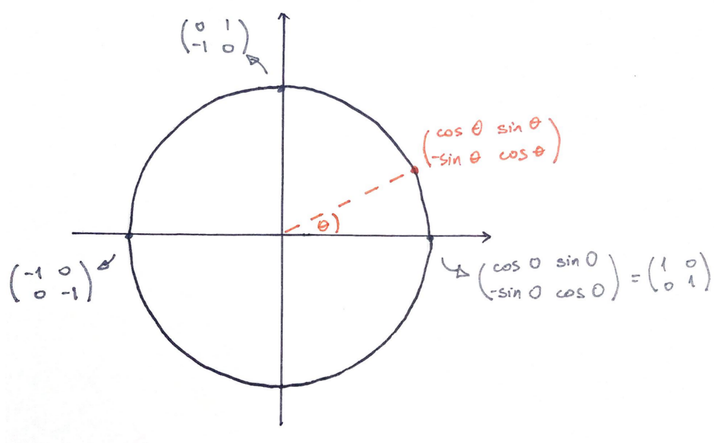


Figure 3.1: $SO(2)$ as a manifold.

Even though we have written $SO(2)$ as a rotation matrix, it is possible to write it using an *exponential map*. In fact, for any $R(t) \in G$ we can write $R(t) = e^{ta} \in G$, where $a = \left.\frac{dR(t)}{dt}\right|_{t=0}$. Such a is then an element of a Lie algebra $\mathfrak{g}$ and it is *tangent* to G at identity.[12] The exponential map gives us the link between Lie algebra and its Lie group. As $R(t)$ can be represented with matrices it should be noted that the exponential function is generalized to matrices. To understand how we should define an exponential map. Note that the definition below applies to matrix groups and this definition will be further generalized upon. Any group whose elements can be written in a matrix form is a subgroup of the aforementioned $GL(n, \mathbb{R}) = \{M \in M_{n\times n}(\mathbb{R}) \mid M \text{ is invertible, } i.e.\ det(M) \neq 0.\}$.[13]

Definition 3.3. *Let $M \in G$, where $G \subseteq GL(n, \mathbb{R})$. Then exponential map is the following map*

$$exp(M) = I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + ... \tag{3.23}$$

[19, 4, 12, 20, 7]

Now, if we come back to our example of $SO(2)$ we can find its Lie algebra $\mathfrak{so}(2)$.

Example 3.2. *[19] We have shown that any element of $SO(2)$ can be represented with a rotation matrix, and we can write it as*

$$R(\theta) = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}. \tag{3.24}$$

We can obtain a by $a = \left.\frac{dR(\theta)}{d\theta}\right|_{\theta=0}$.

[12]Illustrative "proof" of this statement can be shown in the example $\mathfrak{g} = L(G)$. As $T_eG \cong L(G)$ we know $L(G)$ is tangent to G.

[13]$GL(n, \mathbb{R})$ is a Lie group under matrix multiplication and one can use similar argumentation as was used to $O(2)$ to prove that. In fact, $O(2)$ is a subgroup of $GL(n, \mathbb{R})$.

$$a = \frac{dR(\theta)}{d\theta}\bigg|_{\theta=0} = \begin{pmatrix} -sin\theta & cos\theta \\ -cos\theta & -sin\theta \end{pmatrix}\bigg|_{\theta=0} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.25}$$

To show that $e^{\theta a}$ really does equal to rotation matrix we just need to follow the definition of an exponential map.

$$\begin{aligned} exp(\theta a) &= I + \theta a + \frac{1}{2!}(\theta a)^2 + \frac{1}{3!}(\theta a)^3 + \frac{1}{4!}(\theta a)^4 + \frac{1}{5!}(\theta a)^5 + ... \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -\theta^2 & 0 \\ 0 & -\theta^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & -\theta^3 \\ \theta^3 & 0 \end{pmatrix} \\ &+ \frac{1}{4!}\begin{pmatrix} \theta^4 & 0 \\ 0 & \theta^4 \end{pmatrix} + \frac{1}{5!}\begin{pmatrix} 0 & \theta^5 \\ -\theta^5 & 0 \end{pmatrix} + ... \\ &= \begin{pmatrix} 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + ... & \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + ... \\ -\theta + \frac{\theta^3}{3!} - \frac{\theta^5}{5!} + ... & 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + ... \end{pmatrix} \\ &= \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} = R(\theta) \end{aligned} \tag{3.26}$$

Both the Lie group $SO(2)$ and its Lie algebra $\mathfrak{so}(2)$ can be visualised in the following way

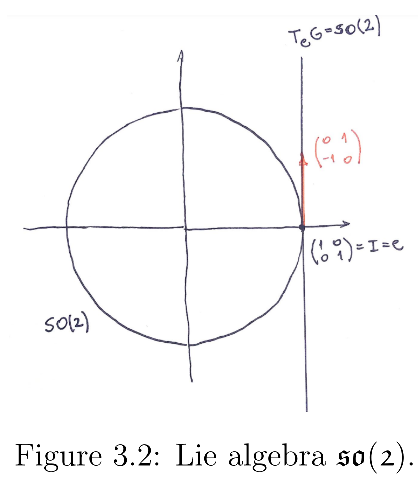


Figure 3.2: Lie algebra $\mathfrak{so}(2)$.

Note that similar to the case of visualizing points in $SO(2)$ as a group, here, Lie algebra $\mathfrak{so}(2)$ is spanned by one vector $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Same as before, this ***does not*** *mean that $\mathfrak{so}(2)$ is two dimensional, but rather that every vector is represented by a single 2×2 matrix.*

Because our Lie group G is not necessarily a matrix group there is a way to further generalize the definition of an exponential map. As elements of a Lie algebra are tangent to the Lie group, we can define curves on G to which they are tangent to. Such curves are then called **integral curves**. More generally, given a curve $\gamma \in M$ we can define tangent vectors for each point on such a curve. If those tangent vectors are elements of the same vector field Y, the curve γ is considered an *integral curve.*

Definition 3.4. *Let M be a smooth manifold and let Y be a smooth vector field on M. A smooth curve $\gamma : (a, b) \in \mathbb{R} \to M$ is called an* ***integral curve*** *if*

$$\forall \lambda \in (a,b) : \underbrace{X_{\gamma(\lambda)}}_{\substack{\text{Vector} \\ \text{tangent at} \\ \text{a point } \gamma(\lambda)}} = \underbrace{Y\Big|_{\gamma(\lambda)}}_{\substack{\text{Vector} \\ \text{field valued at} \\ \text{a point } \gamma(\lambda)}} . \tag{3.27}$$

[21]

As a tangent to a curve $\gamma(\lambda)$ can be written as $\frac{d\gamma(\lambda)}{dt}$. Sometimes, in the definition of an integral curve, a tangent vector is written in such form [3, 6, 7].

Note that this is a definition of an integral curve for an arbitrary manifold. For our generalization of the exponential map, we will be looking at the special case of $M = G$ and $Y = L(G)$.

Definition 3.5. *Let $A \in T_eG$ and let $X_g^A := l_{g*}A$ define a uniquely determined left-invariant vector field X^A. Then let $\gamma^A : \mathbb{R} \to G$ be integral curve of X^A passing through the point $\gamma^A(0) = e$, where $e \in G$ is the identity. An* ***Exponential map*** *is then defined as a map*

$$\begin{aligned} exp &: T_eG \to G \\ exp(tA)\Big|_{t=1} &:= \gamma^A(t)\Big|_{t=1} \\ exp(A) &:= \gamma^A(1). \end{aligned} \tag{3.28}$$

[21, 6, 4]

Intuitively, an exponential map is just defined as an integral curve on a Lie group. If we go back to an example of $SO(2)$ we had $R = e^{\theta a} \in G$. The integral curve is a circle as vectors tangent to a circle really are part of a left-invariant vector field. Reasons as to why that is the case may be understood by first thinking of multiplying points on a circle. As $SO(2)$ is closed under multiplication there is no way for us to "fall out" of a circle. We can move from point to point with left translation, *i.e.* $l_g(h) = gh$. So, the induced map l_{g*} then, moves a vector X_h to a point gh, *i.e.* $l_{g*}(X_h) = X_{gh}$.

Example 3.3. *[16]*

Let

$$SO(3) := \{R = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mid RR^T = R^T R = I \text{ and } \det R = 1.\} . \tag{3.29}$$

Similarly to $SO(2)$, this group also describes rotations, only this time, it does so in 3 dimensions. Such rotations are described using 3 matrices, each of which describes 2D rotations around a 3rd, fixed axis. One possible representation of $SO(3)$ is then described by the following three matrices. Any *possible rotation in three dimensions can then be obtained as a composition of these matrices.*

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & sin\theta \\ 0 & -sin\theta & cos\theta \end{pmatrix}, \; R_2 = \begin{pmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{pmatrix} \; and \; R_3 = \begin{pmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.30}$$

As each of these matrices describes 2D rotations, they are all isomorphic to the $SO(2)$*. As one parameter (one axis) is fixed,* $SO(2)$ *is said to be a* ***one-parameter subgroup*** *of a Lie group* $SO(3)$*. Their respective Lie algebra matrices can be obtained in the same way as before as each of these matrices can be written as* $R_i = e^{\theta a_i}$*, where* $a_i = \frac{dR_i}{d\theta}\Big|_{\theta=0}$.

$$a_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad and \quad a_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.31}$$

Every $a_i \in \mathfrak{so}(3)$ *generates a one-parameter subgroup of* $SO(3)$*. As opposed to the* $SO(2)$*,* $SO(3)$ *is harder to visualize. As a manifold, the dimension of* $SO(3)$ *can be determined by the dimension of its Lie algebra* $\mathfrak{so}(3)$*. As* $so(3)$ *is spanned by 3 vectors we know that* $dim(SO(3)) = dim(\mathfrak{so}(3)) = 3$*. This means that topologically,* $SO(3)$ *has a surface that is three-dimensional. That makes it hard to visualize, as we are used to perceiving surfaces in* $2D$ *with their corresponding volumes being* $3D$*. This also means that* $SO(3)$ *can not be represented as a sphere, given that a sphere is a two-dimensional manifold. Such a surface could then, in some interpretations, be represented as a ball (or a "3-sphere"). The main difference is that walking inside of a ball is now the 3rd degree of freedom we have been missing. This also means that it is not possible to visualize the* $4D$ *volume, which this is a surface of. Representing tangents in such a surface is now also tricky, as we traditionally represent tangents on the bounds of a volume, or on the bounds of a surface. The closest representation of such tangent spaces is then imagining an* $\mathbb{R}^3$

defined in each point (both on the surface and the interior) of a ball. This now, intuitively, makes it unclear to what exactly these $\mathbb{R}^3$ spaces are tangent to.

No matter the visualization, $SO(3)$ still *does* describe all possible rotations in three dimensions and it can still be represented with three one-parameter subgroups that have been shown in the example above. Formally, the definition of a one-parameter subgroup is the following one:

Definition 3.6. ***A one-parameter subgroup*** *of a Lie group G is a Lie group homomorphism, i.e. a smooth map*

$$\begin{gathered}\xi : \mathbb{R} \to G \\ \xi(\lambda_1 + \lambda_2) = \xi(\lambda_1) \bullet \xi(\lambda_2)\end{gathered} \tag{3.32}$$

[21, 16, 5, 3, 12]

3.2 Representations of a Lie group and Lie algebra

Representations of certain objects have already been used throughout this work, namely, each time we have *represented* a vector in a certain basis, the idea of representation was intuitively understood as being concrete examples of more general definitions. One other, more recent, example of representation is the choice to *represent* a group $SO(2)$ with either rotation matrices or an exponential map, both of which are specific representations of such a group. So, we can intuitively think of representations as a way of applying, most likely very abstract, definitions to familiar spaces. The term "*representation*" is usually understood to describe "*a very general relationship that expresses similarities (or equivalences) between mathematical objects or structures*" [22]. However, in this subsection, we are concerned with specific kinds of representation that only refer to representing *algebraic structures* as *linear transformations* of vector spaces. There exists a whole branch of mathematics dealing with specific representations, [23, 24], so, as a way not to oversimplify the topic at hand, we will focus on known representations of Lie groups and Lie algebras. It should be noted that representation theory refers only to *linear transformations*, meaning that it is concerned only with structures that can be mapped to vector spaces. There, of course, do exist other kinds of representations. The aforementioned vector represented in a different basis, *in this specific context*, is not considered a representation as it refers to an object as opposed to an algebraic structure. This example, can, however, be referred to as a *representation* in a more general sense because it does "*express equivalence between mathematical objects*". A Lie group $SO(2)$ then would fall into representation theory as it does *represent* a group in the form of a vector space.

In the case of Lie groups, representation is defined to be a *linear transformation*[14] that preserves the group operation while mapping into a vector space. Such a vector space then "*behaves*" as a group, given that all the group axioms are preserved. This means that such a linear map is, by definition set to be a homomorphism. Formally, we can write

Definition 3.7. *Let* $(G, \bullet)$ *be a Lie group and* V *some vector space. Then the* ***representation of a Lie group*** **G** *is a group homomorphism*

$$\begin{aligned} \rho : G \to Aut(V) \\ \rho(g_1 \bullet g_2) = \rho(g_1) \circ \rho(g_2), \end{aligned} \tag{3.33}$$

where $g_1, g_2 \in G$ *and* $\rho(g_1), \rho(g_2) \in Aut(V)$. *[25, 6, 7, 26]* [15]

[14] Not to be confused with a linear map. Linear maps map between two vector spaces. *Linear transformations*, on the other hand, map from some algebraic structure to a vector space, thus they "*transform*" that structure into a vector space. It should be noted that there are cases where these two are used interchangeably as "some structure" can refer to a vector space.

[15] Recall that $Aut(V) \coloneqq \{f : V \xrightarrow{\sim} V \mid f \text{ is invertible.}\}$

For the Lie algebra case, a representation is defined similarly, the only real difference being that instead of the group operation $\bullet$, Lie algebra multiplication, *a.k.a.* Lie bracket, is used.

Definition 3.8. *Let* $(L, [\![\,\_\,,\,\_\,]\!])$ *be a Lie algebra. Then a* ***representation*** *of a Lie algebra* $(L, [\![\,\_\,,\,\_\,]\!])$ *is a map*

$$\begin{aligned} \rho : & L \to End(V) \\ & \rho([\![a,b]\!]) \coloneqq [\rho(a), \rho(b)] , \end{aligned} \tag{3.34}$$

where $[\,\_\,,\,\_\,] : End(V) \times End(V) \xrightarrow{\sim} End(V)$.
[25, 7, 27] [16]

Example 3.4. *[25, 3, 12, 5, 28]*

A specific example of a Lie group representation that is going to be used fairly frequently in the later part of this work, is an ***adjoint representation****. In general, adjoint representation can be defined for both Lie groups and Lie algebras, so to keep the broad use of the term intact, we can informally say that adjoint representation refers to any representation of elements of some algebraic structure in terms of a Lie algebra considered as a vector space.*

To formally define an adjoint representation of a Lie group G*, we need to first introduce an* ***adjoint map****.*[17] *An adjoint map is defined as follows*

$$\begin{aligned} Ad_g : & G \to G \\ & Ad_g(h) \coloneqq g \bullet h \bullet g^{-1} . \end{aligned} \tag{3.35}$$

[16] Recall that $End(V) \coloneqq \{f : V \xrightarrow{\sim} V\}$.

[17] Note that an adjoint map can also be defined for Lie algebras, it then maps to the Lie bracket.

Specifically in the case of $g = e$ we have

$$Ad_g(e) = g \bullet e \bullet g^{-1} = g \bullet g^{-1} = e. \tag{3.36}$$

This is mentioned because the push-forward of an adjoint map can then be used for the Lie algebra case, i.e.

$$\begin{aligned} Ad_{g*} &: T_eG \to T_{Ad_{g(e)}}G \\ Ad_{g*} &: T_eG \to T_eG. \end{aligned} \tag{3.37}$$

In general, $Ad_{g*} \in End(T_eG)$. The following shows how Ad_{g*} acts on some $A \in T_eG$. As A is a left-invariant vector field, in the equation below, we will use $\gamma(t) = e^{tA}$ as that is an integral curve of A.

$$\begin{aligned} &Ad_{g*}A(f) = A(f \circ Ad_g) = \frac{d}{dt}\left(f \circ \left(Ad_g \triangleleft e^{tA}\right)\right)\Bigg|_{t=0} \\ &= \frac{d}{dt}\left(f \circ \left(g \bullet e^{tA} \bullet g^{-1}\right)\right)\Bigg|_{t=0} = f \circ \frac{d}{dt}\left(g \bullet e^{tA} \bullet g^{-1}\right)\Bigg|_{t=0} = f \circ \left(g \cdot A \cdot g^{-1}\right) \\ &\iff Ad_{g*}A = g \cdot A \cdot g^{-1} \end{aligned} \tag{3.38}$$

Note that the "multiplication" that is used in $g \cdot A \cdot g^{-1}$ is not clearly defined, so oftentimes, $Ad_{g*}A = g \cdot A \cdot g^{-1}$ is said to hold true for matrix groups, and $\cdot$ is then just matrix multiplication, while both Lie algebra element A and Lie group element g are then represented in $\mathfrak{gl}(n, \mathbb{C})$ and $GL(n, \mathbb{C})$ respectively.

Both of these maps are now used to define the adjoint representation of a Lie group G.

Definition 3.9. *The **adjoint representation of a Lie group** G is a map*

$$\begin{aligned} Ad : G &\to End(T_e G) \\ g &\mapsto Ad_{g*} \end{aligned} \tag{3.39}$$

[25]

If G is a matrix group, we can then write $Ad(g) = g \cdot A \cdot g^{-1}$, where $\cdot$ denotes matrix multiplication.

§4 Principal fiber bundles

"To the uninitiated, it would seem that the use of fiber bundles and connections to describe the basic forces of nature is a half-baked scheme devised by some clique of mathematicians bent on producing an application for their work. However, physicists themselves found these notions forced upon them by their own perception of nature."

– David Bleecker, *Gauge Theory and Variational Principles*

4.1 Context needed to define principal fiber bundles

The geometry of principal fiber bundles *is* the geometry of gauge fields. As one of the main ingredients needed to explain gauge transformations and covariant derivatives, *connections* can not fully be understood without the presence of principal fiber bundles.

Even though there exist multiple ways of defining principal fiber bundles, here we chose a way that requires terms like group action, orbits, and free action in the main definition as those terms will be used in the later part of this work, regardless of when one introduces them.

4.1.1 Actions of a Lie group on a manifold

We will first start with the idea of the **right action** of a group on a manifold. Similar to the case of a *left translation*, the right action comes with an analogous *left action.* As opposed to the left or right translation, where group elements act on group elements, here, group elements act on an arbitrary manifold. One important thing to point out is that, traditionally, the notation of left/right action is written as a map acting on something *eg.* $\tilde{R} : M \times G \to M$ acts in the following way $(p, g) \mapsto \tilde{R}(p, g) = p \cdot g$. Here, we will be writing the left or right action as $\triangleright$ or $\triangleleft$ respectively, more similar to the notion of *operation* rather than a map.[18]

Definition 4.1. *Let* $(G, \bullet)$ *be a Lie group and* M *be a smooth manifold. Then a smooth map* $\triangleright : G \times M \to M$ *that satisfies the following:*

(i) $e \triangleright p = p\ ,\ \forall p \in M$ *and* $e \in G$

(ii) $g_2 \triangleright (g_1 \triangleright p) = (g_2 \bullet g_1) \triangleright p\ ,\ \forall g_1, g_2 \in G$ *and* $p \in M$

is called a ***left action*** *of* G *on* M*, or a* ***left G-action****. [29, 5]*

Similarly, one can define a **right action** analogously.

Definition 4.2. *Let* $(G, \bullet)$ *be a Lie group and* M *be a smooth manifold. Then a smooth map* $\triangleleft : M \times G \to M$ *that satisfies the following:*

(i) $p \triangleleft e = p\ ,\ \forall p \in M$ *and* $e \in G$

(ii) $(p \triangleleft g_1) \triangleleft g_2 = p \triangleleft (g_1 \bullet g_2)\ ,\ \forall g_1, g_2 \in G$ *and* $p \in M$

is called a ***right action*** *of* G *on* M*, or a* ***right G-action****.*
[29, 6, 4, 12]

It should be noted that (i) and (ii) do not need to be defined as it is possible to derive them using the predefined left action [29].[19] Any manifold M that is being acted upon with a left or right action is called a **left** or a **right G-space** respectively [5].

[18]Though, every operation *is* indeed a map. The only real distinction between the two is how one writes it, *eg.* one could easily write $+(2, 3) = 5$ as opposed to $2 + 3 = 5$ and the meaning of the statement would not change.

[19]This can also go the other way around as one can define right action and derive the properties of the left action from there.

4.1.2 Equivariance

Given two groups G and H and two left actions, we arrive at the notion of an **equivariance**. Intuitively one can think of ρ-equivarance as preservation of left action or as a homomorphism between the left G-space and the left H-space.[20]

Definition 4.3. *Let $(G, \bullet)$ and $(H, \blacksquare)$. Let $\rho : G \to H$ be a Lie group homomorphism, i.e. $\rho(g_1 \bullet g_2) = \rho(g_1) \blacksquare \rho(g_2)$, $\forall g_1, g_2 \in G$. Also, let M and N be smooth manifolds. Given two left actions*

$$\triangleright : G \times M \to M \tag{4.1}$$

$$\blacktriangleright : H \times N \to N \tag{4.2}$$

and a smooth map $f : M \to N$, f is called
***ρ-equivariant** if the following diagram commutes:*

$$\begin{array}{ccc} G \times M & \xrightarrow{\rho \times f} & H \times N \\ \triangleright \big\downarrow & & \big\downarrow \blacktriangleright \\ M & \xrightarrow[f]{} & N \end{array}, \tag{4.3}$$

i.e.

$$f(g \triangleright p) = \rho(g) \blacktriangleright f(p). \tag{4.4}$$

[29, 5][21]

[20]Note that this is not the standard terminology.

[21]The reading of commutative diagram might be easier to understand if one wrote $(f \circ \triangleright)(g, p) = (\blacktriangleright \circ (\rho \times f))(g, p) \implies f(\triangleright(g, p))) = \blacktriangleright (\rho(g), f(p)) \implies f(g \triangleright p) = \rho(g) \blacktriangleright f(p)$.

4.2 Orbit

If one were to act on some $p \in M$ with all the group elements simultaneously, with either a left or right action, one would get an **orbit** [29, 5, 3, 4]. One good example of an orbit is one for a Lie group we already covered. For manifold $\mathbb{R}^2$ and a Lie group $SO(2)$ what describes an orbit is a rotation matrix. As we can act on some $p \in \mathbb{R}^2$ with such a matrix, the outcome is a circle. For multiple p's we get concentric circles as shown in the figure below.

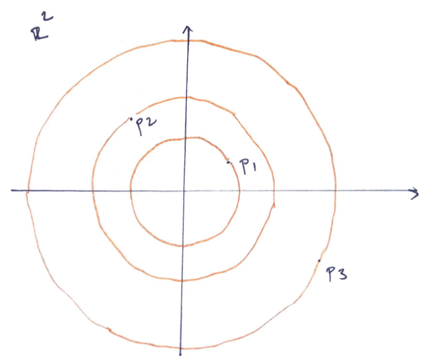


Figure 4.1: Orbit of a left action such that group elements from $SO(2)$ act onto manifold $\mathbb{R}^2$.

Formally, an orbit is just a set

$$\mathcal{O}_p := \{q \in M \mid \exists g \in G \text{ s.t. } g \triangleright p = q\}. \tag{4.5}$$

It should be noted that orbit is a subset of a manifold, *i.e.* $\mathcal{O}_p \subset M$.

4.3 Equivalence class

To further analyze the action of a Lie group on a manifold, we can focus our attention on orbits. The left action *is* what defines an orbit. A more general way to describe such behavior is to say that a left action is an *equivalence relation.* An equivalence relation always subdivides some starting set into disjoint subsets. No two orbits we got for $SO(2)$ can ever overlap and that is the case for any arbitrary equivalence relation acting on any arbitrary set. Because they are disjoint, one can represent one such subset with one element. That element is then called an *equivalence class* and it is part of a new "*modded out*" space called a *quotient space*.

Definition 4.4. *Given a set X and some relation $\sim$ we define an **equivalence relation** for all a, b and $c \in X$ iff the following properties hold true*

(i) reflexivity, i.e. $a \sim a$,

(ii) symmetry, i.e. $a \sim b \iff b \sim a$ and

(iii) transitivity, i.e. if $a \sim b$ and $b \sim c \implies a \sim c$.

*The subset of all elements related with an equivalence relation, i.e. $[a] := \{x \in X \mid x \sim a\}$ defines an **equivalence class**. The set of all equivalence classes then further defines a **quotient space**.*
[6, 5, 3, 30, 31]

4.3.1 Left action defines equivalence classes called orbits

The following example can be found in [29]. In our case, the left action is an equivalence relation and the orbits define equivalence classes. One can check that the noted properties really do hold by following the definitions of both the left action (that is now our relation) and the equivalence relation.

Let $p \sim q :\iff \exists g \in G$ such that $q = g \triangleright p$. In other words, two points are equivalent iff they lie in the same orbit.

(i) reflexivity

$$p \sim p \iff \exists g \in G \text{ such that } p = g \triangleright p. \tag{4.6}$$

Since G is a group, such g really does exist - it is called the neutral, and is usually denoted by e. So the left action really is reflexive.

(ii) symmetry

We have to show that from $p \sim q$ follows $q \sim p$.

$$\begin{aligned} p \sim q \iff q &= g \triangleright p \\ q &= g \triangleright p \Big/ g^{-1}\triangleright \\ g^{-1} \triangleright q = g^{-1} \triangleright (g \triangleright p) &\underset{\substack{\text{By def.} \\ \text{of a } \triangleright.}}{=} \underbrace{(g^{-1} \bullet g)}_{=e} \triangleright p = p \end{aligned} \tag{4.7}$$

If $p = g^{-1} \triangleright q$ this means that $q \sim p$. Since we have gotten from $p \sim q$ to $q \sim p$ we can conclude that the left action really is symmetric.

(iii) transitivity

Let $p \sim q$ and $q \sim r$. We have to show that $p \sim r$.

If $p \sim q \implies q = g_1 \triangleright p$, and similarly $q \sim r \implies r = g_2 \triangleright q$. We can now rewrite r.

$$\begin{aligned} r = g_2 \triangleright q = g_2 \triangleright (g_1 \triangleright p) &= \underbrace{(g_2 \bullet g_1)}_{=\tilde{g}} \triangleright p \\ &\implies r = \tilde{g} \triangleright p \\ &\iff p \sim r \end{aligned} \tag{4.8}$$

This concludes a proof that the left action really is an equivalence relation.

All equivalence classes of a left action form the **orbit space** which will be denoted as $M\backslash_{\sim}$ or $M\backslash_{G}$. One can imagine them as concentric circles we have drawn in the case of $SO(2)$ acting on $\mathbb{R}^2$.

4.4 Free action

One other thing that should be introduced before defining principal fiber bundles is the idea of a *free action*. Informally, we call the left/right action free, if the only element that does not change the points $p \in M$ is the neutral of a Lie group. This might seem odd at first, but recall that we are acting with group elements on something that is not a group. It is not a given fact that only the neutral element does not move our points. Meaning there might exist g_1 and g_2 such that $p = g_1 \triangleright p = g_2 \triangleright p$. In fact, in our example of $G = SO(2)$ and $M = \mathbb{R}^2$ we can set $p = (0, 0) \in \mathbb{R}^2$. No matter the group element we rotate the origin with, we will still stay at the origin. This means that now all of our group elements "*act like*" a neutral as for all of them holds that $(0, 0) = g \triangleright (0, 0)$, $\forall g \in G$ [29, 4, 5, 12, 32].

To formally define a free action, we need to introduce the idea of a **stabilizer**. A stabilizer S_p is simply a set of all such group elements that do not move the points in a manifold, *i.e.*

$$S_p \coloneqq \{g \in G \mid p = g \triangleright p\}. \tag{4.9}$$

Note that S_p is a subset of G. An orbit refers to points on a manifold, while a stabilizer refers to group elements. In the case of the Lie group $SO(2)$, for any point $p \in \mathbb{R}^2$, a stabilizer is *always* a neutral element as it never rotates points, *i.e.* $S_p = \{e\}, \forall p \in M$. The only time that a stabilizer consists of more than the neutral element is when $p = (0, 0)$ as then $S_{(0,0)} = SO(2)$.

A stabilizer that *always* consists of **only** a neutral element (no matter the choice of $p \in M$) defines a free action.

Definition 4.5. *An action $\triangleright$ is called* ***free*** *if, separately, $\forall p \in M : S_p = \{e\}$ [29].*

So, for $G = SO(2)$ and $M = \mathbb{R}^2$ the left action *is not* free. However, if

$G = SO(2)$ and $M = \mathbb{R}^2 \setminus \{(0,0)\}$, then $\triangleright$ is a free action.

There is an important observation that if $\triangleright$ is a free action on M, then each orbit $\mathcal{O}_p$ is homeomorphic to the Lie group G.

Given a free action, there is a bijective correspondence between an orbit $\mathcal{O}_p$ and the Lie group G as each element in G then "produces" exactly one point in the orbit. As all the maps and manifolds included are smooth, such a bijection is then a diffeomorphism. Diffeomorphisms can be thought of as isomorphisms that preserve continuity in both directions, meaning one can write $\mathcal{O}_p \cong G$.

4.4.1 G-bundle

Let E, a total space of a bundle be a right G-space. With the right action on E one can construct a new bundle - the bundle whose base space is the orbit space of E. If our starting bundle is isomorphic to the newly constructed bundle, *i.e.* if

$$\begin{array}{c} E \\ {\scriptstyle\pi}\big\downarrow \\ M \end{array} \cong_{bundle} \begin{array}{c} E \\ \big\downarrow{\scriptstyle\rho} \\ E\backslash_G \end{array}, \tag{4.10}$$

then (E, M, π) is called a **G-bundle**. Note that ρ just projects orbits into points.[22] If (E, M, π) is a G-bundle then $M \cong E\backslash_G$ [5].

4.5 Principal fiber bundle

Let $E \xrightarrow{\pi} M$ be a G-bundle. Note that, because E is a right G-space, one can write $q = p \triangleleft g$. What happens if one projects the point q from a total

[22] Reminder that one can do that using equivalence relation, which is, in this case, the right action.

space into the base space?

$$q = p \triangleleft g \;\Big/\; \pi$$
$$\pi(q) = \pi(p \triangleleft g) \tag{4.11}$$

Recall that a fiber is defined as all the points that get projected into the same point in a base space. This means that points q and $p \triangleleft g$ **live in the same fiber**. The other way to state this is to say that the *group action preserves the fibers*. The points q and $p \triangleleft g$ are exactly the points that define an orbit. **Orbits of a G-bundle are the fibers** themselves.

We also know that if $\triangleleft$ is free, that $\mathcal{O}_p \cong G$. This in turn means that the **fibers** of a G-bundle which have been constructed by a free right action **are isomorphic to a Lie group**. This is exactly how one defines a **principal fiber bundle**.

Definition 4.6. *A smooth bundle (E, M, π) is called a* ***principal fiber bundle*** *or a* ***principal G-bundle*** *if the following holds true:*

(i) E is a right G-space,

(ii) $\triangleleft$ is free,

(iii) There exists a bundle isomorphism

$$\begin{array}{ccc} E & & E \\ \pi\big\downarrow & \cong_{bundle} & \big\downarrow\rho \\ M & & E\backslash_G \end{array} .$$

[29, 12, 20, 5]

Note that as fibers of a principal bundle are Lie groups, one can think of the total space P of a principal fiber bundle as, at least *locally*, being a product space $P = G \times M$.

4.5.1 Frame bundle as an example of a principal fiber bundle

The following example can be found in [29, 5, 3, 12, 20, 7, 6]. Similar to how the tangent bundle is defined as a manifold having a tangent space that is defined for each of its points, here a frame bundle is defined as a set of all possible coordinate frames of some manifold.

Let $x \in M$. Recall that the tangent space in that point is labeled T_xM. We define a space of all possible frames for that point as

$$L_xM \coloneqq \{(e_a, ..., e_{\dim M}) \mid (e_a, ..., e_{\dim M}) \text{ is a basis for } T_xM.\}. \tag{4.12}$$

If one recalls how we got to the structural group of a tangent bundle in the subsection (2.3.1), it was by change of frames of the same tangent space. The transition matrix we got for that change of frames was the element of a structure group $GL(n, \mathbb{R})$, where $\dim(T_pM) = n$. In that same chapter, we learned that the relationship between tangent vector coordinates can be written with the help of a transition matrix from the $GL(n, \mathbb{R})$, *i.e.* $x^\nu = \tilde{x}^\mu \left(\frac{\partial x^\nu}{\partial y^\mu}\right)_p = \tilde{x}^\mu G^\nu_\mu$, where $G^\nu_\mu \in GL(n, \mathbb{R})$ and $x \in T_pM$. From this, it follows that there exists a one-on-one correspondence between the set of all frames and the set of all invertible matrices. In more general terms this then allows us to write $L_xM \cong GL(\dim M, \mathbb{R})$. Also recall that $GL(n, \mathbb{R})$ is a Lie group.

We define a frame bundle in complete analogy to a tangent bundle. A total space is then just a disjoint union of all L_xM.[23]

$$LM \coloneqq \dot{\bigcup_{x \in M}} L_xM \tag{4.13}$$

[23] Or $L_xM \times M$, but as was discussed in the tangent bundle section (2.3), those two amount to the same as each L_xM contains within itself $x \in M$.

This now means we have a bundle

$$\begin{aligned} LM &\xrightarrow{\pi} M, \\ (e_1, ..., e_{\dim M}) &\mapsto x. \end{aligned} \tag{4.14}$$

If we want to see if such a bundle really is a principal bundle, we should define a right action of $GL(\dim M, \mathbb{R})$. Let $\dim M = d$. We define a right action on LM as one would define a change of basis

$$(e_1, ..., e_d) \triangleleft g \coloneqq (e_m g^m_1, ..., e_m g^m_d), \tag{4.15}$$

where $g^m_n \in GL(d, \mathbb{R}) = \{g^m_n \in \mathbb{R} \mid m, n = 1, ..., d \text{ and } det(g^m_n) \neq 0.\}$.

Now we have a right G-space, but is this a free action?

$$\begin{aligned} &(e_1, ..., e_d) \triangleleft g = (e_1, ..., e_d) \\ &\iff (e_m g^m_1, ..., e_m g^m_d) = (e_1, ..., e_d) \\ &\iff e_m g^m_n = e_n \end{aligned} \tag{4.16}$$

Because e_m and e_n are basis vectors, they should be linearly independent, from which we can conclude that $e_m = e_n \implies g^m_n = g^m_m$, *i.e.* g is an identity matrix. This now means $\triangleleft$ is a free action.

The last thing to check is the isomorphism between the following bundles

$$\begin{array}{ccc} LM & \xrightarrow{id} & LM \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\rho} \\ M & \xrightarrow[f]{} & LG\backslash_{GL(d,\mathbb{R})} \end{array} . \tag{4.17}$$

In other words, does this hold true $\rho \stackrel{?}{=} f \circ \pi$?

ρ should be defined as

$$\begin{aligned} \rho : LM &\to LG\backslash_{GL(d,\mathbb{R})} \\ (e_1, ..., e_d) &\mapsto (e_1, ..., e_d) \triangleleft g \end{aligned} \tag{4.18}$$

And for $f \circ \pi$ we should have some other $\tilde{g} \in GL(d, \mathbb{R})$.

$$\begin{aligned} f \circ \pi : LM &\to M \to LG\backslash_{GL(d,\mathbb{R})} \\ (e_1, ..., e_d) &\mapsto x \mapsto (e_1, ..., e_d) \triangleleft \tilde{g} \end{aligned} \tag{4.19}$$

If we come back to the definition of an orbit, one should recall that we get the orbit with the action of a whole group on some $p \in M$. The whole orbit then gets "modded out" to one point as we go from the starting manifold to orbit space. Since we started in the same $(e_1, ..., e_d) \in LM$ for both functions ρ and $f \circ \pi$ it does not matter with which $g \in GL(d, \mathbb{R})$ we acted with on that element - we still remain in the same orbit, *i.e.* $(e_1, ..., e_d)$ gets mapped to the same equivalence class $[(e_1, ..., e_d) \triangleleft g] = [(e_1, ..., e_d) \triangleleft \tilde{g}]$.

Thus, the frame bundle really is a principal fiber bundle.

4.5.2 Trivial principal bundle

The following can be found in [29]. To better understand isomorphisms between two principal bundles, we will first define a weaker notion and build to the one that is more abstract. The goal is for the following diagram to commute, *i.e.* if the following diagram commutes, we can say that the two principal bundles are isomorphic.[24] Note that now we have added one more row to the bundle diagram to emphasize that P is a right G-space.

[24]Similar to bundle map isomorphism in subsection (2.2.1).

$$\begin{array}{ccc} P & \xrightarrow{u} & P' \\ {\scriptstyle \triangleleft G}\uparrow & & \uparrow{\scriptstyle \triangleleft' G} \\ P & \xrightarrow{u} & P' \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi'} \\ M & \xrightarrow[f]{} & M' \end{array} \tag{4.20}$$

If we want this entire diagram to commute, the following should hold true

$$f \circ \pi = \pi' \circ u \text{ and} \tag{4.21}$$

$$u(p \triangleleft g) = u(p) \triangleleft' g, \quad \forall g \in G \text{ and } \forall p \in P. \tag{4.22}$$

Here, we are acting with the same group G on two different principal bundles, but more general isomorphism should include two different groups G and G' acting on their respective principal bundles. One does that by introducing a Lie group homomorphism.

Let $(G, \bullet)$ and $(G', \bullet')$ be two Lie groups. Then a Lie group homomorphism is a map $\rho : G \to G'$ defined as

$$\rho(g_1 \bullet g_2) = \rho(g_1) \bullet' \rho(g_2), \tag{4.23}$$

where $g_1, g_2 \in G$ and $\rho(g_1), \rho(g_2) \in G'$.

Now the following commutative diagram should commute.

$$\begin{array}{ccc} P & \xrightarrow{u} & P' \\ {\scriptstyle \triangleleft G}\uparrow & \xrightarrow{\rho} & \uparrow{\scriptstyle \triangleleft' G'} \\ P & \xrightarrow{u} & P' \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi'} \\ M & \xrightarrow[f]{} & M' \end{array} \tag{4.24}$$

i.e.

$$f \circ \pi = \pi' \circ u \text{ and} \tag{4.25}$$

$$u(p \triangleleft g) = u(p) \triangleleft' \rho(g), \quad \forall g \in G \text{ and } \forall \rho(g) \in G'. \tag{4.26}$$

This is called a **principal bundle map** and similar to a bundle map, it is used to represent the principal bundle isomorphism. Here, both the bundle projection and the right action are preserved. A principal bundle map is used to define a **trivial principal bundle**. A principal bundle is considered trivial if we can globally write $P = G \times M$. Formally, we use principal bundle maps to express that notion.

Definition 4.7. *A **principal fiber bundle** is called **trivial** if there exists the following isomorphism*

$$\begin{array}{ccc} P & & G \times M \\ {\scriptstyle \triangleleft G}\uparrow & & \uparrow{\scriptstyle \blacktriangleleft G} \\ P & \cong & G \times M \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi_1} \\ M & & M \end{array} , \qquad \begin{array}{c} (x,g) \triangleleft g' \\ := (x, g \bullet g') \end{array} \tag{4.27}$$

[29, 5, 12]

There exists an important theorem concerning trivial principal bundles. The theorem states that one can have a globally defined smooth section on a bundle **if and only if** a given bundle is trivial. As cross-sections were used to define vector fields this theorem is of great importance if one wishes to work with well-defined vector fields.

Theorem 4.1. *A principal G-bundle $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ is trivial iff there exists a smooth section $\sigma : M \to P$ such that $\pi \circ \sigma = id_M$.*
[29, 6, 5, 3]

§5 Associated fiber bundles

Associated bundles will be used as a tool for solving a very important issue of comparing the points on manifolds. Intuitively, one can think of associated bundles as replacing the fibers of a principal bundle with new fibers. Those new fibers can then be set to be vector spaces which will allow for a familiar environment and thus for the comparison of points. This will, in turn, allow for the proper definition of *covariant derivatives.*

Suppose we glue new fibers F to the total space of a principal fiber bundle P. Suppose those fibers F are a left G-space, *i.e.* we have a well-defined left action $\triangleright : G \times F \to F$. What we now have is some kind of a product space, that can, at least locally, be thought of as "$((M \times G) \times (G \times F))$", or, because the group is the same "$(M \times G \times F)$". We can define some kind of a "product space equivalence relation" that acts on M from the right and acts on F from the left simultaneously.

$$(p, f) \sim_G (p', f') \tag{5.1}$$

$$(p', f') \coloneqq (p \triangleleft g, g^{-1} \triangleright f) \tag{5.2}$$

As both the right action and the left action are equivalence relations, this "product relation" inherits that property on a slot-by-slot basis.[25] This relation now defines a new *quotient space.* Because we "modded out" the action of a group, we can say, at least locally, that "$((M \times G) \times (G \times F)) \backslash_{\sim_G} = M \times F$". That is why we can think of such bundles as if we "replaced the fibers". This intuitive explanation can be found in [12, 33].

Examples of an associated bundle to a frame bundle include the tangent bundle, cotangent bundle and tensor bundle. For the case of a tangent bundle, one can imagine gluing a "flat space" $\mathbb{R}^d$ to each point of the total space of a frame bundle. Modding out these frames then results in a tangent space

[25]Proof that group actions are equivalence relations is in the subsection (4.3.1).

at each point [12, 5].

As the associated bundle is defined on a global basis the "$M \times G \times F$" bit gets replaced with the $P \times F$. The formal definition also describes the projection of such a quotient space as we still *do* want to define a bundle.

Definition 5.1. *Given a principal G-bundle $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ and a smooth manifold F that is a left G-space, i.e.*

$$\triangleright : G \times F \to F. \tag{5.3}$$

We define the following:

(i) Let $\sim_G$ be the relation on $P \times F$. The relation $\sim_G$ acts in the following manner: $(p, f) \sim_G (p \triangleleft g, g^{-1} \triangleright f)$, $\forall g \in G$ and $\forall (p, f) \in P \times F$. Such relation $\sim_G$ is the equivalence relation on $P \times F$, i.e.

$$[(p, f)] = [(p \triangleleft g, g^{-1} \triangleright f)]$$

This then defines a quotient space $P_F \coloneqq (P \times F)\backslash_{\sim_G}$.

(ii) We also define a projection map

$$\begin{aligned} \pi_F &: P_F \to M \\ [p, f] &\mapsto \pi(p). \end{aligned}$$

The map π_F is well defined as the following holds.

$$\begin{array}{ccc} \pi_F([(p \triangleleft g, g^{-1} \triangleright f)]) & \overset{?}{=} & \pi_F([(p, f)]) \\ \shortparallel & & \shortparallel \\ \pi(p \triangleright g) & = & \pi(p) \end{array}$$

$$\implies \pi_F([(p \triangleleft g, g^{-1} \triangleright f)]) = \pi_F([(p, f)])$$

The triple (P_F, M, π_F) is called an ***associated bundle*** *to the principal bundle (P, M, π).[34, 5, 3, 7, 12, 4]*

5.1 Tangent bundle as an example of an associated bundle

The following example can be found in [34, 5].

Let

$$P = LM$$
$$G = GL(d,\mathbb{R})$$
$$F = \mathbb{R}^d$$

We define:

$$\triangleleft : LM \times GL(d,\mathbb{R}) \to LM \tag{5.4}$$
$$(e \triangleleft g)_a := e_i g^i_a$$
$$\triangleright : GL(d,\mathbb{R}) \times \mathbb{R}^d \to \mathbb{R}^d \tag{5.5}$$
$$(g^{-1} \triangleright x)^a := (g^{-1})^a_j x^j$$

With a left action defined in such a way, the associated bundle is then a bundle

$$(LM \times \mathbb{R}^d)\backslash_{GL(d,\mathbb{R})} \xrightarrow{\pi_{\mathbb{R}^d}} M. \tag{5.6}$$

This bundle is isomorphic to the tangent bundle, *i.e.* the following diagram commutes

$$\begin{array}{ccc} (LM \times \mathbb{R}^d)\backslash_{GL(d,\mathbb{R})} & \xrightarrow{u} & TM \\ \pi_{\mathbb{R}^d}\downarrow & & \downarrow \pi_{\text{TM}} \\ M & \xrightarrow[id]{} & M \end{array} . \tag{5.7}$$

The map u is defined as

$$u : (LM \times \mathbb{R}^d)\backslash_{GL(d,\mathbb{R})} \to TM \tag{5.8}$$
$$[(e,x)] \mapsto e_a x^a.$$

Note that u is diffeomorphic, since if that were not the case, we could not have a bundle isomorphism.

5.2 Associated bundle map

The following definition can be found in [34, 5] For an associated bundle map the following diagram should commute

$$\begin{array}{ccc} P_F & \xrightarrow{\tilde{u}} & P'_F \\ {\scriptstyle \pi_F}\downarrow & & \downarrow{\scriptstyle \pi'_F} \\ M & \xrightarrow[\tilde{h}]{} & M' \end{array}, \qquad \pi'_F \circ \tilde{u} = \tilde{h} \circ \pi_F \tag{5.9}$$

Notice that both bundles share the same fiber F. That is because such a bundle map is constructed from the underlying principle fiber bundle map

$$\begin{array}{ccc} P & \xrightarrow{u} & P' \\ {\scriptstyle \triangleleft G}\uparrow & & \uparrow{\scriptstyle \triangleleft' G} \\ P & \xrightarrow{u} & P' \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi'} \\ M & \xrightarrow[h]{} & M' \end{array} \qquad \begin{aligned} h \circ \pi &= \pi' \circ u \\ u(p \triangleleft g) &= u(p) \triangleleft' g. \end{aligned} \tag{5.10}$$

The map $\tilde{u}$ is defined as $\tilde{u}([p, f]) \coloneqq [u(p), f]$, for all $p \in P, f \in F$ and $m \in M$.

For those two bundles to be isomorphic, maps $\tilde{u}$ and $\tilde{h}$ should be diffeomorphisms. Because of the new property that was introduced for associated bundle maps, a bundle map isomorphism and the associated bundle map isomorphism are different. It may be possible that two fiber bundles with the same fibers F are isomorphic as bundles, but not as associated bundles.

Similar to how the notions of a bundle isomorphism have changed, the notions of triviality for the case of associated bundles have also changed.

Definition 5.2. *An associated bundle is called **trivial** if the underlying principal fiber bundle is trivial. [34, 5]*

One can prove that every trivial associated bundle is actually a trivial bundle. Note that the converse of that statement does not necessarily hold.

One last note to this chapter is that, because of the way they are constructed, associated bundles reflect the changes happening in the principal bundle they are associated with. As we change the frames, vector components change accordingly. Any twists in a non-trivial principal bundle are reflected in all of its associated bundles.

§6 Connections and connection 1-forms

A very much anticipated topic of connections will be used for *connecting* the fibers of both the principal bundle and the associated bundles. Their direct application includes the topics of gauge transformations, parallel transport, and covariant derivatives.

The motivation and explanation given below can be found in [34, 5]. Suppose the principal bundle at hand is a torus whose fibers S^1 can be regarded as a Lie group $SO(2)$.[26] The question of comparison of points on a base manifold is lifted to the total space of the principal bundle.

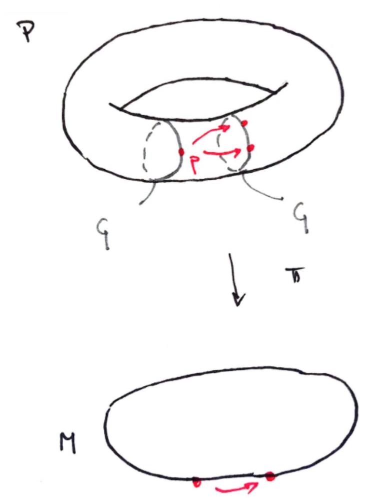


Figure 6.1: $S^1 \times S^1 \xrightarrow{\pi} S^1$

Now the question is one of comparing points of neighboring fibers. The basic idea is to put a tangent space T_pP at some point $p \in P$ and ”*decompose*”

[26] Rotation is a free action as the only element that does not move the points is an identity matrix. Moreover, the isomorphism between such a bundle and the ”*modded-out bundle*”, as required by the definition of a principal bundle, exists as each circle $SO(2)$ gets equated with one point. As circles $SO(2)$ are ”*stacked*” beside each other, so are the ”*modded out*” points resulting in a base space that is isomorphic to the S^1. Such a bundle really can be regarded as a principal bundle.

it to the spaces that will then span T_pP. Similar to how one can represent the $\mathbb{R}^2$ with an x and a y axis. In the case of a tangent space, such subspaces are called a **vertical subspace** and **horizontal subspace** and are labeled V_pP and H_pP respectively. Their direct sum spans the whole T_pP.

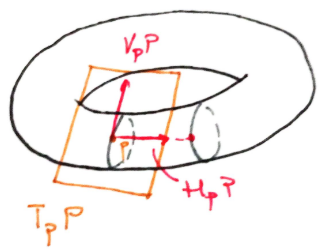


Figure 6.2: $H_pP \oplus V_pP = T_pP$

For the neighboring points that are *close enough*, the horizontal subspace should "*intersect*" the neighboring fiber. Note that it is an abuse of terminology to say that H_pP intersects some other point on a torus, as it does not. It lives in an entirely different space.

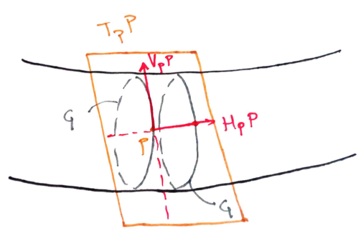


Figure 6.3: Local representation of a total space $P = S^1 \times S^1$.

However, if those fibers are infinitesimally close to each other, the mistake one makes by such identification gets small enough so that one could say that you really do hit a point nearby. It should be noted that "*connecting*" neighboring fibers with vectors from a horizontal subspace is a **choice**. One could have chosen a different way of achieving that.

As each T_pP can be written as $V_pP \oplus H_pP = T_pP$, each $X_P \in T_pP$ can be obtained by the sum of its vertical and horizontal parts, *i.e.* $X_p = hor(X_p) + ver(X_p)$.

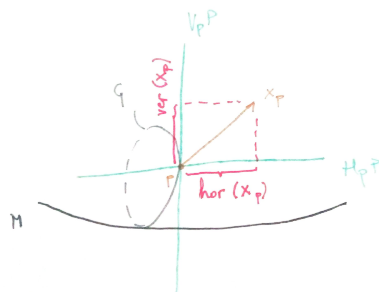


Figure 6.4: $X_p = hor(X_p) + ver(X_p)$

Let X_p live entirely in vertical subspace. If one were to push $X_p = ver(X_p)$ to the base manifold, it would map to a null vector.

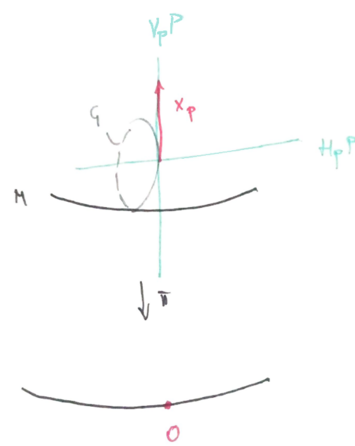


Figure 6.5: $X_p = ver(X_p)$

Because that is the case, a vertical space is then defined by its property to map vertical vectors to null vectors. Because $V_pP \oplus H_pP = T_pP$, a vertical subspace uniquely defines a horizontal subspace, so there is no need to define it separately. However, it *is* possible to define a horizontal subspace with no reference to the vertical one, just not at this point.

Definition 6.1. *Let $p \in P$. Then a* ***vertical subspace*** V_pP *is the set*

$$V_pP := ker(\pi_*) = \{X_p \in T_pP \mid \pi_*(X_p) = 0\} \tag{6.1}$$

[34, 5, 4]

Note that, vectors from a vertical subspace are always tangent to fibers. Even though we do not provide a complete proof, the following explanation will help us gain intuition. A good resource to make this statement precise is [7, p. 258-259]. Because a total space can, at least locally, be considered $M \times G$, we can say that for a vector tangent to the total space $v \in T_mM \oplus T_gG$. As that is also a product space let us temporarily write $v = (v_m, v_g)$, where $v_m \in T_mM$ and $v_g \in T_gG$. As π is a projection we can write $\pi(m, g) = m$. Similar holds for the induced map π_* as we can write $\pi_*(v_m, v_g) = v_m$. If v is tangent only to a fiber, we can write $v = (0, v_g)$, meaning, its projection is now $\pi_* v = \pi_*(0, v_g) = 0$. As that equation satisfies the definition of a vertical subspace, sometimes, a vertical subspace is defined as a subspace that is tangent to the fibers [3, 7]. If fibers were to be frames, or some other entity without a physical shape, the statement that a *vertical subspace is tangent to the fibers* would not really have a clear meaning. For that reason, a definition that does not have instances of clashing with intuition will be used.

In general, any vector that lives entirely in a vertical subspace, *i.e.* $X_p \in V_pP$ belongs to the vector field induced from a left-invariant vector. Such a vector field is then said to be an **induced vector field** or a **fundamental vector field** and it is composed of vectors from all the vertical subspaces [12].

Definition 6.2. *Let $P \xrightarrow{\pi} M$ be a principal G-bundle. Then each $A \in T_eG$* ***induces a vector field*** *on total space P, $\forall p \in P, f \in C^\infty(M)$*

$$X_p^A f := \frac{d}{dt}\left(f \circ (p \triangleleft e^{tA})\right)\bigg|_{t=0}. \tag{6.2}$$

[34, 5, 3, 4, 7]

Here, the curve γ from the definition of a tangent vector is replaced with an integral curve, which is, in our example of a torus, a circle. As $p \triangleleft g$ always traces the orbit, $p \triangleleft e^{ta}$ can always be considered a curve.

Lemma 6.1. *Every vector from the induced vector field lies entirely in the vertical subspace [34].*

Proof.

$$X_p^A f := \frac{d}{dt}\left(f \circ (p \triangleleft e^{tA})\right)\bigg|_{t=0}$$

$$\pi_* X_p^A(f) \underset{\substack{\nearrow \\ \text{By def. of} \\ \text{a push-} \\ \text{forward.}}}{=} X_p^A(f \circ \pi) \underset{\substack{\nwarrow \\ \text{By def.} \\ \text{of } X_p^A.}}{=} \frac{d}{dt}\left[(f \circ \pi)(p \triangleleft e^{tA})\right]\bigg|_{t=0} = \frac{d}{dt}\underbrace{\left[f(\pi(p))\right]}_{\substack{\text{As } f \in C^\infty(M), \\ \text{this is constant.}}}\bigg|_{t=0} = 0$$

$$\implies \pi_*(X_p^A) = 0 \implies X_p^A \in V_pP. \tag{6.3}$$

□

If one were to move a vector from the vertical subspace with the use of the right action, such a vector would then only move on its starting fiber, making it again, a vector from the vertical subspace. As that is the case, the same should hold true for the horizontal subspace as they should, together, span the whole tangent space. Vectors from horizontal subspaces should be pushed to horizontal subspaces. That is done by mapping the horizontal subspace itself. Hence, the last requirement to be met is for the horizontal subspaces to get mapped to horizontal subspaces. Push-forward of the right action is used to map horizontal subspaces.

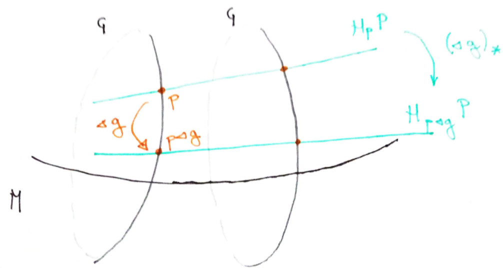


Figure 6.6: $(\triangleleft g)_* H_p P = H_{p \triangleleft g} P$.

This whole narrative can be summed up in the formal definition of a **connection**. Informally, one can think of a connection as a structure on a principal bundle that allows one to compare quantities on the neighboring fibers. Such a structure then, behaves exactly as described.

Definition 6.3. *The **connection** on a principal G-bundle* $P \xrightarrow{\pi} M$ *is an smooth assignment of a vector subspace* $H_p P$ *to each point* $p \in P$ *so that* $H_p P$ *abides to the following*

(i) Each tangent space can be decomposed to the direct sum of a vertical and horizontal subspace i.e. $H_p P \oplus V_p P = T_p P$,

(ii) Horizontal subspaces are pushed to the horizontal subspaces, i.e. $(\triangleleft g)_* H_p P = H_{p \triangleleft g} P$,[27]

(iii) Each $X_p \in T_p P$ *has a unique decomposition to its vertical and horizontal components, i.e.* $X_p = hor(X_p) + ver(X_p)$.

[34, 5, 3, 12, 20]

[27]Note that while the right action is defined as $\triangleleft : G \times M \to M$ for some M, $\triangleleft g : M \to M$.

Note that the third condition follows from the first one so its inclusion in the definition is not necessary. The whole structure of a connection on a principal bundle can actually be packed into an object called a **connection 1-form**.

A connection 1-form maps every vector from T_pP into a vector from a Lie algebra of left-invariant vector fields $L(G) \cong T_eG$. Before the formal definition of a connection 1-form, we will first define an auxiliary map i_p that will, to every left-invariant vector, assign a vector from the induced vector field [34, 5].

$$\begin{aligned} i_p :& T_eG \xrightarrow{\sim} T_pP \\ & A \mapsto X_p^A \end{aligned} \tag{6.4}$$

As vectors from the induced vector field are by definition implied to live in T_pP, the codomain of a map i_p is set to T_pP. However, recall that every X_p^A is actually part of a vertical subspace. Also, note that such a map is an isomorphism. This allows us to write the following

$$\begin{aligned} i_p :& T_eG \overset{\sim}{\rightleftarrows} V_pP \\ & A \mapsto X_p^A. \end{aligned} \tag{6.5}$$

Now, as the connection 1-form acts on any tangent vector on P, the map i_p^{-1} in the definition ensures that only vertical parts get mapped, *i.e.* only those that belong to the induced vector field.

Definition 6.4. *A* ***Connection 1-form*** *is a map*

$$\begin{aligned} \omega_p : T_pP &\xrightarrow{\sim} T_eG \\ X_p \mapsto \omega_p(X_p) &:= i_p^{-1}(ver(X_p)). \end{aligned} \tag{6.6}$$

[34, 5]

As i_p^{-1} maps only vertical parts of vectors, any vector that lives on a whole in H_pP gets mapped to the null vector. This means that the horizontal subspace can be defined with the use of a connection 1-form [34, 5, 3, 12].

$$H_pP := ker(\omega_p) = \{X_p \in T_pP \mid \omega_p(X_p) = 0\}. \tag{6.7}$$

This means that vectors from a horizontal subspace "*do not*" get their corresponding vectors in Lie algebra, or rather, null vectors get assigned to them. The rest of them that do get mapped to vectors other than null vectors are then part of the vertical subspace. This is how a map ω_p replaces the first (and subsequently the third) condition in the definition of a connection. The last remaining condition is satisfied in the form of the following proposition.

Proposition 6.1. *Given a connection 1-form ω_p on a principal bundle we can say that the horizontal subspaces satisfy the following:*

$$(\triangleleft g)_* H_pP = H_{p\triangleleft g}P \tag{6.8}$$

[3]

Proof. Let $X_p \in H_pP$. Then, by definition, we have that $\omega_p(X_p) = 0$. If one were to push that vector with the right action, one would get vector $(\triangleleft g)_* X_p$. If we prove that $\omega_p((\triangleleft g)_* X_p) = 0$ we would also prove $(\triangleleft g)_* X_p \in H_{p\triangleleft g}P$.

$$\omega_p((\triangleleft g)_* X_p) \overset{\text{By def. of a pull-back.}}{=} (\triangleleft g)^* \underbrace{\omega_p(X_p)}_{=0} = 0 \tag{6.9}$$

□

As one might question why we were able to use the definition of a pull-back on connection 1-forms in the proof of the proposition (6.1), recall that the pull-back is defined via the push-forward, *i.e.* $\Phi^*\omega(X) := \omega(\Phi_* X)$. Notice how the output of a 1-form does not affect the definition. The pull-back is fully generalizable to the connection 1-forms [35]. That is how a single map can replace all the points the connection requires. On another note, a connection 1-form abides by the properties listed in the theorem below. Note that sometimes, the first two points of this theorem are presented as a definition of a connection 1-form [3, 7].

Theorem 6.1. *Connection 1-form ω_p satisfies the following*

(i) $\omega_p(X_p^A) = A$,

(ii) $\left((\lhd g)^*\omega\right)_p (X_p) = (Ad_{g^{-1}*})(\omega_p(X_p))$,

(iii) ω_p *is a smooth map.*

[34, 5, 12]

6.1 Local representation of a connection on a base manifold

The connection we have introduced is globally defined and the connection 1-form is defined pointwise. The next step is to compare the neighboring points in a local setting.

Because we have defined the connection 1-form pointwise and we are trying to get it to work for local sections, this is a good place to mention that the connection 1-form can be made into a ”*connection 1-form field*” by defining a tangent bundle of a principal G-bundle and taking its smooth section, similar to how the n-form field was defined. If one recalls that definition one of the names for an ”n-form field” was just the ”n-form”. Similarly, here, the ”*connection 1-form field*” will be referred to as just a *connection 1-form*. As a way to distinguish between the pointwise definition and a field definition, the latter will be denoted as ω. Note that none of this is formal terminology and is usually assumed to be understood.

So, a globally defined connection 1-form is then a map that maps from local sections of a "tangent bundle of a principal bundle" to its Lie-algebra-valued functions, *i.e.* $\omega : \Gamma(TP) \to C^\infty(P) \times T_eG$, where $C^\infty(P) := \{f \mid f : P \to \mathbb{R}\}$.[28] Let $\Gamma(TP) \to C^\infty(P) \times T_eG \equiv \Gamma(TP^*)$, so that we can write $\omega \in \Gamma(TP^*)$. Let $U \subseteq M$ be an open neighborhood on a base manifold M. Each local section $\sigma : U \subseteq M \to P$ induces a map $\sigma^* : \Gamma(TP^*) \to \Gamma(TU^*)$, where $\Gamma(TU^*) \equiv \Gamma(TU) \to C^\infty(U) \times T_eG$ and $C^\infty(U) := \{f \mid f : U \to \mathbb{R}\}$.

Recall that the connection 1-form was defined on a principal bundle. An induced map σ^* seemingly assumes the existence of a connection 1-form on a base manifold. However, as there is no $ver(X_p)$ on a base manifold, the new induced map we got can not be considered a connection 1-form. Even though, it still, very much is, a Lie-algebra-valued 1-form. Such a map will be labeled ω^U and called a **gauge potential** [20, 6, 3] or a **Yang-Mills field** [35, 5].

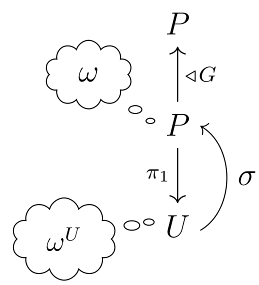
Definition 6.5. *Let $P \xrightarrow{\pi} M$ be a principal G-bundle. Each local section $\sigma : U \subseteq M \to P$ induces a **gauge potential** or a **Yang-Mills field** $\boldsymbol{\omega^U}$ defined as*

$$\begin{aligned} \omega^U :& \Gamma(TU) \to C^\infty(U) \times T_eG \\ \omega^U :=& \sigma^*\omega \end{aligned} \tag{6.10}$$

[35, 5, 3, 12]

Note that ω (and ω^U) in the diagram bellow *do not* actually "*live*" in P (and U) as this is symbolic representation. If one wishes to be precise, they live in $C^\infty(P) \times T_eG$ and $C^\infty(U) \times T_eG$ respectively. As those spaces can not really be presented in a "*bundle form*", we will provide the following picture to strengthen our intuition.

[28] Note that all of this is very much a non-formal notation.



(6.11)

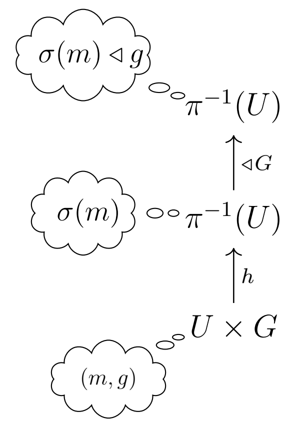
Because we are now interested in a local setting, we introduce a **local trivialization h**. It should be noted that this is exactly the same map that was labeled φ_j^{-1} in the chapter (2.2), only this time, it is defined specifically for the case of a principal bundle. As that is the case, we can now explicitly define how a local trivialization acts, as h can now move the points in the fiber with the use of the right action.

Definition 6.6. *Let $U \subseteq M$ be an open subset on a base manifold of a principal G-bundle. A* ***local trivialization of the principal bundle*** *is a map*

$$\begin{aligned} h : U \times G &\to \pi^{-1}(U) \\ (m, g) &\mapsto \sigma(m) \triangleleft g. \end{aligned} \tag{6.12}$$

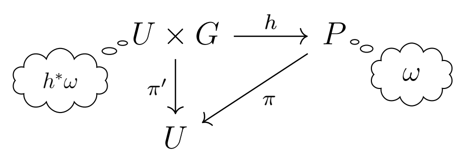
[35, 5]

Similar to the case of a gauge potential, we can symbolically visualize the action of a local trivialization h.



(6.13)

Similar to the case of a cross-section, local trivialization is then used to pull a connection 1-form to a local setting. That is done with an induced map $h^* : T_pP \to T_{(m,g)}(U \times G)$. Now, the pull-back of the connection 1-form will be visualized as shown in the following diagram.

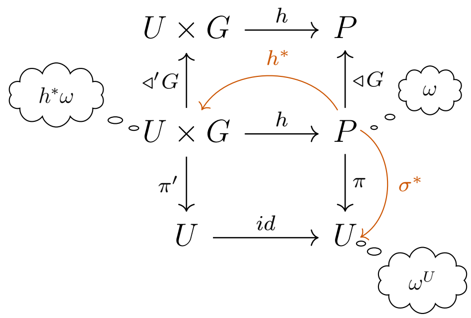


(6.14)

Here a map π' is a projection of a locally trivial bundle.

6.2 Alliance between a gauge potential and a local trivialization

We can draw the following commutative diagram together with the induced maps we have constructed. Note that such a diagram *does not* represent a bundle isomorphism, as it is impossible to construct a bijection between a total space and a local part of that same space.



(6.15)

Also, note that both $h^*\omega$ and ω^U *locally* contain all the information on the starting ω. Although one might feel that the ω^U on the base manifold lost some information due to the base manifold having less structure, that is in fact, not the case, as the following theorem connects the two.

Theorem 6.2. *Let* $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ *be a principal bundle equipped with the local section* σ *and a connection 1-form* ω*. Then the following holds*

$$(h^*\omega)_{(m,g)}(v,\gamma) = Ad_{g^{-1}*}(\omega^U_m(v)) + \Xi_g(\gamma), \tag{6.16}$$

for all $(m,g) \in U \times G$*,* $v \in T_mU$ *and* $\gamma \in T_gG$*,*
[35, 5]

where Ξ_g is the **Maurer-Cartan form** which we will now define.

Similar to the connection 1-form, Maurer-Cartan is also a Lie-algebra valued 1-form, only this time, instead of taking the vectors from T_pP, the Maurer-Cartan form takes vectors from T_gG. As those vectors are left-invariant, one can think of the Maurer-Cartan form as the embodiment of theorem (3.2), *i.e.* $L(G) \cong T_eG$.

$$\begin{aligned} \Xi_g :& T_gG \to T_eG \\ & l_{g*}A \mapsto A \end{aligned} \tag{6.17}$$

6.2.1 Preparation for the proof of the theorem (6.2)

All of the statements with their corresponding proofs given in this subsection can be found in [5].

As those properties will be used, it would be good to point out that the following holds:

$$(\omega^1 \circ \omega^2)^* = \omega^{2*}\omega^{1*} \qquad \text{and} \qquad (v_1 \circ v_2)_* = v_{1*}v_{2*} \tag{6.18}$$

where ω^1 and ω^2 represent any objects that get pulled, and v_1 and v_2 any objects that get pushed.

Although the following theorem might seem obvious, the maps defined to prove it will be used in the future.

Theorem 6.3. *On a product manifold $M \times N$ there exists a natural isomorphism, i.e.*

$$\begin{gathered} T_{(p,q)}M \times N \cong T_pM \oplus T_qN \\ v \mapsto (pr_{1*}v, pr_{2*}v), \end{gathered} \tag{6.19}$$

where

$$\begin{aligned} pr_1 : M \times N &\to M \\ (p,q) &\mapsto p \end{aligned} \qquad \textit{and} \qquad \begin{aligned} pr_2 : M \times N &\to N \\ (p,q) &\mapsto q \end{aligned} \tag{6.20}$$

Proof. We construct the following map

$$\chi : T_{(p,q)}M \times N \to T_pM \oplus T_qN. \tag{6.21}$$

As this map acts on tangent vectors, we should define curves they are tangent to. Let $(-\epsilon, \epsilon) \subseteq \mathbb{R}$. We define curves

$$\begin{aligned} \sigma &: (-\epsilon,\epsilon) \to M \times N \\ t &\mapsto \sigma(t) \end{aligned} \quad , \quad \begin{aligned} \sigma_1 &: (-\epsilon,\epsilon) \to M \\ \sigma_1(t) &:= pr_1 \circ \sigma(t) \end{aligned} \quad \text{and} \quad \begin{aligned} \sigma_2 &: (-\epsilon,\epsilon) \to N \\ \sigma_2(t) &:= pr_2 \circ \sigma(t) \end{aligned} \quad . \tag{6.22}$$

Let $v \in T_{(p,q)} M \times N$ be tangent to the curve $\sigma(t)$, and let vectors $v_1 \in T_p M$ and $v_2 \in T_q N$ be tangent to the curves σ_1 and σ_2 respectively. This now means that we can write

$$\chi(v) = (v_1, v_2). \tag{6.23}$$

If we prove that $(v_1, v_2) = ((pr_{1*}v, pr_{2*}v))$ then $\chi(v) = (pr_{1*}v, pr_{2*}v)$. As both pr_1 and pr_2 are just projections, such a map would then clearly be bijective. As the proofs for v_1 and v_2 are almost identical, here, we will provide proof just for the v_1.

$v_1 \stackrel{?}{=} pr_{1*}v$

Let $f : M \to \mathbb{R}$. By definition it then holds

$$v_1 f \equiv \frac{d}{dt}(f \circ \sigma_1(t))\bigg|_{t=0} . \tag{6.24}$$

On the other hand

$$(pr_{1*}v)f = v(f \circ pr_1) \equiv \frac{d}{dt}(f \circ \underbrace{pr_1 \circ \sigma(t)}_{=\sigma_1(t)})\bigg|_{t=0} = \frac{d}{dt}(f \circ \sigma_1(t))\bigg|_{t=0} = v_1 f \tag{6.25}$$

□

We will also define some useful injections.

$$\begin{aligned} i_q &: M \to M \times N \\ x &\mapsto (x, q) \end{aligned} \quad \text{and} \quad \begin{aligned} j_p &: N \to M \times N \\ x &\mapsto (p, x) \end{aligned} \quad . \tag{6.26}$$

Together with predefined projections, the following relations hold

$$
\begin{aligned}
pr_1 \circ i_q &= id_M, \\
pr_1 \circ j_p &= p
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
pr_2 \circ j_p &= id_N, \\
pr_2 \circ I_q &= q.
\end{aligned}
\tag{6.27}
$$

As a map χ is isomorphic, we define its inverse as the following

$$
\begin{aligned}
\chi^{-1} : T_pM \oplus T_qN &\to T_{(p,q)}M \times N \\
(\alpha, \beta) &\mapsto i_{q*}\alpha + j_{p*}\beta
\end{aligned}
\tag{6.28}
$$

Proposition 6.2. *χ^{-1} is a well-defined map.*

Proof. We know that $v_1 = pr_{1*}v$. As $v \in T_{(p,q)}M \times N$ let $v = i_{q*}\alpha + j_{p*}\beta$. What is $pr_{1*}(v)$?

Push-forward is a linear map.

$$
\begin{aligned}
pr_{1*}(v) = pr_{1*}(i_{q*}\alpha + j_{p*}\beta) &= pr_{1*}(i_{q*}\alpha) + pr_{1*}(j_{p*}\beta) \\
&= (pr_1 \circ j_p)_*\beta + (pr_1 \circ i_q)_*\alpha = id_M\alpha + \cancelto{0}{p_*}\beta = \alpha \in T_pM
\end{aligned}
\tag{6.29}
$$

Points do not get pushed.

Similarly, $pr_{2*}(v) = \beta \in T_qN$. ∎

As there is an isomorphism $T_{(p,q)}M \times N \cong T_pM \oplus T_qN$, the maps χ and χ^{-1} will be assumed to be identity maps, meaning we will write

$$
(\alpha, \beta) = i_{q*}\alpha + j_{p*}\beta. \tag{6.30}
$$

6.3 Proof of theorem (6.2)

Note that the proof of a theorem (6.2) can be found in [5, p. 256-257].

Theorem 6.2. *Let* $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ *be a principal bundle equipped with the local section* σ *and a connection 1-form* ω*. Then the following holds*

$$(h^*\omega)_{(m,g)}(v,\gamma) = Ad_{g^{-1}*}(\omega^U_m(v)) + \Xi_g(\gamma), \tag{6.31}$$

for all $(m,g) \in U \times G$*,* $v \in T_mU$ *and* $\gamma \in T_gG$*.*
[35, 5]

Proof. We can factorize the local trivialization in the following manner[29]

$$\begin{array}{rcccc} h : U \times G & \xrightarrow{\sigma \times id_G} & P \times G & \xrightarrow{\triangleleft} & P \\ (m,g) & \mapsto & (\sigma(m), g) & \mapsto & \sigma(m) \triangleleft g \end{array} \tag{6.32}$$

[29]Note that in the above factorization, the clouds with the connection 1-form and all its pull-backs have been added so the proof may be easier to follow. Similar to when such visualization was introduced, the spaces these forms inhabit are not actually the spaces the clouds point to but are assumed to be understood. The usefulness of this notation comes into play when one needs to think about the specific points ω is defined at.

$$h^*\omega_{(m,g)}(v,\gamma) = \left((\triangleleft \circ (\sigma \times id_G))^* \omega\right)_{(m,g)}(v,\gamma) \underset{\substack{(\omega^1 \circ \omega^2)^* \\ = \omega^{2*}\omega^{1*}.}}{=} \left((\sigma \times id_G)^* \triangleleft^* \omega\right)_{(m,g)}(v,\gamma)$$

$$\underset{\substack{\Phi^*\omega(X) \\ := \omega(\Phi_* X).}}{=} (\triangleleft^*\omega)_{(\sigma(m),g)}\left((\sigma \times id_G)_*(v,\gamma)\right) = (\triangleleft^*\omega)_{(\sigma(m),g)}(\sigma_*(v),\gamma)$$

$$\underset{\substack{(\alpha,\beta)= \\ i_{q*}\alpha + j_{p*}\beta.}}{=} (\triangleleft^*\omega)_{(\sigma(m),g)}(i_{g*}\sigma_*(v) + j_{\sigma(m)*}\gamma) \underset{\substack{\Phi^*\omega(X) \\ := \omega(\Phi_* X).}}{=} \omega_{\sigma(m)\triangleleft g}\left(\triangleleft_*\left(i_{g*}\sigma_*(v) + j_{\sigma(m)*}\gamma\right)\right)$$

$$= \omega_{\sigma(m)\triangleleft g}\left((\triangleleft \circ i_g)_*\sigma_*(v) + (\triangleleft \circ j_{\sigma(m)})_*\gamma\right) \tag{6.33}$$

Now follows the question of calculating $(\triangleleft \circ i_g)_*$ and $(\triangleleft \circ j_{\sigma(m)})_*$. For that task to be easier, let us first write out how i_g and $j_{\sigma(m)}$ map.

$$\begin{aligned} i_g &: P \to P \times G \\ p &\mapsto (p.g) \end{aligned} \qquad \text{and} \qquad \begin{aligned} j_{\sigma(m)} &: G \to P \times G \\ g &\mapsto (\sigma(m), g) \end{aligned} \tag{6.34}$$

(i) First term, *i.e.* the term containing the map $(\triangleleft \circ i_g)_*$.

Let $x_p \in T_pP$. Compare how the following two maps act on X_p

$$\begin{aligned} &\triangleleft \circ i_g : P \to P \\ &(\triangleleft \circ i_g)(p) = \triangleleft(p,g) = p \triangleleft g \\ &\implies (\triangleleft \circ i_g)_*(x_p) = \triangleleft_*(i_{g*}(x_p)) \\ &= \triangleleft_*(x_{(p,g)}) = x_{p\triangleleft g} \end{aligned} \qquad \text{and} \qquad \begin{aligned} &\triangleleft : P \times G \to P \\ &\triangleleft g : P \to P \\ &(\triangleleft g)(p) = g \triangleleft p \\ &\implies (\triangleleft g)_*(x_p) \coloneqq x_{p\triangleleft g} \end{aligned} \quad . \tag{6.35}$$

Meaning that we can write $(\triangleleft \circ i_g)_* \coloneqq (\triangleleft g)_*$.

$$\omega_{\sigma(m)\triangleleft g}\left((\triangleleft \circ i_g)_* \sigma_*(v)\right) = \omega_{\sigma(m)\triangleleft g}\left((\triangleleft g)_* \sigma_*(v)\right) \underset{\swarrow}{\overset{\substack{\Phi^*\omega(X) \\ :=\omega(\Phi_* X)}}{=}} \left((\triangleleft g)^* \omega_{\sigma(m)\triangleleft g}\right) \sigma_*(v)$$

$$= \left((\triangleleft g)^* \omega\right)_{\sigma(m)} \sigma_*(v) \underset{\substack{\nearrow \\ ((\triangleleft g)^*\omega)_p(x_p)= \\ (Ad_{g^{-1}*})(\omega_p(x_p)).}}{=} (Ad_{g^{-1}*})(\omega_{\sigma(m)}(\sigma_* v)) \underset{\substack{\nearrow \\ \omega(\Phi_* X) \\ :=\Phi^*\omega(X).}}{=} (Ad_{g^{-1}*})\left(\sigma^* \omega_{\sigma(m)}(v)\right) \quad (6.36)$$

If one recalls how gauge potential is defined, this term can be written as

$$(Ad_{g^{-1}*})\left(\sigma^* \omega_{\sigma(m)}(v)\right) = (Ad_{g^{-1}*})\left(\omega^U_m(v)\right). \quad (6.37)$$

(ii) Second term, *i.e.* the term containing the map $(\triangleleft \circ j_{\sigma(m)})_*$.

Let $P_{\sigma(m)} := \triangleleft \circ j_{\sigma(m)}$. How such a map acts is easily obtained by looking at how $(\triangleleft \circ j_{\sigma(m)})_*$ acts,

$$\begin{aligned} &\triangleleft \circ j_{\sigma(m)} : G \to P \\ &(\triangleleft \circ j_{\sigma(m)})(g) = \triangleleft(\sigma(m), g) = \sigma(m) \triangleleft g \end{aligned} \qquad \text{and} \qquad \begin{aligned} &P_{\sigma(m)} : G \to P \\ &g \mapsto \sigma(m) \triangleleft g \end{aligned} \quad (6.38)$$

In general, we can write that $P_{\sigma(m)} \circ l_g = P_{\sigma(m)\triangleleft g}$.[30] Notice that $\gamma \in T_e G$ is a left-invariant vector meaning we can write $P_{\sigma(m)*}\gamma = P_{\sigma(m)*}(l_{g*}\gamma)$.

30

Let $h \in G$.

$\left(P_{\sigma(m)} \circ l_g\right)(h) = P_{\sigma(m)}(l_g(h)) = P_{\sigma(m)}(g \bullet h) = \sigma(m) \triangleleft (g \bullet h);$

$P_{\sigma(m)\triangleleft g}(h) = \sigma(m) \triangleleft g \triangleleft h \underset{\substack{\nearrow \\ \text{By def. of a right} \\ \text{action.}}}{=} \sigma(m) \triangleleft (g \bullet h)$

This means we can write

$$(\triangleleft \circ j_{\sigma(m)})_* \gamma = P_{\sigma(m)*}\gamma = P_{\sigma(m)*}(l_{g*}\gamma)$$
$$= (P_{\sigma(m)} \circ l_g)_* \gamma = P_{(\sigma(m)\triangleleft g)*}\gamma \tag{6.39}$$

Recall how connection 1-form is defined

$$\omega_{\sigma(m)\triangleleft g}\left(P_{(\sigma(m)\triangleleft g)*}\gamma\right) \equiv i^{-1}_{\sigma(m)\triangleleft g}\left(ver(P_{(\sigma(m)\triangleleft g)*}\gamma)\right) \tag{6.40}$$

By def. of a connection 1-form.

As $\omega_{\sigma(m)\triangleleft g}$ acts only on vertical bits of $P_{(\sigma(m)\triangleleft g)*}\gamma$, we can safely write that $P_{(\sigma(m)\triangleleft g)*}\gamma = X^{\gamma}_{\sigma(m)\triangleleft g}$, where $X^{\gamma}_{\sigma(m)\triangleleft g}$ is an induced vector field. By the first point of the theorem (6.1), *i.e.* $\omega_p(X^A_p) = A$, we can write the following

$$\omega_{\sigma(m)\triangleleft g}\left(P_{(\sigma(m)\triangleleft g)*}\right) = \omega_{\sigma(m)\triangleleft g}\left(X^{\gamma}_{\sigma(m)\triangleleft g}\right) = \gamma = \Xi_g(\gamma).^{31} \tag{6.41}$$

Altogether,

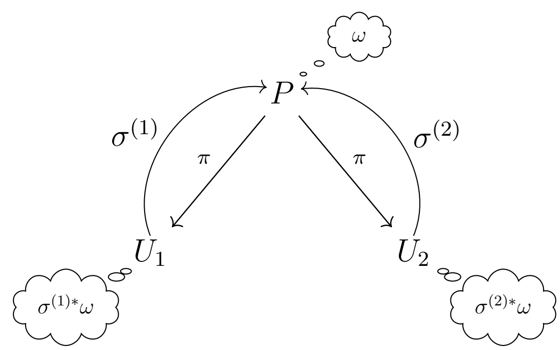
$$h^*\omega_{(m,g)}(v,\gamma) = (Ad_{g^{-1}*})\left(\omega^U_m(v)\right) + \Xi_g(\gamma). \tag{6.42}$$

□

[31] Recall how Maurer-Cartan form was defined
$\Xi_g : T_gG \to T_eG$
$l_{g*}A \mapsto A$
In our case this means
$\Xi_g(l_{g*}\gamma) = \Xi_g(\gamma) = \gamma$.

6.4 Gauge potentials on the intersection of open neighborhoods of a base manifold

The following explanation relies on [35] while proof can be found in [5, p. 259]. Let $U_1, U_2 \subseteq M$ be open neighborhoods such that $U_1 \cap U_2 \neq \emptyset$. We require compatibility between gauge potentials defined at such an intersection.



(6.43)

We want to look at how $\sigma^{(1)*}\omega$ and $\sigma^{(2)*}\omega$ relate to each other. Cross sections from the diagram map in the following manner

$$\sigma^{(1)} : U_1 \to P \qquad \text{and} \qquad \sigma^{(2)} : U_2 \to P \quad . \tag{6.44}$$

We introduce a transition function $\Omega : U_1 \cap U_2 \to G$. Such a map is uniquely defined by the following

$$\sigma^{(2)} \coloneqq \sigma^{(1)} \triangleleft \Omega(m), \tag{6.45}$$

for some $m \in U_1 \cap U_2$. As $\Omega(m) \in G$ it can act with the right action. This equation will prove to be useful in finding the wanted compatibility between gauge potentials. As such compatibility is found in the form of a theorem, let us first state the theorem.

Theorem 6.4. *Let $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ be a principal bundle equipped with the connection 1-form ω. On the intersection $U_1 \cap U_2 \subseteq M$ the following holds*

$$\omega_m^{U_2}(v) = Ad_{\Omega^{-1}(m)*}\left(\omega_m^{U_1}(v)\right) + \left(\Omega^*\Xi\right)_m(v) \tag{6.46}$$

for all $m \in U_1 \cap U_2$ and $v \in T_m(U_1 \cap U_2)$.
[35, 5, 12, 7]

Proof. Similar to the case of a theorem (6.2), we begin by the factorization of a map. Only, this time, the map in question is $\sigma^{(2)}$.

$$\begin{array}{ccccc} \overset{\sigma^{(2)*}\omega}{\sigma^{(2)} : U_1 \cap U_2} & \xrightarrow{\sigma^{(1)}\times\Omega} & \overset{\triangleleft^*\omega}{P \times G} & \xrightarrow{\triangleleft} & \overset{\omega}{P} \\ m & \mapsto & (\sigma^{(1)}(m), \Omega(m)) & \mapsto & \sigma^{(1)}(m) \triangleleft \Omega(m) \end{array} \tag{6.47}$$

$$(\sigma^{(2)*}\omega)_m(v) = \left(\left(\triangleleft \circ \left(\sigma^{(1)} \times \Omega\right)\right)^* \omega\right)_m v \underset{\substack{(\omega^1 \circ \omega^2)^* \\ = \omega^{2*}\omega^{1*}.}}{=} \left(\left(\sigma^{(1)} \times \Omega\right)^* \triangleleft^* \omega\right)_m v$$

$$\underset{\substack{\Phi^*\omega(X) \\ :=\omega(\Phi_* X).}}{=} (\triangleleft^*\omega)_{(\sigma^{(1)}(m),\Omega(m))}\left((\sigma^{(1)} \times \Omega)_* v\right) = (\triangleleft^*\omega)_{(\sigma^{(1)}(m),\Omega(m))}\left(\sigma_*^{(1)}v, \Omega_* v\right)$$

$$\underset{\substack{(\alpha,\beta)= \\ i_{q*}\alpha + j_{p*}\beta.}}{=} (\triangleleft^*\omega)_{(\sigma^{(1)}(m),\Omega(m))}\left(i_{\Omega(m)*}\sigma_*^{(1)}v + j_{\sigma^{(1)}(m)*}\Omega_* v\right)$$

$$\underset{\substack{\Phi^*\omega(X) \\ :=\omega(\Phi_* X).}}{=} \omega_{\sigma^{(1)}(m)\triangleleft\Omega(m)}\left(\triangleleft_*\left(i_{\Omega(m)*}\sigma_*^{(1)}v + j_{\sigma^{(1)}(m)*}\Omega_* v\right)\right)$$

$$= \omega_{\sigma^{(1)}(m)\triangleleft\Omega(m)}\left((\triangleleft \circ i_{\Omega(m)})_*\sigma_*^{(1)}v + (\triangleleft \circ j_{\sigma^{(1)}(m)})_*\Omega_* v\right) \tag{6.48}$$

Similar to the proof of theorem (6.2), we used maps i_q and j_p, where $p \in M$ and $q \in N$. As we are dealing with $P \times G$ as opposed to $M \times N$, it is good to spell out how they act in this case.

$$\begin{aligned} &i_{\Omega(m)} : P \to P \times G \\ &p \mapsto (p.\Omega(m)) \end{aligned} \quad \text{and} \quad \begin{aligned} &j_{\sigma^{(1)}(m)} : G \to P \times G \\ &\Omega(m) \mapsto (\sigma^{(1)}(m), \Omega(m)) \end{aligned}, \tag{6.49}$$

(i) The first term, *i.e.* the term containing a map $\triangleleft \circ i_{\Omega(m)}$.

Compare how the two following maps act for some $x_p \in T_pP$.

$$\begin{aligned} &\triangleleft \circ i_{\Omega(m)} : P \to P \\ &(\triangleleft \circ i_{\Omega(m)})(p) = \triangleleft(p, \Omega(m)) = p \triangleleft \Omega(m) \\ &\implies (\triangleleft \circ i_g)_*(x_p) = \triangleleft_*(i_{\Omega(m)*}(x_p)) \\ &= \triangleleft_*(x_{(p,\Omega(m))}) = x_{p \triangleleft \Omega(m)} \end{aligned} \quad \text{and} \quad \begin{aligned} &\triangleleft : P \times G \to P \\ &\triangleleft(\underbrace{\Omega(m)}_{\in G}) : P \to P \\ &(\triangleleft(\Omega(m)))_* (x_p) = x_{p \triangleleft \Omega(m)} \end{aligned} . \tag{6.50}$$

Meaning that we can write $(\triangleleft \circ i_{\Omega(m)})_* := (\triangleleft(\Omega(m)))_*$.

$$\begin{aligned} &\omega_{\sigma^{(1)}(m) \triangleleft \Omega(m)} \left((\triangleleft \circ i_{\Omega(m)})_* \sigma_*^{(1)} v \right) = \omega_{\sigma^{(1)}(m) \triangleleft \Omega(m)} \left((\triangleleft(\Omega(m)))_* \sigma_*^{(1)} v \right) \\ &\underset{\Phi^*\omega(X) := \omega(\Phi_* X)}{=} \left((\triangleleft(\Omega(m)))^* \omega_{\sigma^{(1)}(m) \triangleleft \Omega(m)} \right) \sigma_*^{(1)} v = \left((\triangleleft(\Omega(m)))^* \omega \right)_{\sigma^{(1)}(m)} \sigma_*^{(1)} v \\ &\underset{((\triangleleft g)^*\omega)_p(x_p) = (Ad_{g^{-1}*})(\omega_p(x_p)).}{=} (Ad_{\Omega(m)^{-1}*}) \left(\omega_{\sigma^{(1)}(m)} (\sigma_*^{(1)} v) \right) \underset{\omega(\Phi_* X) := \Phi^*\omega(X).}{=} (Ad_{\Omega(m)^{-1}*}) \left(\sigma^{(1)*} \omega_{\sigma^{(1)}(m)}(v) \right) \end{aligned} \tag{6.51}$$

Identical to the proof of theorem (6.2), with the use of a definition of a gauge potential, this can be written as

$$(Ad_{\Omega(m)^{-1}*}) \left(\sigma^{(1)*} \omega_{\sigma^{(1)}(m)}(v) \right) = Ad_{\Omega(m)^{-1}*} \left(\omega_m^{U_1}(v) \right). \tag{6.52}$$

(ii) The first term, *i.e.* the term containing a map $\triangleleft \circ j_{\sigma^{(1)}(m)}$.

Let $P_{\sigma^{(1)}(m)} \coloneqq \triangleleft \circ j_{\sigma^{(1)}(m)}$. We can, again, conclude how $P_{\sigma^{(1)}(m)}$ acts

$$\begin{aligned} \triangleleft \circ j_{\sigma^{(1)}(m)} &: G \to P \\ (\triangleleft \circ j_{\sigma^{(1)}(m)}(\Omega(m)) & \\ = \triangleleft(j_{\sigma^{(1)}(m)}, \Omega(m)) & \\ = \sigma^{(1)}(m) \triangleleft \Omega(m) & \end{aligned} \implies \begin{aligned} P_{\sigma^{(1)}(m)} &: G \to P \\ \Omega(m) &\mapsto \sigma^{(1)}(m) \triangleleft \Omega(m) \end{aligned}. \tag{6.53}$$

As $\Omega_* v \in T_e G$ we can write $l_{\Omega(m)*}(\Omega_* v) = \Omega_* v$. In other words $P_{\sigma^{(1)}(m)*}(\Omega_* v) = P_{\sigma^{(1)}(m)*}(l_{\Omega(m)*}(\Omega_* v)) = P_{(\sigma^{(1)}(m)\triangleleft\Omega(m))*}\Omega_* v$.[32] We again recall how connection 1-form is defined

$$\omega_{\sigma^{(1)}(m)\triangleleft\Omega(m)} \left(P_{(\sigma^{(1)}(m)\triangleleft\Omega(m))*}\Omega_* v\right) \underset{\nearrow}{\equiv} i^{-1}_{\sigma^{(1)}(m)\triangleleft\Omega(m)} \left(ver(P_{(\sigma^{(1)}(m)\triangleleft\Omega(m))*}(\Omega_* v)\right) \tag{6.54}$$

By def. of a connection 1-form.

As inputs of a connection 1-form are vertical vectors, we can write $P_{(\sigma^{(1)}(m)\triangleleft\Omega(m))*}(\Omega_* v) = X^{\Omega_* v}_{\sigma^{(1)}(m)\triangleleft\Omega(m)}$, where $X^{\Omega_* v}_{\sigma^{(1)}(m)\triangleleft\Omega(m)}$ is induced vector field. By the first point of the theorem (6.1), *i.e.* $\omega_p(X^A_p) = A$, we can write the following

$$\begin{aligned} &\omega_{\sigma^{(1)}(m)\triangleleft\Omega(m)} \left(P_{(\sigma^{(1)}(m)\triangleleft\Omega(m))*}\Omega_* v\right) = \omega_{\sigma^{(1)}(m)\triangleleft\Omega(m)} \left(X^{\Omega_* v}_{\sigma^{(1)}(m)\triangleleft\Omega(m)}\right) = \Omega_* v \\ &= \Xi_{\Omega(m)}(\Omega_* v) \underset{\nearrow}{=} \Omega^* \Xi_{\Omega(m)}(v) = \left(\Omega^* \Xi\right)_m (v) \end{aligned} \tag{6.55}$$

$\omega(\Phi_* X) \coloneqq \Phi^* \omega(X)$.

One last thing to cover is a starting map, which can be re-written with a

[32] Recall that $P_{\sigma^{(1)}(m)} \circ l_{\Omega(m)} = P_{\sigma^{(1)}(m)\triangleleft\Omega(m)}$.

definition of a gauge potential.

$$(\sigma^{(2)*}\omega)_m(v) = \omega_m^{U_2}(v) \tag{6.56}$$

Altogether, we have

$$\omega_m^{U_2}(v) = Ad_{\Omega^{-1}(m)*}\left(\omega_m^{U_1}(v)\right) + \left(\Omega^*\Xi\right)_m(v). \tag{6.57}$$

□

6.4.1 Theorem (6.4) when G is a matrix group

The following can be found in [35, 5]. If G is a matrix group we can rewrite this equation in a form that may be more familiar to physicists. Since we have already shown how an adjoint map looks like for matrix groups in subsection (3.2), we can just write $Ad_{\Omega^{-1}(m)*}\left(\omega_m^{U_1}(v)\right) = \Omega(m)^{-1}\omega_m^{U_1}(v)\Omega(m)$. What is left to show is the pull-back of Maurer-Cartan's form for matrix groups. As that will be proven component-wise, it would be good to write down how the left-invariant vector field and Maurer-Cartan form look like in component notation.

Coordinates of matrix entries will be defined as follows $x^i_j(g) = g^i_j$, where g is an element of a matrix group G. Here, one can think of x^i_j as a map that outputs a (i,j)-th component of a matrix g, *i.e.* $x^i_j : G \to \mathbb{R}$.

As, in a sense, $x^i_j \in C^\infty(G)$, we can act on it with left-invariant vector field $l_{g*}A$.

$$
\begin{aligned}
(l_{g*}A)(x^i_j)_g &\coloneqq \frac{d}{dt}\left(x^i_j(g \bullet e^{tA})\right)\Big|_{t=0} = \frac{d}{dt}\left((g \bullet e^{tA})^i_j\right)\Big|_{t=0} = \frac{d}{dt}\left(g^i_k (e^{tA})^k_j\right)\Big|_{t=0} \\
&= g^i_k \frac{d}{dt}\left((e^{tA})^k_j\right)\Big|_{t=0} = g^i_k A^k_l (e^{tA})^l_j\Big|_{t=0} = g^i_k A^k_j \\
&\iff l_{g*}A = g^i_k A^k_j \left(\frac{\partial}{\partial x^i_j}\right)_g \quad \Big/ \ \Xi_g \\
&\Xi_g(l_{g*}A) = \Xi_g\left(g^i_k A^k_j \left(\frac{\partial}{\partial x^i_j}\right)_g\right) = A \\
&\Xi^i_{gj}(l_{g*}A) = A^i_j
\end{aligned} \tag{6.58}
$$

For the last equation to hold true, the Maurer-Cartan form has to be of the form $\Xi^i_{gj} \coloneqq (g^{-1})^i_k dx^k_j$ so that

$$
\begin{aligned}
\Xi^i_{gj}(l_{g*}A) &= (g^{-1})^i_k dx^k_j \left(g^p_r A^r_q \left(\frac{\partial}{\partial x^p_q}\right)_g\right) \\
&= (g^{-1})^i_k g^p_r A^r_q \delta^k_p \delta^q_j = (g^{-1})^i_p g^p_r A^r_j \\
&= \delta^i_r A^r_j = A^i_j.
\end{aligned} \tag{6.59}
$$

The coordinate definition of Maurer-Cartan form, $\Xi^i_{gj} \coloneqq (g^{-1})^i_k dx^k_j$, will be used in writing the term $(\Omega^*\Xi)_m(v)$ in matrix form. The arbitrary vector v will be substituted with a basis vector $\left(\frac{\partial}{\partial x^\mu}\right)_m$.

$$
\begin{aligned}
(\Omega^*\Xi)^i_{mj}\left(\left(\frac{\partial}{\partial x^\mu}\right)_m\right) &= \Xi^i_{\Omega(m)j}\left(\Omega_*\left(\frac{\partial}{\partial x^\mu}\right)_m\right)_{\Omega(m)} \\
&\underset{\Xi^i_{gj}:=(g^{-1})^i_k dx^k_j.}{=} (\Omega(m)^{-1})^i_k dx^k_j\left(\Omega_*\left(\frac{\partial}{\partial x^\mu}\right)_m\right)_{\Omega(m)} \\
&\underset{df(v)=v(f).}{=} (\Omega(m)^{-1})^i_k\left(\Omega_*\left(\frac{\partial}{\partial x^\mu}\right)_m\right)_{\Omega(m)}(x^k_j) \\
&= (\Omega(m)^{-1})^i_k\left(\frac{\partial}{\partial x^\mu}\right)_m\left(x^k_j(\Omega)\right)_m \\
&= (\Omega(m)^{-1})^i_k\left(\frac{\partial}{\partial x^\mu}\right)_m(\Omega(m))^k_j
\end{aligned}
\tag{6.60}
$$

$$
\begin{aligned}
&\iff (\Omega^*\Xi)^i_{mj} = (\Omega(m)^{-1})^i_k\underbrace{\frac{\partial(\Omega(m))^k_j}{\partial x^\mu}dx^\mu}_{:=d\Omega(p)}. \\
&\iff (\Omega^*\Xi)_m = \Omega(m)^{-1}d\Omega(m).
\end{aligned}
\tag{6.61}
$$

All together, we can write

$$\omega^{U_2}_m(\partial_\mu) = \Omega(m)^{-1}\omega^{U_1}_m(\partial_\mu)\Omega(m) + \Omega(m)^{-1}\partial_\mu\Omega(m). \tag{6.62}$$

Or, equally validly we can write

$$\omega^{U_2}_m = \Omega(m)^{-1}\omega^{U_1}_m\Omega(m) + \Omega(m)^{-1}d\Omega(m). \tag{6.63}$$

There exists another notation that may be more familiar. If we substitute

$A^2_\mu(m) = \omega^{U_2}_m(\partial_\mu)$, the equation (6.62) can be re-written as

$$A^2_\mu(m) = \Omega(m)^{-1} A^1_\mu(m)\Omega(m) + \Omega(m)^{-1}\partial_\mu\Omega(m). \tag{6.64}$$

6.4.2 Passive vs. active gauge transformations

The following explanation can be found in [6, 5]. The equation (6.64) in all of its equivalent forms is what is considered a *passive gauge transformation*. The change between the gauge potentials that occurs as a consequence of looking at different open neighborhoods at the same intersection is what is considered a passive gauge transformation. However, as gauge potentials are ultimately dependent on connection 1-forms from the principal bundle, one can look at gauge potentials obtained by pulling two different connection 1-forms from the same fiber to the base manifold. Such an approach is then considered an *active gauge transformation*, and as it turns out, both ways of thinking result in an equivalent formula.

Even though it took us a lot of work to get to the gauge transformation equation, we have finally arrived at equations that look familiar to physicists as they can be found throughout gauge theories in slightly different forms. After deriving the case for the active gauge transformation, we will provide two physical examples of its use - one that describes a passive gauge transformation, another describing the active one.

To really see that both ways of thinking result in the same equation, we should derive the equation for an active gauge transformation. As we are interested in different points of the same fiber, we will be looking at the principal bundle automorphisms. That is formally done with the use of principal bundle maps, which we introduced in subsection (4.5.2) by substituting

$P' = P$ and $\triangleleft G' = \triangleleft G$.

$$\begin{array}{ccc} P & \xrightarrow{\phi} & P \\ {\scriptstyle \triangleleft G}\uparrow & & \uparrow{\scriptstyle \triangleleft G} \\ P & \xrightarrow{\phi} & P \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow[h]{} & M \end{array}, \qquad \begin{aligned} h \circ \pi &= \pi \circ \phi \\ \phi(p \triangleleft g) &= \phi(p) \triangleleft g. \end{aligned} \tag{6.65}$$

So, our principal bundle automorphism is a map $\phi : P \to P$ for which we can write $\phi(p \triangleleft g) = \phi(p) \triangleleft g$. To get to the gauge potentials, we also need a cross-section $\sigma(m)$, so we will be writing $p = \sigma(m)$. The equation for the principal bundle map automorphism is now of the form $\phi(\sigma(m) \triangleleft g) = \phi(\sigma(m)) \triangleleft g$.

Similar to the case of passive gauge transformation, we will be introducing a map $\Omega : U \to G$, only this time, it does not map from the intersection as we are no longer interested in that. As $\Omega(m) \in G$ we can further re-write the principal bundle automorphism as $\phi(\sigma(m) \triangleleft \Omega(m)) = \phi(\sigma(m)) \triangleleft \Omega(m)$. As $\triangleleft$ is a free action by virtue of being defined on a principal bundle, there has to exist some $\Omega(m)$ such that $\phi(\sigma(m) \triangleleft \Omega(m)) = \phi(\sigma(m)) \triangleleft \Omega(m) = \sigma(m)$. The equation $\phi \circ (\sigma(m) \triangleleft \Omega(m)) = \sigma(m)$ ensures that no matter how an automorphism ϕ acts on some point $\sigma(m) \triangleleft \Omega(m)$, we will still stay in the same fiber as our original point.

The last thing left to comment on before deriving an expression for an active gauge transformation is the notation. The following names will be used for gauge potentials $A \coloneqq \sigma^*\omega$ and $A' \coloneqq \sigma^*(\phi^*\omega) = (\phi \circ \sigma)^*\omega$.

$$\begin{array}{ccccccc} \sigma^*\omega & & \lhd^*(\phi^*\omega) & & \phi^*\omega & & \omega \\ \sigma : U & \xrightarrow{\sigma\times\Omega} & P \times G & \xrightarrow{\lhd} & P & \xrightarrow{\phi} & P \\ m & \mapsto & (\sigma(m),\Omega(m)) & \mapsto & \sigma(m)\lhd\Omega(m) & \mapsto & \phi\circ(\sigma(m)\lhd\Omega(m)) \\ & & & & & & =\sigma(m) \end{array} \tag{6.66}$$

$$\begin{aligned} (\sigma^*\omega)_m(v) &= ((\phi\circ\lhd\circ(\sigma\times\Omega))^*\omega)_m(v) = ((\sigma\times\Omega)^*\lhd^*(\phi^*\omega))_m(v) \\ &= (\lhd^*(\phi^*\omega))_{(\sigma(m),\Omega(m))}((\sigma\times\Omega)_*v) = (\lhd^*(\phi^*\omega))_{(\sigma(m),\Omega(m))}(\sigma_*v,\Omega_*v) \\ &= (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}(\lhd_*(\sigma_*v,\Omega_*v)) \\ &= (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}\left(\lhd_*\left(i_{\Omega(m)*}\sigma_*v + j_{\sigma(m)*}\Omega_*v\right)\right) \\ &= (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}\left((\lhd\circ i_{\Omega(m)})_*\sigma_*v + (\lhd\circ j_{\sigma(m)})_*\Omega_*v\right) \\ &= (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}(\lhd(\Omega(m)))_*\sigma_*v + (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}P_{\sigma(m)*}\Omega_*v \end{aligned} \tag{6.67}$$

As the equation above could get really crowded, the two obtained terms will be expanded upon separately.

(i) First term:

$$\begin{aligned} (\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}(\lhd(\Omega(m)))_*\sigma_*v &= \omega_{\phi\circ(\sigma(m)\lhd\Omega(m))}(\phi_*(\lhd(\Omega(m)))_*(\sigma_*v)) \\ &= \omega_{\sigma(m)}((\phi\circ(\lhd\Omega(m)))_*\sigma_*v) \\ &= ((\phi\circ(\lhd\Omega(m))^*\omega)_{(\sigma(m),\Omega(m))}(\sigma_*v) \\ &= (\lhd(\Omega(m)))^*(\phi^*\omega)_{\sigma(m)\lhd\Omega(m)}(\sigma_*v) \\ &= ((\lhd(\Omega(m)))^*\phi^*\omega)_{\sigma(m)}(\sigma_*v) \\ &= Ad_{\Omega(m)^{-1}*}\left((\phi^*\omega)_{\sigma(m)}(\sigma_*v)\right) \\ &= Ad_{\Omega(m)^{-1}*}((\sigma^*\phi^*\omega)_m(v)) \\ &= Ad_{\Omega(m)^{-1}*}(((\phi\circ\sigma)^*\omega)_m(v)) \end{aligned} \tag{6.68}$$

(ii) Second term:[33]

$$\begin{aligned}(\phi^*\omega)_{\sigma(m)\triangleleft\Omega(m)} P_{\sigma(m)*}\Omega_* v &= \omega_{\phi\circ(\sigma(m)\triangleleft\Omega(m))}\left(\phi_* P_{\sigma(m)*}\Omega_* v\right) \\ &= \omega_{\sigma(m)}\left((\phi\circ P_{\sigma(m)\triangleleft\Omega(m)})_*\Omega_* v\right) \\ &= \omega_{\sigma(m)}\left(P_{(\phi\circ(\sigma(m)\triangleleft\Omega(m)))*}\Omega_* v\right) \\ &= \omega_{\sigma(m)}\left(P_{\sigma(m)*}\Omega_* v\right) \\ &= \omega_{\sigma(m)}\left(X^{\Omega_* v}_{\sigma(m)}\right) \\ &= \Omega_* v = (\Omega^*\Xi)_m v \end{aligned} \tag{6.69}$$

So, we can write

$$(\sigma^*\omega)_m(v) = Ad_{\Omega(m)^{-1}*}\left(((\phi\circ\sigma)^*\omega)_m(v)\right) + (\Omega^*\Xi)_m v, \tag{6.70}$$

which may be more comparable to the passive transformation when written in the form

$$A_\mu(m) = Ad_{\Omega(m)^{-1}*}\left(A'_\mu(m)\right) + (\Omega^*\Xi)_m(\partial_\mu), \tag{6.71}$$

and is G is a matrix group we can write

$$A_\mu(m) = \Omega(m)^{-1}A'_\mu(m)\Omega(m) + \Omega(m)^{-1}\partial_\mu\Omega(m). \tag{6.72}$$

Example 6.1. *[3]*

An example of a passive transformation can be given in the form of a magnetic monopole. The equation $\boldsymbol{B} = \nabla\times\boldsymbol{A}$ *requires that* $\nabla\boldsymbol{B}$ *be* 0 *everywhere.*

[33]Note that $\phi\circ P_{\sigma(m)} = P_{\phi\circ\sigma(m)}$. To see that let $h\in G$ and recall that $\phi(\sigma(m)\triangleleft h) = (\phi(\sigma(m)))\triangleleft h = \sigma(m)$.

$$\begin{aligned}(\phi\circ P_{\sigma(m)})(h) &= \phi(\sigma(m)\triangleleft h) = \sigma(m) \\ P_{\phi\circ\sigma(m)}(h) &= (\phi\circ\sigma(m))\triangleleft h = \sigma(m)\end{aligned}$$

Also notice that this does not hold in general, only when ϕ does not move us to a different fiber does this apply.

If a magnetic monopole ($\boldsymbol{r} = 0$) exists, it requires that

$$\boldsymbol{B} = g\frac{\boldsymbol{r}}{r^3}, \tag{6.73}$$

where g is the strength of a monopole. Solving that equation results in

$$\nabla \boldsymbol{B} = 4\pi g \delta^3(\boldsymbol{r}). \tag{6.74}$$

As $\nabla \boldsymbol{B} = 0 \neq 4\pi g \delta^3(\boldsymbol{r})$, the point $\boldsymbol{r} = 0$ is then removed from the further analysis so that the monopole is now defined in $\mathbb{R}^3 \backslash \{0\}$. We can now look for a vector potential that yields equation (6.73). Let (r, θ, Φ) be our coordinate system and suppose we looked at

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{g(1-\cos\theta)}{r\sin\theta}\hat{\boldsymbol{e}}_\Phi. \tag{6.75}$$

The solution to equation $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ is well-defined, and yields a magnetic field $\boldsymbol{B}$ everywhere ***except*** *at $\Theta = \pi$, which corresponds to a negative z-axis, where it results in a singularity. Similarly, vector potential*

$$\boldsymbol{A'}(\boldsymbol{r}) = -\frac{g(1+\cos\theta)}{r\sin\theta}\hat{\boldsymbol{e}}_\Phi\ , \tag{6.76}$$

yields a magnetic field $\boldsymbol{B}$ everywhere ***except*** *at $\Theta = 0$ which, this time, corresponds to a positive z-axis. As globally defined $\boldsymbol{A}$ and $\boldsymbol{A'}$ both result in singularities, Wu and Yang (1975) try to look at them locally, at the regions where they are well-defined.*

We can divide the surface of a sphere surrounding a monopole into two regions that we will call U^N (a region where $\boldsymbol{A}$ is well-defined) for the northern region, and U^S (a region where $\boldsymbol{A'}$ is well-defined) for the one in the south as shown in figure (6.7). Both U^N and U^S should be extended a little bit past the equator, so that they overlap. As our vector potentials now describe northern and southern regions, we can rename them accordingly, i.e.

$\boldsymbol{A} \coloneqq \boldsymbol{A}^N$ and $\boldsymbol{A'} \coloneqq \boldsymbol{A}^S$.

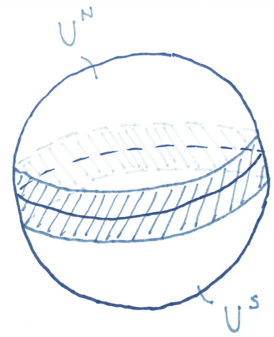


Figure 6.7: Intersection of two open sets of a sphere.

Now, both $\boldsymbol{A}^N$ and $\boldsymbol{A}^S$ are well-defined, and both do yield the correct solution for the magnetic field of a monopole (6.73). On the intersection $U^N \cap U^S$, they are related with a gauge transformation.

$$\boldsymbol{A}^N - \boldsymbol{A}^S = \frac{2g}{r\sin\theta}\hat{\boldsymbol{e}}_\Phi = \nabla(2g\Phi). \tag{6.77}$$

Written in terms of a coordinate basis, vector potentials (6.75) and (6.76) take the form of

$$A^N = ig(1-\cos\Theta)d\phi \qquad \text{and} \qquad A^S = -ig(1+\cos\Theta)d\phi. \tag{6.78}$$

Since we are looking at a sphere surrounding a magnetic monopole, the base space is S^2, and the fibers of a principal bundle are considered $U(1)$, which in turn means our transition function $\Omega \in U(1)$ is of the form $\Omega = e^{i\alpha}$. Following the equation of a passive gauge transformation (6.64), we can write

$$\begin{aligned} A^N_\mu &= e^{-i\alpha}A^S_\mu e^{i\alpha} + e^{-i\alpha}\partial_\mu e^{i\alpha} \\ A^N_\mu &= A^S_\mu + i\partial_\mu\alpha, \end{aligned} \tag{6.79}$$

where $d\alpha = -i(A^N - A^S) = 2gd\phi \implies \alpha = 2\pi\phi$.

Example 6.2. *An example of an active gauge transformation will be the one of a plane wave. Here, the base space is a space-time* $\mathbb{R}^4$*, and a Lie group in question is again* $U(1)$*, which means we can still write* $\Omega = e^{i\alpha}$*.*

A four-vector potential is given by $\mathcal{A}_\mu = i\epsilon_\mu \cos(kx)$*, where* ϵ_μ *are polarization vectors. In this example, we follow the formula for the case of active gauge transformation (6.72). Here, multiplication by* $e^{i\alpha} \in U(1)$ *(such multiplication can be identified as a right action) corresponds to the change of a section* σ*. As a right action can not move the points from one fiber to another, multiplication by* $e^{i\alpha}$ *from this example is considered an automorphism, making this an example of an active gauge transformation. By following expression (6.72), we can write the following*

$$\begin{aligned}
\mathcal{A}_\mu &= i\epsilon_\mu \cos(kx) \\
\tilde{\mathcal{A}}_\mu &= e^{-i\alpha}\mathcal{A}_\mu e^{i\alpha} + e^{-i\alpha}\partial_\mu e^{i\alpha} \\
\tilde{\mathcal{A}}_\mu &= \mathcal{A}_\mu + i\partial_\mu \alpha.
\end{aligned} \tag{6.80}$$

Note that the gauge potential used here differs by a factor i *that came from the Lie algebra. If we wish to write equation (6.80) in its usual form, we can write* $\mathcal{A} \coloneqq iA$*.*

$$\begin{aligned}
\tilde{\mathcal{A}}_\mu &= A_\mu + i\partial_\mu \alpha \\
i\tilde{A}_\mu &= iA_\mu + i\partial_\mu \alpha \\
\tilde{A}_\mu &= A_\mu + \partial_\mu \alpha
\end{aligned} \tag{6.81}$$

.

For $\alpha = \cos(kx)$ *we can write*

$$\begin{aligned}
\tilde{\mathcal{A}}_\mu &= i\epsilon_\mu \cos(kx) - ik_\mu \sin(kx) \\
\implies A_\mu &= \epsilon_\mu \cos(kx) - k_\mu \sin(kx).
\end{aligned} \tag{6.82}$$

§7 Parallel transport

We will first introduce the idea of a **horizontal lift** of a curve inhabiting a base manifold. As we would like to know the derivative of some curve in a base space, the basic idea is to lift both the curve and its tangent vectors to a total space of a principal bundle where we learned how to *connect* the points of neighboring fibers.

7.1 Horizontal lift

Suppose we glue circles $S^1 \cong SO(2)$ to each point of a torus. What we have now is a principal $SO(2)$-bundle, similar to example (6.1), but this time, the torus is a base manifold. As visualization of such a total space would be complicated at best, let us note that the visualization of $SO(2)$ as a circle was a *choice* from the get-go. One can easily visualize such a Lie group as a line by identifying the same points of a unit circle on a line.

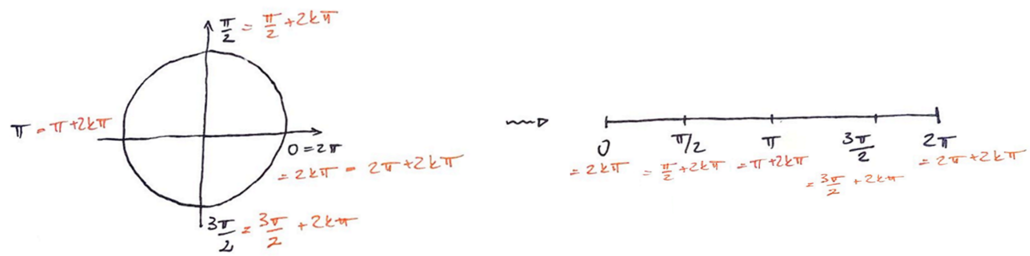


Figure 7.1: Different visualization of a Lie group $SO(2)$.

The reason we mention this now is that, with that in mind, we can freely draw at least *this* principal bundle that has a $2D$ manifolds as a base space, so the visualization of curves in base manifolds is not just parts of circular arcs. Note that a vertical subspace can now be identified with the fiber itself (as a straight line would be tangent to itself), while a horizontal subspace is then visualized as any line that intersects the fiber but is not the fiber itself.

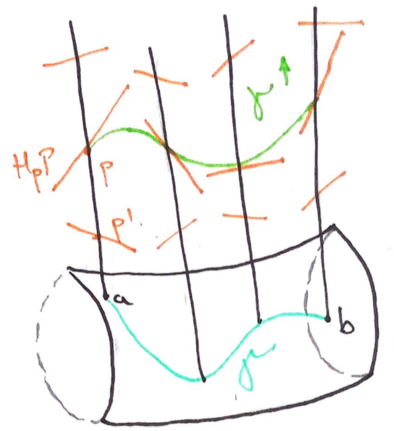


Figure 7.2: Horizontal lift $\gamma^{\uparrow}$ of a curve γ through a point $p \in P$.

The following explanation can be found in [36]. Suppose we have a curve γ on our base manifold of a principal $SO(2)$-bundle as shown in the figure (7.2). As noted in the introduction of this section, we now want to lift that curve to the total space. Let us denote the lifted curve $\gamma^{\uparrow}$. We will call such a curve the **horizontal lift of γ**.

Such a lifted curve $\gamma^{\uparrow}$ can be thought of as a local part of a cross-section. However, we can not *just* use a cross-section to lift γ, as we need to keep in mind that vectors tangent to γ should map to vectors tangent to $\gamma^{\uparrow}$. As every point of a curve $\gamma^{\uparrow}$ should belong to a different fiber, every vector tangent to $\gamma^{\uparrow}$ should "*point*" to the next fiber. This logic might seem familiar as that is exactly how the connection operates. Vectors tangent to lifted curves are vectors from horizontal subspaces, and each of these horizontal subspaces then uniquely determines the following point in a curve $\gamma^{\uparrow}$. This means that it is enough to choose a starting point $p \in P$ and we can now rely on horizontal subspaces of each subsequent neighboring fiber to do the rest. However, by doing that, it is **not guaranteed** that a curve $\gamma^{\uparrow}$ will project back into our starting γ, so that needs to be set as a condition, *i.e.* we want $\pi \circ \gamma^{\uparrow} = \gamma$ to hold true.

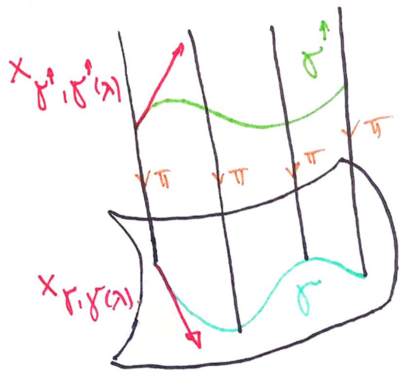


Figure 7.3: Vectors tangent to γ and $\gamma^{\uparrow}$.

As we will be discussing vectors that are tangent to specific curves at specific points, the curve and a point will be denoted in the index of a tangent vector, *i.e.* $X_{\gamma,\gamma(\lambda)}$ is a vector tangent to the curve γ at a point $\gamma(\lambda)$. Tangent vectors should map to tangent vectors, so we require that projection of a vector tangent to $\gamma^{\uparrow}$ results in vector tangent to γ, *i.e.* we require that $\pi_* X_{\gamma^{\uparrow},\gamma^{\uparrow}(\lambda)} = X_{\gamma,\gamma(\lambda)}$. As vectors tangent to $\gamma^{\uparrow}$ live in horizontal subspaces, *i.e.* $X_{\gamma^{\uparrow},\gamma^{\uparrow}(\lambda)} \in H_{\gamma^{\uparrow}(\lambda)}P$ we can write $\omega_{\gamma^{\uparrow}(\lambda)}(X_{\gamma^{\uparrow},\gamma^{\uparrow}(\lambda)}) = 0$ as that is how we defined the horizontal subspace.[34]

With all of that cleared up, we can formally define the *horizontal lift*.

[34]Or one can equally validly write $ver(X_{\gamma^{\uparrow},\gamma^{\uparrow}(\lambda)}) = 0$ as vectors from horizontal subspace have no vertical component.

Definition 7.1. *Let* $P \xrightarrow{\pi} M$ *be a principal* G*-bundle equipped with a connection 1-form* ω*. Let* $\gamma : [0,1] \to M$ *be a curve such that* $\gamma(0) = a$ *and* $\gamma(1) = b$*, where* $a, b \in M$*. Then, a unique curve* $\gamma^{\uparrow} : [0,1] \to P$ *going through a point* $\gamma^{\uparrow}(0) =: p$*, where* $p = \pi^{-1}\{(a)\}$*, which satisfies the following*

(i) $\pi \circ \gamma^{\uparrow} = \gamma$*,*

(ii) $ver(X_{\gamma^{\uparrow}, \gamma^{\uparrow}(\lambda)}) = 0$ *and*

(iii) $\pi_* X_{\gamma^{\uparrow}, \gamma^{\uparrow}(\lambda)} = X_{\gamma, \gamma(\lambda)}$

is called a ***horizontal lift of*** $\boldsymbol{\gamma}$ ***through a point*** p*.*
[36, 5, 3, 12]

Note that the uniqueness of $\gamma^{\uparrow}$ stems solely from a condition that we chose a starting point $p \in P$. In fact, for each $p \in \pi^{-1}\{(a)\}$ there exists a different horizontal lift [5].

Even though we set the conditions that need to be satisfied for our horizontal lift, we still do not have an explicit expression of how $\gamma^{\uparrow}$ would look like. The cross-section σ is used to lift γ to a total space. As that would generate some arbitrary curve $\delta = \sigma \circ \gamma$, such a curve may or may not be a horizontal lift. It does, however, project back into γ. The idea is then to find all such $g \in G$ that act on every point of δ resulting in the $\gamma^{\uparrow}$, *i.e.* $\gamma^{\uparrow} = \delta \triangleleft g$. That is done by defining a curve for the Lie group $g : [0,1] \to G$. Then the wanted equation can be written as $\gamma^{\uparrow}(\lambda) = \delta(\lambda) \triangleleft g(\lambda)$.

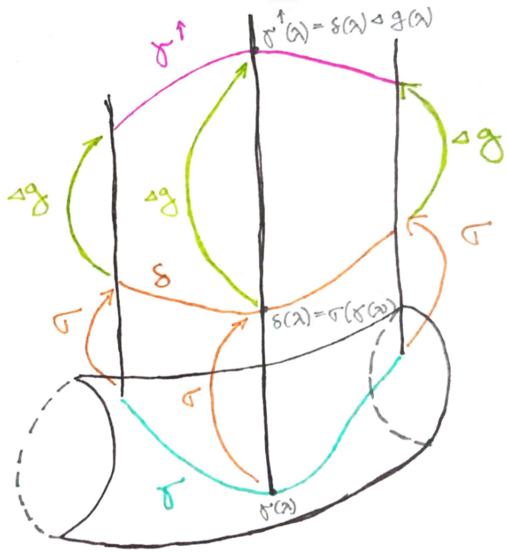


Figure 7.4: $\gamma^{\uparrow}(\lambda) = \delta(\lambda) \triangleleft g(\lambda)$.

What follows is a rough explanation of how one can do such a thing in practice.

Let $\delta : [0,1] \to P$ a curve such that $\pi(\delta(\lambda)) = \gamma(\lambda)$, $\forall \lambda \in [0,1]$. As we can always find such a $p \in \pi^{-1}\{(a)\}$ for which the horizontal lift of γ exists, this means that there must exist such a $g(\lambda) : [0,1] \to G$ so that we get to this $\gamma^{\uparrow}$, *i.e.* $\gamma^{\uparrow}(\lambda) = \delta(\lambda) \triangleleft g(\lambda)$, $\forall \lambda \in [0,1]$. We can factor $\gamma^{\uparrow}$ in a similar manner as was done for theorems (6.2) and (6.4). As the procedure for getting explicit expression is very similar to that of proving the referenced theorems, some repetitive parts will be left out. The following can be found in [5, 36].

$$\begin{aligned} \gamma^{\uparrow} : [0,1] &\xrightarrow{\delta \times g} P \times G \xrightarrow{\triangleleft} P \\ \lambda \;\mapsto\; (\delta(\lambda), g(\lambda))) &\;\mapsto\; \delta(\lambda) \triangleleft g(\lambda) \end{aligned} \tag{7.1}$$

Let X be vector in $[0,1]$ defined at a point λ.

$$\begin{aligned}\gamma^{\uparrow}_{*}X &= (\triangleleft \circ (\delta \times g))_{*}X = \triangleleft_{*}((\delta \times g)_{*}X)\\ &= \triangleleft_{*}(\delta_{*}X, g_{*}X) = \triangleleft_{*}(i_{g(\lambda)*}\delta_{*}X + j_{\delta(\lambda)*}g_{*}X) \Big/ \omega\\ \omega_{\delta(\lambda)\triangleleft g(\lambda)}(\gamma^{\uparrow}_{*}X) &= \omega_{\delta(\lambda)\triangleleft g(\lambda)}\left(\triangleleft_{*}(i_{g(\lambda)*}\delta_{*}X + j_{\delta(\lambda)*}g_{*}X)\right) \qquad (7.2)\end{aligned}$$

As $\gamma^{\uparrow}_{*}X$ is by definition part of a horizontal subspace $\omega_{\delta(\lambda)\triangleleft g(\lambda)}(\gamma^{\uparrow}_{*}X) = 0$.

$$\begin{aligned}0 &= \underbrace{\omega_{\delta(\lambda)\triangleleft g(\lambda)}\left(\triangleleft_{*}(i_{g(\lambda)*}\delta_{*}X\right)}_{=\omega_{\delta(\lambda)\triangleleft g(\lambda)}((\triangleleft g(\lambda))_{*}\delta_{*}X)} + \underbrace{\omega_{\delta(\lambda)\triangleleft g(\lambda)}\left(\triangleleft_{*}j_{\delta(\lambda)*}g_{*}X)\right)}_{=\omega_{\delta(\lambda)\triangleleft g(\lambda)}(X^{g_{*}X}_{\delta(\lambda)\triangleleft g(\lambda)})}\\ &= \left((\triangleleft g(\lambda))^{*}\omega\right)_{\delta(\lambda)}\delta_{*}X + g_{*}X\\ &= Ad_{g(\lambda)^{-1}*}\left(\omega_{\delta(\lambda)}(\delta_{*}X)\right) + \Xi_{g(\lambda)}(g_{*}X) \qquad (7.3)\end{aligned}$$

As $\delta_{*}X$ represents vector tangent to δ and is defined at a point $\delta(\lambda)$, we will write $\delta_{*}X \coloneqq X_{\delta,\delta(\lambda)}$, and similarly $g_{*}X \coloneqq X_{g,g(\lambda)}$.

Altogether we have

$$Ad_{g(\lambda)^{-1}*}\left(\omega_{\delta(\lambda)}(X_{\delta,\delta(\lambda)})\right) + \Xi_{g(\lambda)}(X_{g,g(\lambda)}) = 0. \qquad (7.4)$$

Although it may not look like it in this form, this equation is first-order ordinary differential equation. For G being a matrix group, it can be re-written in a form that may represent that fact more clearly.

$$\begin{aligned}&g(\lambda)^{-1}\cdot\omega_{\delta(\lambda)}(X_{\delta,\delta(\lambda)})\cdot g(\lambda) + g(\lambda)^{-1}\cdot\frac{dg(\lambda)}{dt} = 0\\ \Longleftrightarrow\ &\frac{dg(\lambda)}{dt} = -\omega_{\delta(\lambda)}(X_{\delta,\delta(\lambda)})\cdot g(\lambda). \qquad (7.5)\end{aligned}$$

Solving this equation with the initial condition $g(0) = g_0$ gives rise to an explicit form of a horizontal lift curve $\gamma^{\uparrow}$.

$$\gamma^{\uparrow}(\lambda) = \delta(\lambda) \triangleleft \underbrace{\left(P \cdot \left(e^{-\int_0^\lambda dt \Gamma_\mu(\gamma(t))\dot{\gamma}^\mu(t)}\right)\right)}_{\in G} \cdot g_0, \tag{7.6}$$

where $P \cdot \exp(-)$ is a path-ordered exponential, $\dot{\gamma}^\mu(t) = X_{\gamma,\gamma(t)}$ and $\Gamma_\mu \coloneqq \omega^U_\mu$.

7.2 Parallel transport

Now that we have the horizontal lift of γ, we can define a **parallel transport map**. Such a map simply uses $\gamma^{\uparrow}$ to define the interval between finitely separated fibers. It can be, ***very roughly speaking***, *thought of* as Δx one might use to define derivatives in a classical sense.

Definition 7.2. *Let $\gamma : [0,1] \to M$ be a curve in a base manifold and $\gamma^{\uparrow}_p : [0,1] \to P$ be one horizontal lift of that curve through point $p \in \pi^{-1}(\{\gamma(1)\})$. Then the* ***parallel transport map*** *is*

$$\begin{aligned} T_\gamma &: \underbrace{\pi^{-1}(\{\gamma(0)\})}_{\text{"initial fiber"}} \to \underbrace{\pi^{-1}(\{\gamma(1)\})}_{\text{"final fiber"}} \\ T_\gamma &: \pi^{-1}(a) \to \pi^{-1}(b) \\ & \quad p \mapsto \gamma^{\uparrow}_p(1). \end{aligned} \tag{7.7}$$

[36, 5]

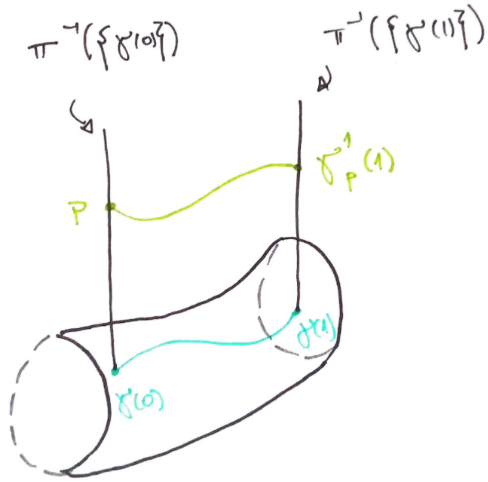


Figure 7.5: Parallel transport.

7.3 Horizontal lift and parallel transport defined on an associated bundle

To accurately convey the following plan, let us first discuss how we compare vectors in the Euclidean space. Or, more specifically, how one adds vectors that are defined at different points in space, which is the exact scenario we have here with our parallel transport map.

Now, in the Euclidean plane, if one wishes to add two vectors that are defined at different points one needs to first shift the beginning of one vector to the end of the other one. That is physically done using a set square and parallel transporting a vector to the point where we want it to be. Axiomatically, the parallelly transported vector and the starting vector are considered the same.

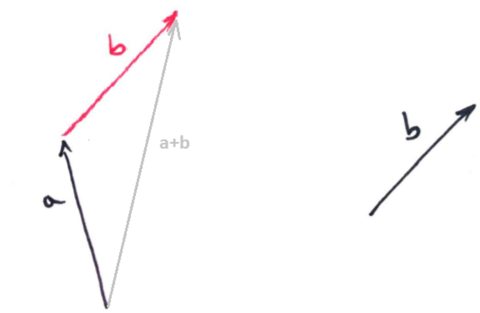


Figure 7.6: Parallel transport in a flat space.

On an arbitrary manifold, such parallel transport is done along some curve as that is how we know to "*move*" the vector. Now, we are moving the vector through fibers with the use of a parallel transport map until we end up with two vectors in the same fiber. However, these two would then, still be defined at different points, even if they live in the same fiber - so, to be able to compare these two vectors, we will do all of what has just been described in a vector bundle.[35]Because all the vectors that live in one fiber of an associated vector bundle are defined over the same point $m \in M$, we know how to compare them. Not only compare, but we know how to add them because the fiber is now a vector space where we know there exists a well-defined pointwise addition.

[35]Vector bundle is just an associated bundle whose fibers are vector spaces. An example of such a bundle is a tangent bundle. The name "vector bundle" is used just so the generality is kept.

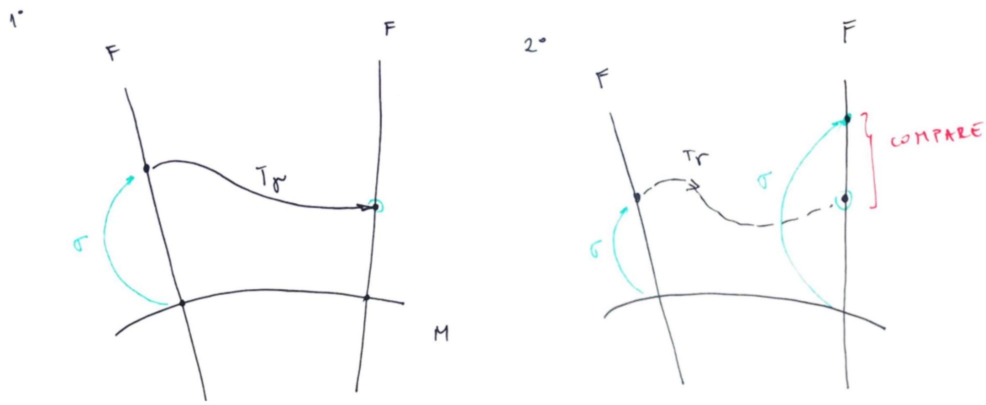


Figure 7.7: Comparing the points of finitely separated fibers by parallelly transporting one point to the same fiber and then comparing them there.

This is the basic idea, and to get there we need to define both the horizontal lift of a curve in a base manifold and a parallel transport map on an associated vector bundle. Since it is very much analogous to the case of defining those on a principal bundle, we will not delve much into the definitions that are given below.

Definition 7.3. *Let $P \overset{\pi}{\to} M$ be a principal G-bundle with a well-defined connection one-form ω. Let $P_F \overset{\pi_F}{\to} M$ be associated vector bundle. Also let $\gamma : [0,1] \to M$ be a curve in a base manifold and $\gamma_p^\uparrow : [0,1] \to P$ its horizontal lift through $p \in \pi^{-1}(\{\gamma(0)\}$ Then the* ***horizontal lift of γ to the associated bundle*** *that passes through a point $[p,f] \in P_F$ is a curve*

$$\begin{aligned} \overset{P_F}{\gamma^\uparrow_{[p,f]}} &: [0,1] \to P_F \\ \overset{P_F}{\gamma^\uparrow_{[p,f]}} &:= \left[\gamma^\uparrow(\lambda), f\right]. \end{aligned} \tag{7.8}$$

[36, 5]

Definition 7.4. *A **parallel transport map**, as defined on the associated vector bundle, is a map*

$$\begin{aligned} T_\gamma^{P_F} :& \pi_F^{-1}(\{\gamma(0)\}) \to \pi_F^{-1}(\{\gamma(1)\}) \\ & [p,f] \mapsto \gamma^{\overset{P_F}{\uparrow}}_{[p,f]} \end{aligned} \tag{7.9}$$

[36, 5]

7.4 Covariant derivative

It is time to finally formally define a covariant derivative. Figure (7.8) should help illustrate the basic idea behind the definition of a covariant derivative given below the figure. Again note that this is done on a vector bundle as there, we are able to compare the vectors.

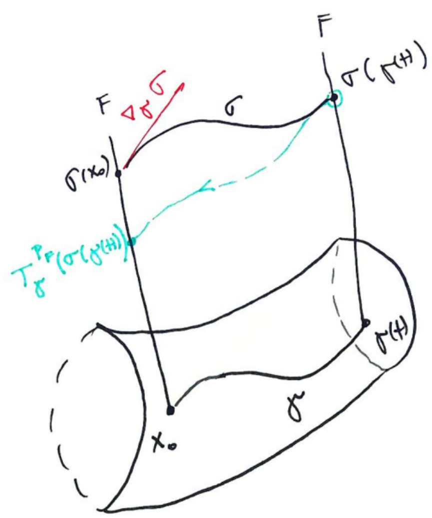


Figure 7.8: Covariant derivative of a cross-section σ in the direction of a curve γ at a point $\gamma(0) := x_0$.

Definition 7.5. *Let $P_F \overset{\pi_F}{\to} P$ be associated vector bundle to a principal bundle $P \overset{\pi}{\to} M$ with a connection 1-form ω. Let γ be a curve in a base manifold such that $\gamma(0) := x_0$ and let $\sigma : M \to P_F$ be a cross-section of an associated vector bundle. Then a* ***covariant derivative of a cross-section $\boldsymbol{\sigma}$ in the direction of a curve $\boldsymbol{\gamma}$ at a point $\mathbf{x_0}$*** *is defined as*

$$\nabla_\gamma \sigma = \lim_{t \to 0} \frac{T_\gamma^{P_F}(\sigma(\gamma(t))) - \sigma(x_0)}{t} \tag{7.10}$$

[5]

7.5 Curvature on a principal bundle

Even though we have solved the issue of not being able to differentiate on arbitrary manifolds, there exist more concepts related to principal bundles and connections that are useful to physicists. Namely, the idea we are interested in in this section is the one of **curvature**, or more precisely, we are interested in the curvature of a connection on a principal bundle. To understand what that means, we should discuss both the idea of curvature in a more general sense and the curvature in the context of principal bundles. A wonderful introduction is provided by none other than Wikipedia [37].

> "In mathematics, curvature is any of several strongly related concepts in geometry that intuitively measure the amount by which a curve deviates from being a straight line or by which a surface deviates from being a plane."

The example that should serve as a way to get more of an intuition for curvature is one of a circle. As the curvature should measure the deviation from the straight line, one can imagine a tangent line of two circles - a small circle and a big circle. Locally, around its tangent line, a smaller circle is more curved than a bigger one as a bigger circle locally resembles a straight line more compared to the smaller one.

In the context of differential geometry, the term *curvature* (sometimes referred to as a *curvature form*) then describes a curvature of a connection on a principal bundle. For a formal definition, we also need the concept of a **covariant exterior derivative**. Note that the sign ❁ will appear in the definition as a placeholder for any arbitrary space. Similar to the concept of a "*Lie algebra*-valued" n-forms, there exists a notion of a "*random-space-valued*" n-form. If one imagines a Lie-algebra valued n-form to be an n-form in a "normal" sense that alongside itself carries some vector from the Lie algebra, one can also imagine a "random-space-valued" n-form to carry the objects from that random space. So, a ❁-valued n-form would then carry some object from a ❁ space.

Definition 7.6. *Let* $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ *be a principal bundle with a connection 1-form* ω*. Let* Φ *be ❁-valued k-form.*
Then a map

$$D\Phi : \underbrace{\Gamma(T_0^{k+1}P)}_{\substack{\text{The space}\\ \text{of } k+1\\ \text{vector fields.}}} \to \text{❁} \tag{7.11}$$

$$D\Phi \underbrace{(X_1, \dots, X_{n+1})}_{\substack{\text{These are now}\\ \text{vector fields.}}} := d\Phi(hor(X_1), \dots, hor(X_{k+1}))$$

is called ***exterior covariant derivative*** *of a k-form* Φ*.*
[38, 5]

Curvature is then defined to be a special case of an exterior covariant derivative for $k = 1$ and ❁ $= T_eG$, so we can write

Definition 7.7. *Let* $P \xrightarrow{\triangleleft G} P \xrightarrow{\pi} M$ *be a principal bundle with a connection 1-form* ω*. Then the* ***curvature*** *of the connection 1-form is the Lie-algebra-valued 2-form on* P*.*

$$\begin{aligned} \Omega &: \Gamma(T^2_0 P) \to T_e G \\ \Omega &:= D\omega \end{aligned} \tag{7.12}$$

[38, 5]

If one thinks of exterior derivatives as a generalization of a total derivative, one can imagine n-forms that are obtained as a result of exterior differentiation being a sort of a measure of how much does that $(n-1)$-form deviate from our starting arbitrary manifold. Similarly, one can then imagine an exterior covariant derivative acting in that same manner only this time it is for the case of connection 1-forms, and the "*straight line*" is a horizontal subspace. This would then mean that what is considered a *curvature* in the definition above is a plane of deviation from a curved manifold that is our base space.

Related to the definition of curvature is a theorem that physicists might recognize as a Riemann curvature tensor in the context of general relativity or as a Yang-Mills field strength tensor in the context of particle physics. As the theorem is of great importance it has a name - **Cartan structural equation**.

Theorem 7.1. *Let* $\Omega \equiv D\omega$ *be a curvature on a principal bundle. Then the following holds true*

$$\Omega = d\omega + \omega \curlywedge \omega. \tag{7.13}$$

[38, 3, 5]

Note that the sign $\curlywedge$ used in the definition is used as opposed to the standard wedge sign, since ω is now a Lie-algebra-valued n-form and the "normal" wedge product rules do not apply. This notation is most likely introduced by Schuller, as in literature the standard wedge sign $\wedge$ is used for this case with an implicit understanding of its meaning. In a more general sense, ω can be any ❀-valued n-form and this theorem would still apply. Depending on the ❀ space then, the sign $\Wedge$ holds different meanings. As we have defined curvature for the case where ❀ is a Lie algebra, we can safely define $\Wedge$ as

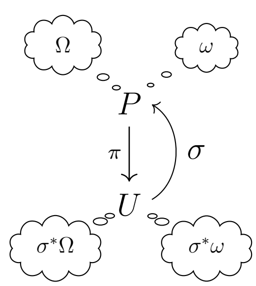
$$(\omega \Wedge \omega)(X, Y) := [\![\omega(X), \omega(Y)]\!]\,, \tag{7.14}$$

where X and Y are just vector fields on which ω acts on and bracket $[\![\,\_\,,\,\_\,]\!]$ is a Lie algebra bracket making the resulting Ω again a Lie-algebra-valued 2-form, *i.e.* a curvature.

If the principal bundle group G is a matrix group, then the equation from the theorem can be re-written in an index form as

$$\Omega^i_j = d\omega^i_j + \omega^i_k \wedge \omega^k_j. \tag{7.15}$$

As both the connection 1-form and curvature are ultimately defined in the total space of a principal bundle, and physics is usually done on a base manifold, we should look at how these objects map to their base manifold equivalents.



(7.16)

The pullback of a connection 1-form is referred to as a *gauge potential*, or sometimes, as a *Yang-Mills field*. The pullback of a curvature to the base manifold is called a **Yang-Mills field strength** or just a **field strength** [5, 38]. Note that Y.Choquet-Bruhat, C. Dewitt-Morette and M.Dillard-Bleick [6, p. 374] call it a **local curvature on the base manifold**. As is often the case in the field of mathematics, singular concepts are able to describe multiple seemingly unrelated problems and the Cartan structural equation is no exception to this phenomenon. That is why pullbacks of curvature and the connection 1-form can go by more names than were mentioned here, depending on the context in which they are referred to. In the context of electrodynamics, these objects that we pulled back to the base manifold can be identified with the Faraday tensor F (EM field strength) that would represent a pull-back of a curvature to the base manifold, and a *gauge potential* A (EM field) which is a pull-back of a connection 1-form to the base manifold. A similar notation is used in particle physics so we can call $\sigma^*\omega := A$ a Yang-Mills field and $\sigma^*\Omega := F$ the Yang-Mills field strength. In the context of general relativity, $\sigma^*\Omega := R$ refers to the Riemann curvature tensor and $\sigma^*\omega := \Gamma$ are referred to as Christoffel symbols.

In general, on the total space of a principal bundle, we can write

$$\begin{aligned} \Omega &= d\omega + \omega \curlywedge \omega \\ \implies \sigma^*\Omega &= \sigma^*(d\omega + \omega \curlywedge \omega) \\ &= \sigma^*(d\omega) + \sigma^*(\omega) \curlywedge \sigma^*(\omega). \end{aligned} \tag{7.17}$$

These examples can be written side by side for easier comparison [3].

In the case of general relativity:	In the case of particle physics:
$\sigma^*\omega \coloneqq \Gamma$ and $\sigma^*\Omega \coloneqq R$	$\sigma^*\omega \coloneqq A$ and $\sigma^*\Omega \coloneqq F$
$\Gamma^\mu_\nu = dx^\lambda \Gamma^\mu_{\lambda\nu}$	$A = A_\mu dx^\mu$
$R^\mu_\nu = d\Gamma^\mu_\nu + \Gamma^\mu_\sigma \curlywedge \Gamma^\sigma_\nu$	$F = dA + A \curlywedge A$
$R^\mu_\nu = \frac{1}{2} R^\mu_{\nu\rho\sigma} dx^\rho \wedge dx^\sigma$	$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$
$R^\kappa_{\lambda\mu\nu} = \partial_\mu \Gamma^\kappa_{\nu\lambda} - \partial_\nu \Gamma^\kappa_{\mu\lambda} + \Gamma^\eta_{\nu\lambda}\Gamma^\kappa_{\mu\eta} - \Gamma^\eta_{\mu\lambda}\Gamma^\kappa_{\nu\eta}$	$F^\alpha_{\mu\nu} = \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + f^\alpha_{\beta\gamma} A^\beta_\mu A^\gamma_\nu$

(7.18)

R and F **are** precisely the same formula and they both represent pullbacks of curvature. Discussion on indices is in order. In the following expressions, proper 1-form indices are written in blue. Notice that those are exactly the indices that appear with 1-form basis dx^μ.

$$\Gamma^\mu_\nu = dx^\lambda \Gamma^\mu_{\lambda\nu} \qquad \text{and} \qquad A = A_\mu dx^\mu. \tag{7.19}$$

The rest of the indices are exactly the ones that come from Lie algebra and we can write them in orange to highlight their existence. For the full expressions, this means that the following orange indices come from Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ in case of the Riemann tensor, and $SU(N)$ in case of Yang-Mills field strength.

$$\begin{aligned} R^\kappa_{\lambda\mu\nu} &= \partial_\mu \Gamma^\kappa_{\nu\lambda} - \partial_\nu \Gamma^\kappa_{\mu\lambda} + \Gamma^\eta_{\nu\lambda}\Gamma^\kappa_{\mu\eta} - \Gamma^\eta_{\mu\lambda}\Gamma^\kappa_{\nu\eta} \\ F^\alpha_{\mu\nu} &= \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + f^\alpha_{\beta\gamma} A^\beta_\mu A^\gamma_\nu. \end{aligned} \tag{7.20}$$

§8 Conclusion

Bundles were used as a tool for formally describing *vector* and *covector fields*. This was done in the form of a *tangent bundle* (or a *cotangent bundle*). Arguably, bundles teach us an even more important lesson, which is the idea of *gluing* objects we are interested in onto a *base manifold*. This allows not only for the notions of vector or covector bundles, but also a frame bundle. As we soon learned, all of these are just specific examples of two more important objects that are closely linked to each other - a *principal fiber bundle*, and all its *associated bundles*. For the full description of principal fiber bundles, we first needed to introduce the topic of Lie groups and Lie algebras. As Lie groups describe continuous symmetries, any description of nature that is invariant under gauge transformations should include Lie groups. Two Abelian Lie groups were chosen to study how Lie groups behave and how they relate to their algebras. Alongside that, a short introduction to Lie group and Lie algebra representations was given, so that the future equations that relate to physics could be written in their more familiar forms. Principal fiber bundles were then introduced as fiber bundles which can, at least locally, be thought of as a product of a base manifold, and a Lie group. However, such bundles needed context in the form of associated fiber bundles. One of the most important uses of their relationship was the fact that all the changes on a principal bundle are reflected in all of its associated bundles. This meant that later in this work we could define a horizontal lift of a curve and parallel transport map on a principal bundle, have them both be well-defined, and because they were well-defined on a principal bundle, their redefinitions on the associated vector bundle followed. And, only by transferring them to the associated vector bundle, were we able to finally compare the points in one fiber. Only by picking and choosing the properties we wanted from each of these, were we able to actually get to a proper definition of a covariant derivative. The property from a principal bundle that we picked was one of a connection that helped us connect the neighboring fibers, and the property we picked from an associated vector bundle was one of comparing the points within one fiber.

Only after defining the *connection 1-form*, were we able to introduce a first use of principal bundles in physics, in the form of *gauge transformations*. As we saw in chapter (6.4), gauge transformations can be obtained in two different manners. The first one we introduced was a *passive gauge transformation*, that can mathematically be described as a compatibility requirement for two different gauge potentials defined at the same intersection. The example given for this case was one of a magnetic monopole. The second one was an *active gauge transformation*, which can be mathematically described as an automorphism on a principal bundle. The example given in this case was the one of plane wave.

The last application of connections to the field of physics was one of the *curvature*, or more precisely, the *Cartan structural equation* that describes curvate in terms of connection 1-forms. We found this equation in both general relativity, in the form of a Riemann curvature tensor, and in particle physics in the form of a Yang-Mills field strength.

§A Appendix

A.1 The operations of push-forward and pull-back generalized to vector and covector fields

A.1.1 Push-forward of a vector field

Push-forward can not be extended to a vector field unless the underlying map Φ is a diffeomorphism. As opposed to defining the push-forward of a tangent vector, where a map used to push-forward a tangent vector does not have to be surjective or injective, in the case of vector fields, we have a special condition that needs to stay preserved while mapping, that condition comes from the definition of a cross-section - $(\pi \circ \sigma)(p) = p$. Note that this holds for points, but it should also hold for vectors we map.

Suppose we wanted to push vectors from a total space to a base space using a map π. Since π is by definition surjective, the induced π_* is also surjective, so from all of the points π mapped to in the base space, *grow* vectors pushed by π_*. One such pushed vector is drawn in the figure below. Similarly, we can imagine pushing a vector from the base space to the total space with a map σ. Since σ attaches exactly one point from a fiber to any p in a base space, σ is an injective map. By extension then, an induced map σ_* is also injective. All the vectors that are pushed to the total space map to exactly one vector.

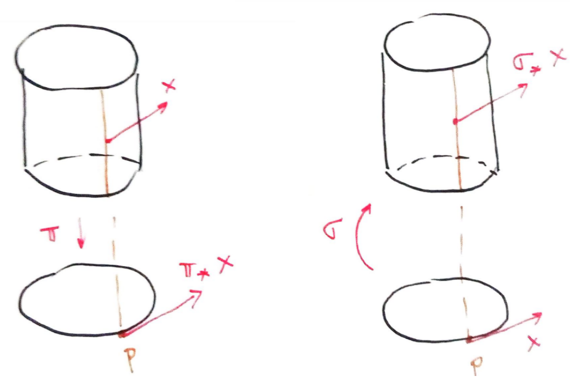


Figure A.1: Push-forward of vectors using both π and σ.

And since we know that $(\pi \circ \sigma)(p) = p$ induced maps π_* and σ_* should then map in the following way $(\pi_* \circ \sigma_*)(x) = x$, for any tangent vector $x \in T_pB$, where T_pB is now a tangent space of a base space. For that to be possible, an underlying smooth map $\pi \circ \sigma$ should be bijective, *i.e.* it should be *diffeomorphic*.

All that is to say that one has to define a push-forward of a vector field to be a diffeomorphic map. One specific and quite useful diffeomorphic map that we know of is the left transition l_g, so by definition push-forward of a vector field is set to be induced by the left translation. Note that since $l_g : G \to G$, such push-forward can only be used if we wish to push vectors from smooth manifold G to that same G. Also, note that using l_g to extend the concept of push-forward to the whole of a vector field is a *choice* and not a requirement. Such a choice limits one to a single manifold G.

One other note is that since we are discussing both tangent vectors and vector fields that sometimes had the same name x (or X), as a way to distinguish between the two, we shall label tangent vectors defined at the point p as x_p (or X_p) while vector fields shall have no such indices as they are not defined at one point.

Since our push-forward for tangent vectors was defined from any $T_{\Phi(p)}N$ to some other T_pM, and we intend to generalize this using left translation, we should re-define a push-forward of a tangent vector to be induced by left-translation only. The figure below shows a push-forward of a vector x_h to be a vector $l_{g*}x_h$. Since we know that $l_g(h) = gh$, pushed vector $l_{g*}x_h$ is defined at a point gh and thus holds another name x_{gh}. As those are the same, we define a push-forward of a tangent vector by left translation as $l_{g*}x_h \coloneqq x_{gh}$.

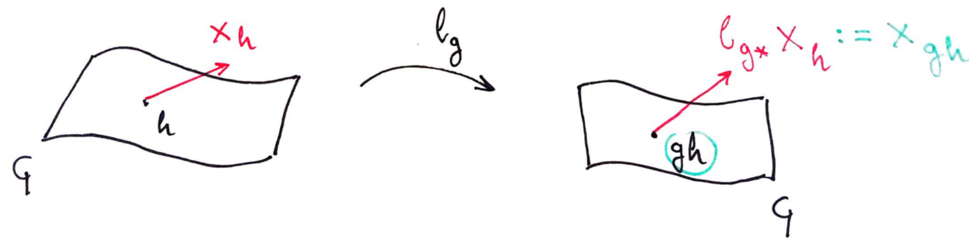


Figure A.2: Push-forward of vectors by left translation l_g.

Another thing to note is that one can think of a tangent vector in terms of a vector field. Namely, vector field X valued at some point h **is** a tangent vector X_h. A vector field $l_{g*}X$ valued at the point gh is a tangent vector $(l_{g*}X)_{gh}$. We can use that line of thinking to define a push-forward of a vector field at a point. From the picture above we know that a vector at a point gh is $l_{g*}x_h$ and we also know that push forward of a vector field at that same point yields a vector $(l_{g*}X)_{gh}$. As both of these are the same vector defined at the same point, they are equated by definition: $(l_{g*}X)_{gh} \coloneqq l_{g*}X_h$ [15].

Vector fields for which it holds that $l_{g*}X = X$ are called *left-invariant* as they do not change the vector field while mapping. There exist three equivalent statements that all yield left-invariant vector fields and they are all mentioned in the definition below.

Definition A.1. *Let* $(G, \bullet)$ *be a Lie group, and let* X *denote the vector field on* G*. Vector field* X *is called* ***left-invariant*** *for any* $g \in G$ *and* $\forall h \in G$ *if the following equivalent statements hold true:*

(i) $l_{g*}X = X$

(ii) $l_{g*}X_h = X_{gh}$

(iii) $X(f \circ l_g) = (Xf) \circ l_g$

[15, 5, 3, 12, 6]

The set of all left-invariant vector fields on G is denoted $L(G)$.

A.1.2 Pull-back of a n-form field

Let the $\Omega^n(N)$ denote the set of all n-form fields. We wish for the following generalization.

$$\begin{array}{ccc} \Phi : M & \longrightarrow & N \\ & \rightsquigarrow & \\ \Phi^* : T^*_{h(p)}N & \longrightarrow & T^*_pM \\ & \rightsquigarrow & \\ \Phi^* : \Omega^n(N) & \longrightarrow & \Omega^n(M) \end{array} \tag{A.1}$$

Definition A.2. *Let* $\omega \in \Omega^n(M)$*. Let* $X_1, ..., X_n$ *denote* n *vector fields. Then the* ***pull-back*** *of such n-form is the following map:*

$$(\Phi^*\omega)(X_1, ..., X_n) \coloneqq \omega(\Phi_*(X_1), ..., \Phi_*(X_n)) \tag{A.2}$$

[3, 4]

Given that the pull-back is defined with the use of a push-forward, and we have ensured that the push-forward is well-defined, such a definition requires no further conditions to be met.

A.2 An abstract Lie algebra

Same as any algebra, a Lie algebra is defined by defining multiplication on a vector space. Such multiplication is then called a **Lie bracket**. This one describes multiplication between vector fields as opposed to vectors themselves. It is possible to construct a real vector space with a set that consists of vector fields. As any T_pM already carries the structure of a real vector space, so does the set of all the tangent vectors one can define on some manifold. The properties of vector addition and scalar multiplication are still there when working with vector fields and by defining them the structure of a vector field is not affected. Isham [5, p. 99] specifically made that clear by defining those again. Let X and Y be two vector fields on M, let $a, b \in \mathbb{R}$ and $f \in C^\infty(M)$. We can define $aX + bY$ as

$$(aX + bY)f \coloneqq aXf + bYf. \tag{A.3}$$

Although this is now written in one equation, here it is visible that we scaled both of our vector fields and added them to get a new field. This was done pointwise for all points. The next step in defining vector field multiplication is to show that vector fields can also be viewed as a derivation meaning it is a linear map that follows a Leibniz rule. We can start again by looking at vector fields valued at some point $p \in M$. As was discussed before, a vector field valued at a point is just a tangent vector. That means we can write $(Xf)(p) = X_pf$.

Since X_pf *is* a derivation we know that for any $f, g \in C^\infty(M)$

$$X_p(fg) = fX_p(g) + gX_p(f). \tag{A.4}$$

And because $X_p f = (Xf)(p)$ we can now write

$$(Xfg)(p) = f(Xg)(p) + g(Xf)(p). \tag{A.5}$$

We have shown that a vector field follows the Leibniz rule *i.e.* that vector field is a derivation. This should be taken into account while defining vector field multiplication. The following explanation comes from Isham [5, Chapter 3.1.2].

If one were to define the multiplication of two vector fields as a composition $X \circ Y(f) \coloneqq X(Y(f))$ such an object would not follow the Leibniz rule as we would have the following

$$\begin{aligned} X \circ Y(fg) &= X(Y(fg)) = X(fY(g) + gY(f)) \\ &= X(f)Y(g) + fX(Y(g)) + X(g)Y(f) + gX(Y(f)) \\ &\neq f(X(Y(g))) + g(X(Y(f))) \end{aligned} \tag{A.6}$$

Similar would hold true for $Y \circ X(fg) \coloneqq Y(X(fg))$.

$$\begin{aligned} Y \circ X(fg) &= Y(X(fg)) = Y(fX(g) + gX(f)) \\ &= Y(f)X(g) + fY(X(g)) + Y(g)X(f) + gY(X(f)) \\ &\neq f(Y(X(g))) + g(Y(X(f))) \end{aligned} \tag{A.7}$$

Although, by themselves these do not follow Leibniz's rule, subtracting the second equation from the first results in a derivation.

$$(X \circ Y - Y \circ X)(fg) = g(X \circ Y - Y \circ X)(f) + f(X \circ Y - Y \circ X)(g) \tag{A.8}$$

This now means that $[X, Y] \coloneqq X \circ Y - Y \circ X$ *is* a vector field. In general, bracket $[X, Y]$ satisfies the following properties:

(i) Antisymmetry, *i.e.* $[A, B] = -[B, A]$

(ii) Jacobi identity, *i.e.* $[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0$

Definition A.3. *Let X, Y and Z be vector fields on a manifold M. An* ***abstract Lie algebra*** $(L,+,\cdot,[\![\,\_\,,\,\_\,]\!])$ *is a vector space L over a field K equipped with an abstract* ***Lie bracket*** $[\![\,\_\,,\,\_\,]\!]$ *that satisfies the following:*

(i) Bilinearity, i.e. $[\![\,\_\,,\,\_\,]\!]: L\times L\to L$

(ii) Antisymmetry, i.e. $[\![X,Y]\!]=-[\![Y,X]\!]$

(iii) Jacobi identity, i.e. $[\![X,[\![Y,Z]\!]]\!]+[\![Y,[\![Z,X]\!]]\!]+[\![Z,[\![X,Y]\!]]\!]=0$

[5, 15, 39]

Some authors define Lie algebra on a manifold that is a Lie group G with vector fields being left-invariant vector fields [12, 3]. Note that as vector space structure is implied by the definition of a Lie algebra, we will be writing $(L,[\![\,\_\,,\,\_\,]\!])$ instead of $(L,+,\cdot,[\![\,\_\,,\,\_\,]\!])$. One more note is that to distinguish between vector spaces defined over $\mathbb{R}$ and those defined over some other field the brackets used for $\mathbb{R}$-vector fields will be denoted $[\![\,\_\,,\,\_\,]\!]\coloneqq[\,\_\,,\,\_\,]$.

Let $\Gamma(TM)$ denote the set of all vector fields on M. Together with pointwise addition, scalar multiplication, and a Lie bracket such a set is considered a Lie algebra $(\Gamma(TM),[\,\_\,,\,\_\,])$ [15].

A subset of any Lie algebra L is considered a **Lie subalgebra** if it is closed under $[\![\,\_\,,\,\_\,]\!]$. So, to check if a given subset is a Lie subalgebra, one only needs to check that the output of the Lie bracket is again in this subset. One such subset is the set of left-invariant vector fields.

References

[1] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3 2014.

[2] Frederic Schuller. Topological manifolds and manifold bundles- lec 06 - frederic schuller. `https://www.youtube.com/watch?v=uGEV0Wk0eIk&list=PLPH7f_7ZlzxTi6kS4vCmv4ZKm9u8g5yic&index=6&ab_channel=FredericSchuller`, 2015.

[3] M. Nakahara. *Geometry, Topology and Physics, Second Edition*. Graduate student series in physics. Taylor & Francis, 2003.

[4] Loring W. Tu. An introduction to manifolds, 2011.

[5] C.J. Isham. *Modern Differential Geometry for Physicists*. World Scientific lecture notes in physics. World Scientific, 1999.

[6] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick. *Analysis, Manifolds and Physics Revised Edition*. Analysis, Manifolds and Physics. Elsevier Science, 1982.

[7] Mark Hamilton, J. D. *Mathematical Gauge Theory: With Applications to the Standard Model of Particle Physics*. Universitext. Springer International Publishing, Cham, 2017.

[8] (All authors are visible at ”View history” link given on the right hand side below the title). Bundle map. `https://en.wikipedia.org/wiki/Bundle_map` (version: 2022-12-15).

[9] Frederic Schuller. Construction of the tangent bundle - lec 10 - frederic schuller. `https://www.youtube.com/watch?v=XZcKSoI17r0&list=PLPH7f_7ZlzxTi6kS4vCmv4ZKm9u8g5yic&index=10&ab_channel=FredericSchuller`, 2015.

[10] Min Ru. Tangent bundle, vector bundles and vector fields. https://www.math.uh.edu/~minru/Riemann08/vd.pdf.

[11] Frederic Schuller. Tensor space theory ii: over a ring - lec 11 - frederic schuller. https://www.youtube.com/watch?v=V0TPgeiyWCo&ab_channel=FredericSchuller, 2015.

[12] S. Kobayashi and K. Nomizu. Foundations of differential geometry. https://books.google.hr/books?id=wn4pAQAAMAAJ, 1963.

[13] Frederic Schuller. Grassmann algebra and derham cohomology - lec 12 - frederic schuller. https://www.youtube.com/watch?v=QLnzIOGIvfo&list=PLPH7f_7ZlzxTi6kS4vCmv4ZKm9u8g5yic&index=12, 2015.

[14] J. Schwichtenberg. *Physics from Symmetry*. Undergraduate Lecture Notes in Physics. Springer International Publishing, 2015.

[15] Frederic Schuller. Lie groups and their lie algebras - lec 13 - frederic schuller. https://www.youtube.com/watch?v=mJ8ZDdA1OGY&t=1s&ab_channel=FredericSchuller, 2015.

[16] J.F. Cornwell. *Group Theory in Physics: An Introduction*. ISSN. Elsevier Science, 1997.

[17] (All authors are visible at "View history" link given on the right hand side below the title). Orthogonal group. https://en.wikipedia.org/wiki/Orthogonal_group#Lie_algebra),.

[18] (All authors are visible at "View history" link given on the right hand side below the title). Circle group. https://en.wikipedia.org/wiki/Circle_group,.

[19] Jonathan Evans. Lie groups and lie algebras: Matrix exponential. https://www.youtube.com/watch?v=eBq_DeJHfiU&list=PLN_4R2IuNu

uRgJb00X2J53Iq9qe7k1nyr&index=2&ab_channel=JonathanEvans, 2020.

[20] D. Bleecker. *Gauge Theory and Variational Principles.* Dover Books on Mathematics. Dover Publications, 2013.

[21] Frederic Schuller. Reconstruction of a lie group from its algebra - lec 18 - frederic schuller. https://www.youtube.com/watch?v=7qO5y6Es9Ns&t=2302s&ab_channel=FredericSchuller, 2015.

[22] (All authors are visible at "View history" link given on the right hand side below the title). Representation (mathematics). https://en.wikipedia.org/wiki/Representation_(mathematics) (version: 2024-01-09).

[23] (All authors are visible at "View history" link given on the right hand side below the title). Representation theory. https://en.wikipedia.org/wiki/Peter%E2%80%93Weyl_theorem#Representation_theory_of_connected_compact_Lie_groups (version: 2024-05-17).

[24] (Authors depend on a choice of specific article). Category:definitions/representation theory. https://proofwiki.org/wiki/Category:Definitions/Representation_Theory.

[25] Frederic Schuller. Representation theory of lie groups and lie algebras - lec 17 - frederic schuller. https://www.youtube.com/watch?v=h-d8TUg022A&ab_channel=FredericSchuller, 2015.

[26] (All authors are visible at "View history" link given on the right hand side above the title). Definition:linear representation. https://proofwiki.org/w/index.php?title=Definition:Linear_Representation&action=history (version: 2017-06-28).

[27] (Authors depend on a choice of specific article). Definition:linear representation/algebra. `https://proofwiki.org/wiki/Definition:Linear_Representation/Algebra` (version: 2020-03-08).

[28] (All authors are visible at "View history" link given on the right hand side below the title). Adjoint representation. `https://en.wikipedia.org/wiki/Adjoint_representation` (version: 2024-07-12).

[29] Frederic Schuller. Principal fibre bundles - lec 19 - frederic schuller. `https://www.youtube.com/watch?v=vYAXjTGr_eM&t=8474s&ab_channel=FredericSchuller`, 2015.

[30] (All authors are visible at "View history" link given on the right hand side below the title). Equivalence relation. `https://en.wikipedia.org/wiki/Equivalence_relation` (version: 2024-01-24).

[31] Another Roof. Defining every number ever. `https://www.youtube.com/watch?v=IzUw53h12wU&t=3s&ab_channel=AnotherRoof`, 2022.

[32] (All authors are visible at "View history" link given on the right hand side below the title). Homeomorphism. `https://en.wikipedia.org/wiki/Homeomorphism` (version: 2022-04-11).

[33] Claire (https://math.stackexchange.com/users/572257/claire). What is a simple way of understanding associated bundles? Mathematics Stack Exchange. `https://math.stackexchange.com/q/4797173` (version: 2023-10-31).

[34] Frederic Schuller. Associated fibre bundles - lec 20 - frederic schuller. `https://www.youtube.com/watch?v=q2GYZz6q3QI&t=1812s&ab_channel=FredericSchuller`, 2015.

[35] Frederic Schuller. Local representations of a connection on the base manifold: Yang-mills fields - lec 22. `https://www.youtube.com/watch?v=KhagmmNvOvQ&t=381s&ab_channel=FredericSchuller`, 2015.

[36] Frederic Schuller. Parallel transport - lec 23 - frederic schuller. https://www.youtube.com/watch?v=jGHaZc2fuX8&t=20s, 2015.

[37] (All authors are visible at "View history" link given on the right hand side below the title). Curvature. https://en.wikipedia.org/wiki/Curvature (version: 2024-08-28).

[38] Frederic Schuller. Curvature and torsion on principal bundles - lec 24 - frederic schuller. https://www.youtube.com/watch?v=j36o4DLLK2k, 2015.

[39] Frederic Schuller. Differential structures: the pivotal concept of tangent vector spaces - lec 09 - frederic schuller. https://www.youtube.com/watch?v=UPGoXBfm6Js&list=PLPH7f_7ZlzxTi6kS4vCmv4ZKm9u8g5yic&index=9&ab_channel=FredericSchuller, 2015.