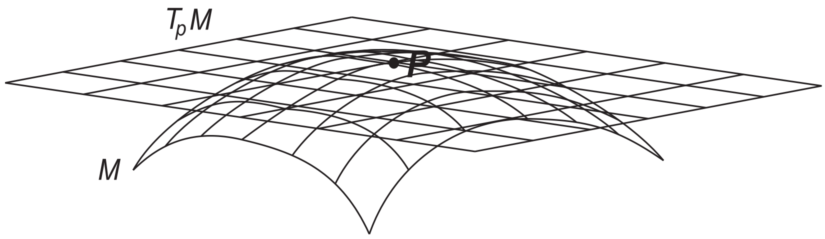
LULEÅ UNIVERSITY OF TECHNOLOGY

A Route Towards Gauge Theory

An Introduction to the Geometric Foundations of Modern Theoretical Physics



Stefan Eriksson Jon Häggblad Daniel Strömbom

March 2005

PROJECT REPORT

Department of Mathematics

Abstract. The aim of this report is to provide a brief account of the mathematics underlying a substatial part of modern theoretical physics, more specifically field theory, and in particular the gauge theories used to describe elementary particle interactions. Therefore, the elements of mathematics we consider mainly belong to the realms of differential geometry and topology, and is divided into five main chapters; Manifolds, Tensors, Differential Forms, Lie Theory and Bundles and Gauge Theory. Each chapter serve as an elementary introduction to the subject it concerns and the degree of depth is largely determined by what is optimal with respect to comprehension of the other chapters.

Sammanfattning. Syftet med denna rapport är att kort redogöra för matematiken som ligger till grund för en väsentlig del av den moderna teoretiska fysiken, mer specifikt fältteori och speciellt gaugeteorierna som används för att beskriva elementarpartikelväxelverkan. Matematiken vi behandlar tillhör därför i huvudsak områdena differentialgeometri och topologi och är uppdelad i fem huvudkapitel; mångfalder, tensorer, differentialformer, lieteori och buntar och gaugeteori. Varje kapitel utgör en elementär introduktion till området det rör och framställningens djup beror i huvudsak på vad som är optimalt med avseende på förståelse av de övriga kapitlen.

Contents

Introduction

The mathematics we will treat in this report has largely become essential in physics because of three physical insights. First of all it has become apparent that the universe cannot be thought of as a realization of Euclidean 3-space, not even if one adds a time dimension to the spatial part. The current theory dictating the relation between space and time and the properties of the world related to these, i.e. general relativity, states that space and time must be equal parts of one entity, space-time, and demand a more general space than the Euclidean to describe the universe and to this end spaces called manifolds are used. The use of manifolds then requires a generalization of the concept of vectors and this is obtained by the notion of tensors, and in order to do analysis with these new objects analogous to vector analysis one needs (exterior) differential forms. Secondly physicists have realized that symmetry is very important in the universe and thus for the description of it, therefore, algebra naturally enters into modern physics since this is a mathematical subject well equipped, and in a way designed for treating symmetries and thereby invariants. In particular the use of Lie groups have become essential in modern physics, largely related to the fact that in addition of being groups these objects are also manifolds. Finally, the above two realizations and a third, which is the importance of a distinction between local- and global properties of objects and phenomena in the physical universe, have made the mathematical objects called bundles powerful and, in a way, essential tools in modern theoretical physics. Thus, our opening claim seems fair and therefore reading this report should be time well spent for people engaged in physics as well as in mathematics.

The use of bundles in physics considered here is in field theory, in particularly (classical) gauge field theory, and the idea of letting fields represent physical entities must be one of the most important ever conceived. Classical field theory begun with the development of electrodynamics in the 20th century and through the general theory of relativity the subject got a very geometrical flavor. These developments in physics was possible largely due to the fact that in mathematics a theory of geometry had been developed, by B. Riemann and others, that naturally allowed one to represent and work with the entities of the physical theories and these entities became known as fields, in the cases here, the electromagnetic- and gravitational field. In this way mathematics paved way for physics, but the favor was to be returned. A monumental event in the history of field theory, commonly refered to as the gauge revolution, occurred when it was realized that gauge theories could be used to accurately describe elementary particle interaction. This insight changed the perceived nature of the physical fields, with respect to global vs. local properties, and forced one to replace the function representation of them on space-time by a more sophisticated object, a so called section of a fibre bundle. Nowadays, although modern field theory involving quantum fields is in its infancy, at least mathematically, it has in itself and through its generalization, e.g. string- and M-theory, repaid the favor to mathematics, in particular by providing new ways, using new invariants, to study the structure of low dimensional topological spaces, especially manifolds. These things will not be treated here though, since they lie well beyond the focus of this report.

The reader should note, that even though the contents of this report is largely motivated by its relation to the gauge theories of modern physics the report is a mathematical one. We begin by defining the basic concepts of manifold theory, tensors, differential forms and Lie theory in a purely mathematical fashion. These are then intertwined with the subsequent chapter on bundles into an exposition of gauge theoretical concepts through notions such as connection, curvature and covariant derivative. Finally the reader should note that an appendix has been added in which the necessary prerequisites from topology, analysis and algebra may be found.

Luleå University of Technology
March 11th 2005

S. Eriksson steeri-3@student.ltu.se
J. Häggblad jonhag-1@student.ltu.se
D. Strömbom danstr-2@student.ltu.se

CHAPTER 1

Manifolds

Traditionally we are used to do calculus in the familiar Euclidean space $\mathbb{R}^n$. Living as we do, in a world closely resembling this abstractation, it can be hard to realize the limitations that this framework imposes on us. So it would not come as a surprise that it took so long, from the ancient Greeks to the 19th century, to free ourselves from the Euclidean mind set of $\mathbb{R}^n$ as the most general space. With the groundwork of C. F. Gauss, B. Riemann was finally able to generalize this notion to the concept of manifolds. The fundamental idea behind manifolds is that although globally they can be very complicated, locally they always look like the familiar Euclidean space. What is meant by "globally" and "locally" can be illustrated by considering a sphere. For a tiny 2-dimensional creature living on this sphere, the sphere looks flat. Yet for a larger 3-dimensional creature observing from a distance, the sphere is obviously curved. It should be noted that examples like these should not be considered in detail, as the analogy is not exact. Yet it conveys the idea of "local" and "global" in a nice way. The second crucial idea conceived with the concept of manifolds is that although they can be embedded in $\mathbb{R}^n$ to help us visualize them, this is not required. A manifold can be a completely separate entity in itself. The natural example of this is of course the 4-dimensional model of the world used in the formulation of Einstein's General Relativity. Here the world is modeled as a curved 4-dimensional manifold that we can never visualize globally. The only thing we see is "flat" Euclidean space. This is quite natural since in a fundamental physical theory about the world we want the formulation to be independent and not depending on any external phenomena.

The goal of this chapter is to define in mathematical terms what a manifold is. Then we will develop the structure necessary for doing calculus on them.

1. Defining a Manifold

Before we can define a manifold we need the very important notion of a homeomorphism (not to be confused with homomorphism). Intuitively we can think of two homeomorphic spaces as such that one can in a continuous way transform one into the other and vice versa. There are many topological notions used here, especially in the first chapter, that if unsure of, should be looked up in the prerequisites.

Definition 1.1. *If a function between two topological spaces $f : M \to N$ is bijective with inverse f^{-1}, and the both maps f and f^{-1} are continuous, we say that f is a* ***homeomorphism*** *and the two topological spaces M and N are said to be homeomorphic.*

There are various ways of defining a manifold depending on how many "strange" spaces one wants to include. A metrizable condition on the manifold eliminates most, if not all of these. Yet, two different conditions, *Hausdorff* and *second countable* (which are covered in the preliminaries), are used instead, which one can show to be equivalent to the metrizable condition [**Spi99**]. This is to clearer emphasize the properties of the space.

Definition 1.2. *A* ***manifold****, or* ***topological manifold****, of dimension n is a topological space M such that*

(1) *M is a Hausdorff space*
(2) *M is second countable (has a countable base for the topology)*
(3) *$\forall x \in M$ there exist a neighborhood $U \subset M$ of x such that U is homeomorphic to an open subset of $\mathbb{R}^n$.*

The basic example of a manifold is the Euclidean space, as well as any open set inside. Any smooth boundary of a subset of Euclidean space is a manifold as well, such as the circle or the sphere.

Example 1.3 (S^1 and S^2). Two examples of manifolds are the unit circle S^1 in $\mathbb{R}^2$ and the unit sphere S^2 in $\mathbb{R}^3$.

Another type of manifold is **manifold with boundary**. It is defined similarly to a manifold except that U is homeomorphic to the closed set

$$\mathbb{H} = \{(x^1, \cdots, x^n) \in \mathbb{R}^n \mid x^n \geq 0\}.$$

We will, unless otherwise specified, mean a "ordinary" manifold when we talk about manifolds in our continued discussion.

2. The Differentiable Structure

To be able to do calculus on a manifold we need to introduce additional structure to facilitate this. Basically what we need to do first is to make sure that different homeomorphisms of a subset of a manifold M are compatible with each other in the sense that one can transform between them in a smooth way. If we let U, V be open subsets of a manifold M, then the two homeomorphisms

$$x : U \to x(U) \subset \mathbb{R}^n$$
$$y : V \to y(V) \subset \mathbb{R}^n$$

are called C^∞**-related** if

$$y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$$

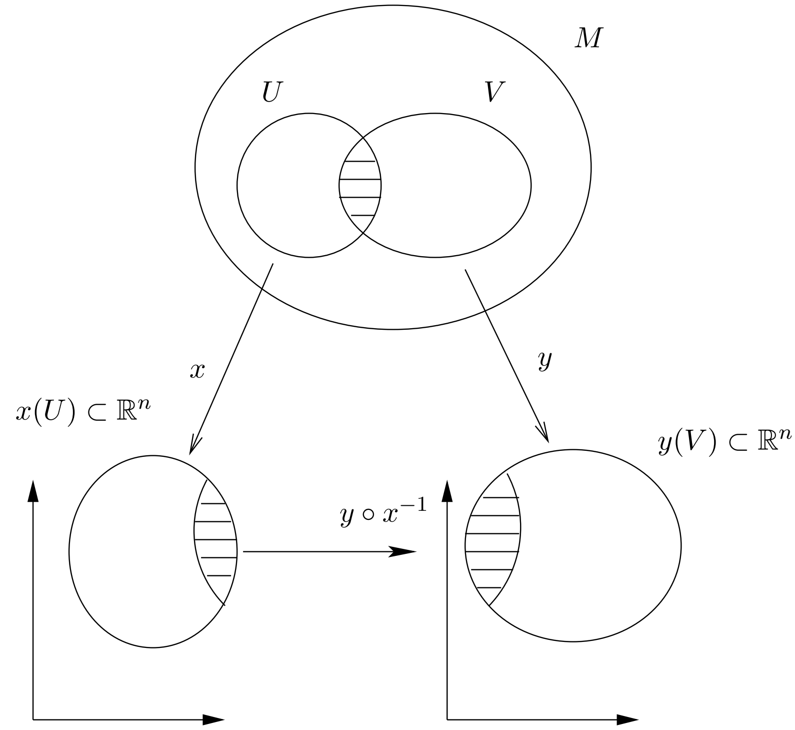


FIGURE 1. The composition of two homeomorhisms in an overlapping region on a manifold M.

$$x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$$

are C^∞. That is, the composite functions are infinitely many times differentiable (smooth). These mappings are illustrated in figure 1. An **atlas** $\mathcal{A}$ for a manifold M is a set of C^∞ related homeomorphisms whose domains cover M. An element (x, U) of an atlas is called a **chart** or a **coordinate system**. The notation (x, U) is chosen to emphasize the domain of the chart. The denotation "x" of the homeomorphisms instead of the otherwise common "ϕ" is chosen so that one more easily identifies the point $p \in M$ with the point $(x^1(p), x^2(p), \ldots, x^n(p)) \in \mathbb{R}^n$.

EXAMPLE 1.4 (S^1). An example of a manifold is the unit circle S^1 in $\mathbb{R}^2$. We cannot cover the entire circle using a single chart, and thus must use at least two overlapping charts. Here we show an example of such an atlas, altough there are of course many other atlases. Using four homeomorphisms

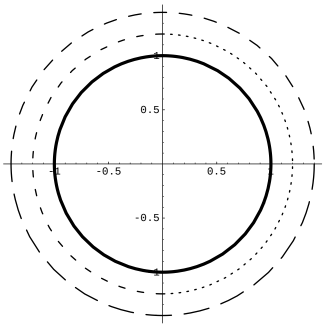


FIGURE 2. The four different charts covering the unit circle S^1 (solid) in example 1.4. U_1 dotted, U_2 dashed, U_3 longly dashed and U_4 very longly dashed.

$x^i : U_i \to x(U_i) \subset \mathbb{R}^1$ we get the charts

$$\begin{array}{lll} (x^1, U_1): & U_1 = \left\{(\mathrm{x},\mathrm{y})|\mathrm{x} > 0\right\} & x^1(\mathrm{x},\mathrm{y}) = \mathrm{y} \\ (x^2, U_2): & U_2 = \left\{(\mathrm{x},\mathrm{y})|\mathrm{x} < 0\right\} & x^2(\mathrm{x},\mathrm{y}) = \mathrm{y} \\ (x^3, U_3): & U_3 = \left\{(\mathrm{x},\mathrm{y})|\mathrm{y} > 0\right\} & x^3(\mathrm{x},\mathrm{y}) = \mathrm{x} \\ (x^4, U_4): & U_4 = \left\{(\mathrm{x},\mathrm{y})|\mathrm{y} < 0\right\} & x^4(\mathrm{x},\mathrm{y}) = \mathrm{x} \end{array}$$

where we have $\mathrm{x}^2 + \mathrm{y}^2 = 1$ for the unit circle. Note that the pair (x,y) is a point on the manifold, which here is embedded in $\mathbb{R}^2$. The inverse maps $(x^i)^{-1} : x(U_i) \subset \mathbb{R}^1 \to U_i$, are

$$\begin{aligned} (x^1)^{-1}(\mathrm{x}) &= ((1-\mathrm{x}^2)^{\frac{1}{2}}, \mathrm{x}) \\ (x^2)^{-1}(\mathrm{x}) &= (-(1-\mathrm{x}^2)^{\frac{1}{2}}, \mathrm{x}) \\ (x^3)^{-1}(\mathrm{x}) &= (\mathrm{x}, (1-\mathrm{x}^2)^{\frac{1}{2}}) \\ (x^4)^{-1}(\mathrm{x}) &= (\mathrm{x}, -(1-\mathrm{x}^2)^{\frac{1}{2}}) \end{aligned}$$

Again, $((1-\mathrm{x}^2)^{\frac{1}{2}}, \mathrm{x})$ are points on the manifold. If we want to convince ourself that the overlapping maps are smooth, we can as an example look at a point in $U_1 \cap U_3$. The reasoning are similar for all other points. In a point (x,y) in the overlap between U_1 and U_3 we have

$$\mathrm{y} = (1-\mathrm{x}^2)^{\frac{1}{2}} \qquad 0 < \mathrm{x} < 1, \;\; 0 < \mathrm{y} < 1.$$

From the inverse maps we get $(x^3)^{-1}(\mathrm{x}) = (\mathrm{x}, (1-\mathrm{x}^2)^{\frac{1}{2}})$, and $x^1 \circ (x^3)^{-1}(\mathrm{x}) = (1-\mathrm{x}^2)^{\frac{1}{2}}$, which obviously are smooth when $0 < \mathrm{x} < 1$, hence the manifold is smooth.

THEOREM 1.5. *If $\mathcal{A}$ is an atlas on the manifold M, then $\mathcal{A}$ is contained in a unique* ***maximal atlas*** *$\mathcal{A}'$. The maximal atlas $\mathcal{A}'$ is usually referred as the* ***differentiable structure*** *of the manifold.*

That we have a maximal atlas is dependent on Zorns lemma, which is equivalent to the debated axiom of Choice. A statement of **Zorns lemma** is as follows: *If S is a non-empty inductively ordered set, then S has a maximal element.* The statement above is proved analogous to the way one usually proves, the more familiar theorem, that every vector space V has a basis. That is, choose an arbitrary linearly independent subset $S \subset V$. If S does not constitute a basis then there exist a vector $v \in V$ that cannot be written as a linear combination of the elements in S, which means that we can include v in S and still have a linearly independent set $\tilde{S}$. If $\tilde{S}$ is still not a basis we repeat the procedure and this is done again and again until finally a basis is obtained, and it is Zorns lemma that guarantees that we in this way eventually will get a basis. Thus, the maximal atlas statement is proved by starting with, for instance, one atlas and then keep adding atlases in the same manner as in the vector space case, until a maximal atlas is obtained.

DEFINITION 1.6. *A **differentiable manifold** or C^∞-manifold is the pair $(M, \mathcal{A})$, where $\mathcal{A}$ denotes the maximal atlas for the manifold M.*

In the same way as homeomorphisms are used with manifolds, the notion of diffeomorphisms are used when dealing with differentiable manifolds.

DEFINITION 1.7. *If a function between two smooth (C^∞) manifolds $f : M \to N$ is bijective with inverse f^{-1}, and the both maps f and f^{-1} are smooth, we say that f is a **diffeomorphism** and the two manifolds M and N are said to be diffeomorphic.*

It should be noted that the concepts of homeomorphisms for topological spaces and diffeomorphisms for manifolds are similar to what isomorphisms are for vector spaces. From here on when we talk about a manifold M, we mean a differentiable manifold M. Also when we mention U, we always mean the subset $U \subset M$. Using this differentiable structure on a manifold we are now ready to define what a differentiable function is.

DEFINITION 1.8. *A function $f : M \to N$ is **differentiable** at $p \in U$ if $y \circ f \circ x^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable in the usual sense.*

For a function $g : \mathbb{R}^n \to \mathbb{R}$ we denote $D_i g(a)$ for

$$\lim_{h \to 0} \frac{g(a^1, \cdots, a^i + h, \cdots, a^n) - g(a)}{h}.$$

This with the familiar chain rule for two functions $g : \mathbb{R}^m \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$,

$$D_i(h \circ g)(a) = \sum_{j=1}^{n} D_j h(g(a)) \cdot D_i g^j(a),$$

we can define how the function $f : M \to \mathbb{R}$ is differentiated with respect to its charts.

$$\left.\frac{\partial f}{\partial x^i}\right|_p = D_i(f \circ x^{-1})(x(p)). \tag{1}$$

To avoid lengthy expressions involving multiple summations we will introduce a summation convention directly here in the beginning. This convention, the **Einstein summation convention**, means that whenever two indices in an expression appears both "up" and "down", it is summed over, e.g.

$$a^i e_i = \sum_i a^i e_i.$$

THEOREM 1.9. *Let (x, U) and (y, V) be two charts on the manifold M, and $f : M \to \mathbb{R}$ is differentiable map. Then on $U \cap V$ we have*

$$\frac{\partial f}{\partial y^i} = \frac{\partial f}{\partial x^j}\frac{\partial x^j}{\partial y^i}. \tag{2}$$

PROOF. Using the definition in equation (1) and the chain rule, gives

$$\begin{aligned}
\left.\frac{\partial f}{\partial y^i}\right|_p &= D_i(f \circ y^{-1})(y(p)) \\
&= D_i([f \circ x^{-1}] \circ [x \circ y^{-1}])(y(p)) \\
&= D_j(f \circ x^{-1})([x \circ y^{-1}](y(p))) \cdot D_i[x \circ y^{-1}]^j(y(p)) \\
&= D_j(f \circ x^{-1})(x(p)) \cdot D_i[x^j \circ y^{-1}](y(p)) \\
&= \left.\frac{\partial f}{\partial x^j}\right|_p \cdot \left.\frac{\partial x^j}{\partial y^i}\right|_p
\end{aligned}$$

□

We will often write this important result on operator form

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

3. The Tangent Space

For a curve in $\mathbb{R}^2$ or a surface in $\mathbb{R}^3$, what is meant by a tangent vector is easily defined and visualized. For an abstract space such as a manifold this is no longer the case. Fundamentally because the manifold can be the space itself, and not like for example the 1-dimensional curve embedded in $\mathbb{R}^2$, where the tangent vectors are simply the vectors in $\mathbb{R}^2$ that are tangent to the curve. That is, the tangent vectors "live" in a space outside of the space they are tangent to. Thus, for a manifold we must define tangent vectors on a manifold intrinsically of the manifold itself.

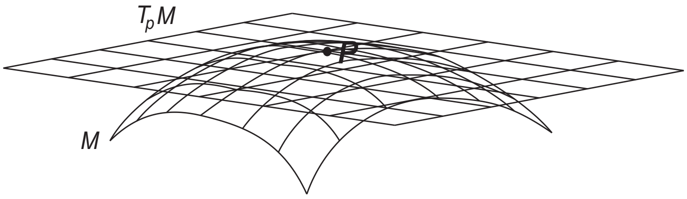


FIGURE 3. The 2-dimensional tangent space at a point p on a 2-dimensional manifold embedded in $\mathbb{R}^3$

DEFINITION 1.10. *A* ***tangent vector, vector*** *or* ***contravariant vector*** *X at p is a derivation (derivative operator) defined on the germs of functions as $X : C_p^\infty(M) \to \mathbb{R}$ such that*

(1) $X(\alpha f + \beta g) = \alpha X(f) + \beta X(g)$
(2) $X(f\ g) = X(f)\ g(p) + f(p)\ X(g)$

The first condition is the usual linearity condition. The second is the Leibniz property that we require derivations to satisfy. We can also define the tangent vector in a cumbersome, but more intuitive, way by letting X be an equivalence class of curves $c : (-\epsilon, \epsilon) \to M$ passing through a point $p \in M$ such that their tangents coincide in p. This definition is harder to work with, but included for illustrative purposes. Next we define the tangent vector space in the natural way.

DEFINITION 1.11. *The* ***tangent vector space*** *at $p \in M$, written T_pM, is the vector space spanned by all tangent vectors at $p \in M$.*

THEOREM 1.12. *In a given chart (x, U), a vector X at $p \in M$ admits the representation*

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

The basis of this representation,

$$\frac{\partial}{\partial x^1}\bigg|_p, \cdots, \frac{\partial}{\partial x^n}\bigg|_p$$

is called the ***coordinate basis****.*

This theorem is stated without proof, and as with the other theorems stated without proofs, they can be found in [**Spi99**]. Using theorem 1.12 we can let

$$X = a^i \frac{\partial}{\partial x^i}\bigg|_p \tag{3}$$

correspond to the vector a at the point $p \in U$ for the chart (x, U). We can now express a vector in terms of charts on the manifold. Next we want a way to transform a vector between different charts on $U \subset M$.

THEOREM 1.13. *Two vectors $a, b \in T_pM$ for different charts (x, U) and (y, V) where
$p \in U \cap V$ are the same vector if and only if they satisfy*

$$b^j = \left.\frac{\partial y^j}{\partial x^i}\right|_p a^i \tag{4}$$

PROOF. We can express a vector a at the point p as

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

according to equation (3) in the chart (x, U), and another vector b at the same point as

$$X = b^j \frac{\partial}{\partial y^j}.$$

Comparing those two expression we immediately get the result

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

at the point p. □

In physics this theorem is often taken as the definition of a contravariant vector. That is, one defines a contravariant vector in the way its coefficients transform between different charts (coordinate systems). Although this approach makes it possible to do calculations fairly early, it does not give the same level of understanding of what a contravariant vector is. We will see this later on when we define "the other" type of vectors, called covariant vectors. Grasping the fundamental difference between these two types of vectors can be quite difficult if one just considers how they transform.

Since we have a one-to-one correspondence [**Spi99**] between tangent vectors $X \in T_pM$ and the differential operators they are associated with, we do not make a distinction between a vector and the differential operator $\left.\frac{\partial}{\partial x^i}\right|_p$ corresponding to it. To emphasize this further we will use the notation

$$\partial_i = \frac{\partial}{\partial x^i}. \tag{5}$$

Most of the work uptil now has been to construct on manifolds, the usual structures we are accustomed to in $\mathbb{R}^n$, and we are now ready to start exploring these tools in the context of manifolds.

4. The Tangent Bundle

The tangent bundle is an important construction. We will later see that it is used to define vector fields on manifolds and in an example from classical mechanics the tangent bundle is the space of generalized velocities. It is also our first example of a fibre bundle.

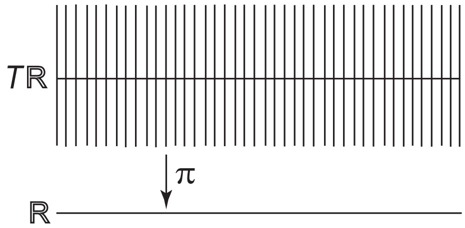


FIGURE 4. The manifold $\mathbb{R}$, its tangent space $T\mathbb{R}$ and the projection π

DEFINITION 1.14. *The **tangent bundle** of a manifold M is the disjoint union of all its tangent spaces*

$$TM = \bigcup_{p\in M} T_pM \tag{6}$$

together with a continuous map onto M

$$\pi : TM \to M \tag{7}$$

*such that at every point $p \in M$, T_pM is mapped to p. The inverse of this map at every point p, is called the **fibre** over the point p, denoted by $\pi^{-1}(p)$, and is obviously the set of all tangent vectors at the point p.*

The tangent bundle will be denoted by its map $\pi : TM \to M$ or simply just TM. Locally the tangent bundle for an open set $U \subset M$ looks like $U\times\mathbb{R}^n$. This is a property that we will return to later, as the tangent bundle belongs to the important class of structures called fibre bundles, which will be our focus later on in chapter 5. We will then also define, more precisely what a bundle topology is, whereas this definition of the tangent bundle is more informal.

DEFINITION 1.15. *Let $f : M \to N$ be differentiable, then at the point $p \in U \subset M$ the **differential** $f_{*p} : T_pM \to T_pN$ of the map f is defined by*

$$(f_{*p}X)(g) = X(g \circ f_p) \qquad \forall g \in C^\infty_{f(p)}(N).$$

The union of all f_* is written as $f_* : TM \to TN$ and gives rise to the commutative diagram

$$\begin{array}{ccc} TM & \xrightarrow{f_*} & TN \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi} \\ M & \xrightarrow{f} & N \end{array} \tag{8}$$

The principles of how the differential of a function behaves, can easier be seen if one consider the definition mentioned earlier about tangent vectors at a point p as an equivalence class of curves. Then the differential is simply the transportation of the tangent vector $c(t_0) \mapsto f\circ c(t_0)$. As mentioned earlier the tangent bundle TM looks locally like $U \times \mathbb{R}^n$ which is an important property. This we can for example use to see that the tangent bundle TM is itself a manifold. Let (x, U) be a chart, then all $X \in TM|U$ is uniquely of the form

$$X = a^i \partial_i, \quad p = \pi(X),$$

and thus the map $(x \circ \pi, X) : TM|U \to x(U) \times \mathbb{R}^n$ is simply

$$x_* : U \times \mathbb{R}^n \to x(U) \times \mathbb{R}^n$$

when we identify $TM|U$ with $U \times \mathbb{R}^n$ in the natural way. This is of course an homeomorphism and hence qualifies the tangent bundle TM to be a manifold in itself. If (y, V) is another chart then we can show [**Spi99**] that $y_* \circ (x_*)^{-1}$ is C^∞ and thus x_* and y_* are C^∞-related and can be extended to form a maximal atlas on TM. This makes TM, not surprisingly, a differentiable manifold as well.

Example 1.16 (Configuration space in mechanics). The configuration of a dynamical system with n degrees of freedom can be thought of as a n-dimensional manifold M, this we call the **configuration space**. Usually in physics these coordinates are denoted by q^i. A tangent vector on this manifold is thought of as a velocity vector and its components are written as $\dot{q}_i$ (instead of v^i, the reason for this will be expanded on later). We say that they are "generalized velocities" and TM is simply the space of all these velocities. For a simple pendulum, M is S^1 and for a dynamical system of two particles in $\mathbb{R}^2$, M is $\mathbb{R}^2 \times \mathbb{R}^2$.

5. Vector Fields and the Lie Bracket

A **section** of a tangent bundle $\pi : TM \to M$ is a continuous function $s : M \to TM$ such that $\pi \circ s$ is the identity of M. This map we can use to define vector fields in an extremely elegant way.

Definition 1.17. *A **vector field** X on M is a section $X : M \to TM$.*

We can understand a section if we think of how a vector field assigns to every point p on the manifold M a point in $\pi^{-1}(p) \subset TM$. The section s then describes in a smooth way in TM how the vector field on M is constructed.

Example 1.18 (Vector field on $\mathbb{R}$). If we have as a manifold M the real line $\mathbb{R}$ then our tangent bundle TM will be $\mathbb{R} \times \mathbb{R}$. The 0-section, which always exists, will then give a vanishing vector field on $\mathbb{R}$, and how a section s produces a vector field on M is illustrated in figure 5.

A vector field can thus be seen as a (locally) continuous set of tangent vectors to M. For a chart (x, U) we have

$$X(p) = a^i(p) \partial_i\big|_p \quad \forall p \in U.$$

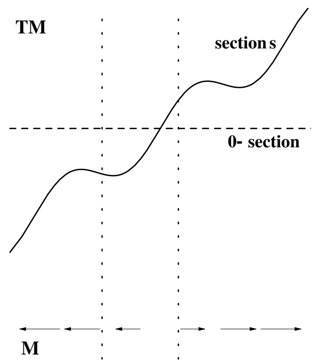


FIGURE 5. The section s in TM and the vector field on M for this section. Included is also the 0-section which gives a vanishing vector field on M

Thus we usually just write a vector field as

$$X = a^i \partial_i.$$

We define in the usual way the vector fields $X + Y$ and fX by

$$(X + Y)(p) = X(p) + Y(p)$$
$$(fX)(p) = f(p)X(p)$$

where X and $Y$are vector fields and $f : M \to \mathbb{R}$. Thus we see that for every $p \in M$ the set of all vector field on M constitutes a real vector space over $\mathbb{R}$, which we denote by $\chi(M)$. However if $X, Y \in \chi(M)$, then the operation XY, if viewed as repeated application, does not generally lie in $\chi(M)$. We can see this easily from the following

$$\begin{aligned} XY(fh) &= X(fY(h) + hY(f)) \\ &= X(f)Y(h) + fX(Y(h)) + X(h)Y(f) + hX(Y(f)) \end{aligned} \tag{9}$$

which certainly is not a vector field according to the Leibniz property of a vector in definition (1.10). To satisfy this it should have been $hX(Y(f)) + fX(Y(g))$. Although we can conclude that XY is not a vector field, we note that

$$\begin{aligned} YX(fh) &= Y(fX(h) + hX(f)) \\ &= Y(f)X(h) + fY(X(h)) + Y(h)X(f) + hY(X(f)), \end{aligned} \tag{10}$$

and if we now subtract equation (10) from equation (9) we get

$$(XY - YX)(fh) = fX(Y(h)) - fY(X(h) + hX(Y(f)) - Y(X(f)))$$

Thus we have

$$(XY - YX)(fh) = h(XY - YX)f + f(XY - YX)h.$$

which shows that $XY - YX$ behaves correctly compared to the definition 1.10 of a vector. This convinces us together with the easily shown fulfillnes of the linear properties that $XY - YX$ is a vector field. We can now define a composition rule under which $\chi(M)$ is closed.

Definition 1.19. *Let X, Y be (differentiable) vector fields on a manifold M and let the map $f : M \to \mathbb{R}$ be C^1. The* ***Lie bracket*** *is defined as*

$$[X, Y](f) = (XY - YX)(f)$$

and forms a vector field $[X, Y]$ which we usually write as

$$[X, Y] = XY - YX.$$

This is also called the ***Lie derivative*** *$\mathcal{L}_X Y$ of Y in the direction of X.*

An **algebra** is a vector space closed under a bilinear composition, and thus $\chi(M)$ together with the Lie bracket is an algebra. This algebra is also **anti-commutative**

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

and satisfies the **Jacobi identity** [**Spi99**]

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

Definition 1.20. *An algebra that has a anti-commutative composition that satisfies the Jacobi identity is called a* ***Lie algebra***

Thus obviously the set $\chi(M)$ of C^∞ vector fields on M is a Lie algebra. The notion of a Lie algebra is important and one which we will return to in our discussion of Lie theory.

CHAPTER 2

Tensors

The notion of a tensor is a natural generalization of scalars and vectors that we need when, either dealing with these objects on manifolds, or simple want to describe more complex behavior and transformations. In physics, one usually never need tensors of higher degrees than two, the one with the lowest degree higher than that of regular vectors. But that doesn't make the general framework any less interesting. Tensors traditionally were introduced using the "index approach", the reason for this is that it is simpler but can sometimes hide the central concept behind what E. Cartan called "the debauch of indices". We will instead use the more modern approach, which emphasize the tensor as a linear map. This also makes the important introduction of differential forms much more natural, as we shall see later on.

1. Covectors and the Dual Space

Before we can introduce the concept of a tensor we must define some notions we need to use. The first is that a **linear functional** α on V, where V is a vector space, is a linear function $\alpha : V \to \mathbb{R}$. From this we can define the very important concept of dual spaces.

Definition 2.1. *The set of all linear functions α on a vector space V forms a new vector space V^*, called the **dual space** of V. Elements of this space we call **covectors** or **covariant vectors**.*

The basis of the dual space V^* is related to the basis of V. If $e_1, e_2, \ldots, e_n$ is a basis for V then we define the basis of V^* to be $\sigma^1, \sigma^2, \ldots, \sigma^n$ where

$$\sigma^i(e_j) = \delta^i_j.$$

Since for a vector v we have

$$\sigma^i(v) = \sigma^i(e_j v^j) = \sigma^i(e_j)v^j = \delta^i_j v^j = v^i$$

Hence σ^i is a linear functional that picks out the i^{th} component of a vector. To see that σ^i forms a complete basis we note that, assuming that a linear combination $a_i\sigma^i$ is the zero functional:

$$0 = a_i\sigma^i(e_j) = a_i\delta^i_j = a_j.$$

Thus σ^i are linearly independent since all coefficients are zero. That they span the entire vector space we get from:

$$\alpha(v) = \alpha(v^i e_i) = v^i \alpha(e_i) = \sigma^i(v)\alpha(e_i) = \left(\alpha(e_i)\sigma^i\right)(v),$$

and hence we can write any linear functional α as a linear combination

$$\alpha = \alpha(e_i)\sigma^i = a_i\sigma^i. \tag{11}$$

Now we have the basic general definitions. Returning to manifolds we say that the dual space to the tangent space T_pM, denoted by T_pM^*, is called the **cotangent space**. Since T_pM^* is also a vector space we can use it to construct a bundle in the same way we did with T_pM.

DEFINITION 2.2. *Let M be a manifold, then the cotangent bundle of M is the disjoint union of all its cotangent spaces T_pM,*

$$TM^* = \bigcup_{p\in M} T_pM^*.$$

Once again this definition might seem a bit sloppy, but when we generalize the notion of bundles later on we shall make a more precise definition. The reason we introduce these different types of bundles early on is that we need them to define fields in an elegant way, but also to familiarize us with the concepts, as they are so important in the general framework of gauge theory. Using the general definition (1.15) of a differential, we can now use our new formalism to define an important special case. This is in fact so important that we shall from now mean this when we talk about differentials.

DEFINITION 2.3. *If $f : M^n \to \mathbb{R}$, then the* ***differential*** *of f at $p \in M$ is the function $df : T_pM \to \mathbb{R}$ defined by*

$$df(p)(X) = X(f) \quad X \in T_pM.$$

If (x, U) is a chart on M, then using the definition directly we see that

$$dx^i(p)\left(\partial_j\big|_p\right) = \delta^i_j,$$

and thus $dx^1(p), \ldots, dx^n(p)$ is a basis for T_pM^*. Hence with equation (11) we can express a linear functional $\alpha \in T_pM^*$ as

$$\alpha(p) = \alpha(p)\left(\frac{\partial}{\partial x^i}\right) dx^i(p) = a_i(p)dx^i(p)$$

Covectors written in this way we call **1-forms** or **differential 1-forms** if the coefficients $\alpha_i(p)$ are differentiable. In the same way that we defined fields of vectors on a manifold M, we can define fields of covectors as sections on the cotangent bundle.

DEFINITION 2.4. *A* ***covector field*** *α on M is a section $\alpha : M \to TM^*$*

In an analog way to what we did with vector fields we define the sums of covector fields and the product of functions and covector fields by

$$(\alpha + \beta)(p) = \alpha(p) + \beta(p)$$
$$(f\alpha)(p) = f(p)\alpha(p).$$

Thus we can write

$$\alpha = a_i dx^i$$

for all covectors on M. Using this we can write the differential df in these coordinate basis, which is a classical formula.

THEOREM 2.5. *Let f be a C^∞ function and (x, U) a chart on the manifold M, then on $U \subset M$ we have*

$$df = \frac{\partial f}{\partial x^i} dx^i \tag{12}$$

PROOF. If $X \in T_pM$ is a vector

$$X = a^i \partial_i\big|_p,$$

then we have

$$a^i = X(x^i) = dx^i(p)(X).$$

Thus

$$df(p)(X) = X(f) = a^i \partial_i f\big|_p = \partial f\big|_p dx^i(p)(X),$$

which in traditional notation is

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

□

How the components of a covector transform between different charts can easily be derived in a similar way to the case with vectors.

THEOREM 2.6. *If (x, U) and (y, V) are two charts on M, then the components of a covector $\alpha = \alpha_i dx^i$ on $U \cap V$ transform according to*

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

PROOF. Since we can always express a covector β in terms of the coordinate basis for different charts, we can in the overlap $U \cap V$ write

$$\beta = \alpha_i dx^i. \tag{13}$$
$$\beta = \beta_j dy^j. \tag{14}$$

From equation (12) we have

$$dx^i = \frac{\partial x^i}{\partial y^j} dy^j,$$

substituting this into (13) we get

$$\beta = \alpha_i \frac{\partial x^i}{\partial y^j} dy^j.$$

Finally comparing with (14) gives

$$\beta_j = \frac{\partial x^i}{\partial y^j} \alpha_i.$$

□

Again as mentioned for the case of vectors, this result is often used to define covectors in physics (and older books on differential geometry). Although it gives a concrete definition early on, it is difficult to see what these transformation rules actually mean. That is why we use the more modern approach of emphasizing the linear transformation properties of vectors and covectors. More specifically we can use the definition of the dual space to say that a vector X is a map

$$X : V^* \to \mathbb{R}, \tag{15}$$

and a covector α is a map

$$\alpha : V \to \mathbb{R}. \tag{16}$$

These two definitions are very important and it is crucial to understand them for the continued discussion.

Example 2.7 (Phase space in mechanics). Let M be the configuration space of a dynamical system. The Lagrangian of this system is then

$$L = L(q, \dot{q}).$$

Thus we can view the Lagrangian as a function on the space of generalized velocities TM

$$L : TM \to \mathbb{R}.$$

This formulation is called the Lagrangian formulation in mechanics. The Hamiltonian formulation is in mechanics usually viewed as a simple change of variable from q and $\dot{q}$ to q and p,

$$p_i(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i}. \tag{17}$$

To analyze this change of variables we see how they transform under a change of coordinates. First, if (q_1, U) and (q_2, V) are two charts on M we get

$$\begin{aligned} q_2 &= q_2(q_1) \\ \dot{q}_2^i &= \frac{\partial q_2^i}{\partial q_1^j} q_1^j \end{aligned} \tag{18}$$

Thus we see the tangent bundle structure these coordinates have. Now with the new p variables we see that they transform according to

$$(p_2)_i = \frac{\partial L}{\partial \dot{q}_2^i} = \frac{\partial L}{\partial q_1^j}\frac{\partial q_1^j}{\partial \dot{q}_2^i} + \frac{\partial L}{\partial \dot{q}_1^j}\frac{\partial \dot{q}_1^j}{\partial \dot{q}_2^i} = \frac{\partial L}{\partial \dot{q}_1^j}\frac{\partial \dot{q}_1^j}{\partial \dot{q}_2^i} = \frac{\partial L}{\partial \dot{q}_1^j}\frac{\partial q_1^j}{\partial q_2^i} = \frac{\partial q_1^j}{\partial q_2^i}(p_1)_j$$

since q_1 does not depend on $\dot{q}_2$ and $\frac{\partial \dot{q}_1^j}{\partial \dot{q}_2^i} = \frac{\partial q_1^j}{\partial q_2^i}$ from equation (18). Hence p is a *covector* and we can view the p's and q's as coordinates on the *cotangent bundle*. Equation (17) can thus be seen as a map

$$p : TM \to TM^*.$$

This space TM^* is called **phase space** in mechanics

Example 2.8 (Vectors and covectors in mechanics). When learning classical physics in $\mathbb{R}^3$, one usually doesn't distinguish between vectors and covectors and this can thus sometimes give rise to some confusion. Although we could see in the previous example that velocity is a vector and momentum is a covector by looking how they transform, there is a simpler way to see this. The most natural vectorial quantity is the *radius vector* $\mathbf{r}$ of a point in space relative the origin. Hence the *velocity* $\mathbf{v} = \frac{d\mathbf{r}}{dt}$ is also a vector, as it is the derivative of $\mathbf{r}$ with respect of a scalar quantity. On the other hand if we consider the potential energy U as a scalar, it's relation to force is $dU = -\mathbf{F} \cdot d\mathbf{r}$. Thus force is a linear map of vectors into scalars $\mathbf{F}(d\mathbf{r}) = -dU$, and hence a covector (1-form). This in turn implies, through the Newton equation $\mathbf{F} = \frac{d\mathbf{p}}{dt}$, that the momentum $\mathbf{p}$ also is a covector.

Example 2.9 (The electric field). The electric field $\mathbf{E}$, although often thought of as a vector field, is actually a field of covectors (1-forms) mapping the infinitesimal vector $d\mathbf{r}$ into the infinitesimal potential difference $-dV = \mathbf{E}(d\mathbf{r})$.

2. The Pull-back of Covectors

Covectors has an important property that vectors don't have and that is that we can globally "pull" a covector field from a manifold N to a manifold M with a map $f : M \to N$. Although we can do something similar locally for vectors, it is not guaranteed that we can do it globally.

Definition 2.10. *Let $f : M \to N$ be a C^∞ map between two manifolds and $\alpha \in T_{f(p)}N^*$ a covector. Then we can use the differential f_* to define the* ***pull-back*** *$f^*\alpha \in T_pM^*$ by*

$$(f^*\alpha)(X) = \alpha \circ f_{*p}(X) \qquad X \in T_pM$$

That is we pull the function $\alpha : N \to \mathbb{R}$ "back" to $f^*\alpha : M \to N \to \mathbb{R}$. The reason for the terminology "back" is because it gives rise to the commutative diagram

$$\begin{array}{ccc} T_pM^* & \xleftarrow{f^*} & T_{f(p)}N^* \\ \pi\big\downarrow & & \big\downarrow\pi \\ p & \xrightarrow{f} & f(p) \end{array} \tag{19}$$

This is also a source of confusion about the name "contravariant" and "covariant". In differential geometry vectors are called "contravariant" and covectors "covariant", but category theorists would do the opposite and call our "contravariant" vectors "covariant" since they with f_* transform "the same way" on the commutative diagram (8) as f. Our "covariant" vectors would then be called "contravariant" since they with f^* transform "the other way" as f in the commutative diagram (19). The reason for the classical terminology in differential geometry can be seen by considering charts on $\mathbb{R}^n$. If $x(v_i) = e_i$ then

$$x(a^1v_1 + \cdots + a^nv_n) = (a^1, \ldots, a^n).$$

If y is another such chart then

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

Comparing this with equation (12) we see that the differentials

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

transform "in the same way" as the coordinates x^i, and hence are called "covariant".

3. Contravariant Tensors

With our current understanding of vectors and covectors as the linear maps (15) and (16), then it is straightforward to generalize this to the concept of tensors. What we mean by **multilinear** here is that a function $T : V_1 \times \cdots \times V_n \to \mathbb{R}$ is multilinear if $T(v_1, \ldots, v_n)$ is linear in each argument, provided the rest is held fixed. The first type of tensor we will deal with is contravariant tensors. They are the natural generalizations of vectors.

DEFINITION 2.11. *A* ***contravariant tensor of degree r*** *or* ***(r,0)-tensor*** *is a multilinear mapping*

$$T : \underbrace{V^* \times \cdots \times V^*}_{r} \to \mathbb{R}.$$

The vector space of all contravariant r-tensors is denoted by

$$V \otimes \cdots \otimes V = \otimes^r V,$$

where these circle signs is the tensor product which is defined in the natural way.

Definition 2.12. *Let $T \in \otimes^r V$ and $S \in \otimes^k V$, then the* ***tensor product*** *$T \otimes S \in \otimes^{r+k} V$ is defined by*

$$T \otimes S(v_1, \dots, v_r, v_{r+1}, \dots, v_{r+k}) = T(v_1, \dots, v_r) \cdot S(v_{r+1}, \dots, v_{r+k})$$

If $e_1, \dots, e_n$ is a basis for V, then the basis for $\otimes^r V$ is

$$e_{i_1} \otimes \cdots \otimes e_{i_k} \qquad 1 \leq i_1, \dots, i_r \leq n$$

which thus has dimension n^r. The tangent bundle can also easily be generalized in a straightforward way. For the same reasons as with vectors, bundles allow us to define fields in an elegant way.

Definition 2.13. *Let M be a manifold, then the bundle*

$$\otimes^r TM = \bigcup_{p \in M} \otimes^r T_p M$$

is called a ***contravariant tensor bundle of order*** *r.*

Definition 2.14. *A section of $\otimes^r TM$ is called a* ***contravariant tensor field of order*** *r.*

Using the coordinate basis of the tangent space T_pM we get in an analogue way a coordinate basis for $\otimes^r T_p M$, that is the tensor products

$$\partial_{i_1}\big|_p \otimes \cdots \otimes \partial_{i_r}\big|_p \in \otimes^r T_p M \qquad 1 \leq i_1, \dots, i_r \leq n$$

is a basis for $\otimes^r T_p M$. So on U, every contravariant tensor field T of order r can written

$$T(p) = T^{i_1 \dots i_r}(p)\, \partial_x^{i_1}\big|_p \otimes \cdots \otimes \partial_x^{i_r}\big|_p.$$

This is usually shortened to

$$T = T^{i_1 \dots i_r}\, \partial_{i_1} \otimes \cdots \otimes \partial_{i_r}.$$

We define the addition of two contravariant tensor fields and the product of a function and a tensor field in the obvious way by

$$(T_1 + T_2)(p) = T_1(p) + T_2(p)$$
$$(fT)(p) = f(p)T(p).$$

We can also define a new tensor field of order $r+k$ by using the tensor product on two tensors fields of order r and k by

$$(T_1 \otimes T_2)(p) = T_1(p) \otimes T_2(p),$$

operating on $T_pM^* \times \cdots \times T_pM^*$ $r + k$ times.

Theorem 2.15. *If (x, U) and (y, V) are two different charts on a manifold M, then on $U \cap V$ the* components *of a contravariant tensor T transform according to*

$$T'^{j_1 \dots j_r} = \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_r}}{\partial x^{i_r}} T^{i_1 \dots i_r}$$

It should be pointed out that this result, like the one for vectors and covectors, are motivated for use as definitions in the physics literature as they emphasize the property, important in physics, that tensors represent physical quantities which are independent of coordinate system (chart). In our case this property is obvious in our more thorough construction of manifolds, but should yet be noted.

It is easy to see the naturality of these generalizations from vectors to contravariant tensors and thus there is not much to say about them outside of that they are very important, both as representation of physical quantities and as basis for other important objects, most notably the metric tensor. Their importance will become clearer later on.

4. Covariant Tensors

Covectors can be generalized in much the same way as with contravariant vectors, and thus we will only state the summarized results.

Definition 2.16. *A* ***covariant tensor of degree r*** *or* ***(0,r)-tensor*** *is a multilinear mapping*

$$T : \underbrace{V \times \cdots \times V}_{r} \to \mathbb{R}.$$

The vector space of all covariant r-tensors is denoted by

$$V^* \otimes \cdots \otimes V^* = \otimes^r V^*$$

and their basis is

$$\sigma^{i_1} \otimes \cdots \otimes \sigma^{i_r} \quad 1 \leq i_1, \ldots, i_r \leq n,$$

where the tensor product is defined in an analogous way as with contravariant tensors.

Definition 2.17. *Let M be a manifold, then the bundle*

$$\otimes^r TM^* = \bigcup_{p \in M} \otimes^r T_pM^*$$

is called a ***covariant tensor bundle of order*** *r.*

A section of $\otimes^r TM^*$ is called a **covariant tensor field of order** r. Thus using the basis

$$dx^{i_1}(p) \otimes \cdots \otimes dx^{i_r}(p) \in \otimes^r T_pM^* \qquad 1 \leq i_1, \ldots, i_r \leq n$$

for $\otimes^r T_pM^*$, we can thus write every covariant tensor T of order r as

$$T = T_{i_1 \ldots i_r}\, dx^{i_1} \otimes \cdots \otimes dx^{i_r}.$$

Addition of these covariant fields and the product of a function and a covariant tensor is defined in the same way (point wise) as contravariant tensors.

THEOREM 2.18. *If (x,U) and (y,V) are two different charts (coordinate systems) on a manifold M, then on $U \cap V$ the* components *of a covariant tensor T transform according to*

$$T'_{j_1 \ldots j_r} = \frac{\partial x^{i_1}}{\partial y^{j_1}} \cdots \frac{\partial x^{i_r}}{\partial y^{j_r}} T^{i_1 \ldots i_r}$$

Something that we could not do with contravariant tensors, but can with covariant tensors, is that similarly in the case of covectors we can define the pull-back. Let $f : M \to N$ then the map f^* takes covariant tensor fields T of order r on N to covariant tensor fields f^*T of order r on M by

$$f^*T(p)(X_{1_p}, \ldots, X_{r_p}) = T(f(p))(f_* X_{1_p}, \ldots, f_* X_{r_p}).$$

That is, the pull-back is a map

$$f^* : \otimes^r N^* \to \otimes^r M^*,$$

which can be shown to be an algebra homomorphism, that is, $f^*(\alpha \otimes \beta) = (f^*\alpha) \otimes (f^*\beta)$.

5. Mixed Tensors

Often we have **mixed tensors**, written as **(r,s)-tensor**. This is simply a multilinear mapping

$$T : \underbrace{V^* \times \cdots \times V^*}_{r} \times \underbrace{V \times \cdots \times V}_{s} \to \mathbb{R}$$

with the tensor product $T \otimes S$ of a contravariant tensor and a covariant tensor defined in the natural way as

$$T \otimes S(v_1, \ldots, v_r, v_{r+1}, \ldots, v_{r+k}) = T(v_1, \ldots, v_r, v_{r+1}) \cdot S(v_{r+1}, \ldots, v_{r+k}).$$

Thus every mixed tensor T, can be written as

$$T = T^{i_1 \ldots i_r}_{j_1 \ldots j_s} \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$$

in the coordinate basis.

THEOREM 2.19. *If (x,U) and (y,V) are two different charts on a manifold M, then on $U \cap V$, the* components *of a mixed tensor T transform according to*

$$T'^{\mu_1 \ldots \mu_r}_{\nu_1 \ldots \nu_s} = \frac{\partial x^{i_1}}{\partial y^{\nu_1}} \cdots \frac{\partial x^{i_r}}{\partial y^{\nu_s}} \frac{\partial y^{\mu_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{\mu_r}}{\partial x^{i_r}} T^{i_1 \ldots i_r}_{j_1 \ldots j_s}$$

This equation exemplifies the monstrosities that classical differential books, and to some extent physics books, are filled with. Altough they concretely state how the components transform, the sheer number of indices easily hides the important transformation properties.

An important mixed tensor is the $(1,1)$-tensor

$$T : V \times V^* \to \mathbb{R}.$$

These type of tensors arises from linear transformations $A : V \to V$ by the formula

$$T(v, \alpha) = \alpha(Av).$$

Actually each such transformation A gives rise to a tensor $T : V \times V^* \to \mathbb{R}$, and this correspondence is linear and bijective and hence is an isomorphism. Thus we shall not distinguish between a linear transformation A and its associated mixed tensor T, a linear transformation A *is* a mixed tensor (with components A^i_j). Not knowing the difference here is a source of confusion in elementary linear algebra, since the matrix of a linear transformation A is usually written there as a_{ij}, and the difference between bilinear forms and linear transformations are not pointed out. From the three matrices a_{ij}, a^{ij} and a^i_j, the first two define bilinear forms (on V and V^* respectively), while the last is the matrix of a linear transformation.

This isomorphism can be used to define the contraction of a mixed tensor $T : V \times V^* \to \mathbb{R}$ by taking the trace of the correspondig transformation matrix $A : V \to V$ to get a scalar.

THEOREM 2.20. *If $T^{\cdots i \cdots}_{\cdots j \cdots}$ are the components of a (r, s)-tensor, the the* ***contraction*** *on a pair of indices i, j, defined by the components $T^{\cdots i \cdots}_{\cdots i \cdots}$, defines a $(r-1, s-1)$-tensor.*

6. Scalar Product and the Metric Tensor

An important second order covariant tensor is the metric tensor. To see how it arises we consider the scalar product. Assume $e_1, \ldots, e_n$ is a basis of V with $v, u \in V$, $v = e_i v^i$ and $u = e_i u^i$. Then

$$\langle v\,,\, u\rangle = \langle e_i v^i\,, e_j u^j\rangle = v^i \langle e_i\,, e_j\rangle u^j.$$

Now we can introduce the metric tensor (g_{ij}) as the matrix in the last expression

$$g_{ij} = \langle e_i\,, e_j\rangle. \tag{20}$$

Here we can also note that if the basis is orthonormal, then the metric tensor would simply be the identity matrix. This is the case in elementary linear algebra, where the metric tensor never is introduced at all. Furthermore, since the scalar product is a bilinear function, that is, linear in each of its arguments, it is possible to introduce the covariant version (2.1) of each vector v in a vector space V with a scalar product. That is we can define the covariant vector α by the functional

$$\alpha(e_i) = \langle v, e_i\rangle, \tag{21}$$

which together with the expression $\alpha = \alpha_i \sigma^i$ allows us to write

$$\alpha = \alpha(e_i)\sigma^i = \langle v, e_i\rangle\sigma^i,$$

and hence by equation (20)

$$\alpha = (v^j g_{ji})\sigma^i.$$

The covariant version of the vector v thus has the components $\alpha_i = v^j g_{ji} = g_{ij} v^j$. Normally the components of this covariant vector is referred to as v_i to make the connection with the vector v more easily seen. This procedure is usually referred to as lowering or raising indices. The inverse of the metric tensor $(g_{ij})^{-1}$ is normally written as g^{ij}, and with the help of this we can, in an analogous way, define the components of the contravariant version of the covector α as

$$v^i = g^{ij} v_j$$

7. Riemannian Manifolds

Now when we have introduced the concept of a metric, it is possible to assign a specific such to a manifold. Here we will discuss the Riemannian metric.

Definition 2.21. *A* ***Riemannian metric*** *on a manifold M assigns, in a differentiable way, a positive definite scalar product $\langle\,,\rangle$ for each tangent space T_pM. A* ***Riemannian manifold*** *is then a manifold equipped with a Riemannian metric. A* ***pseudo-Riemannian metric*** *is a metric where we allow the scalar product to be non-positive definite. The resulting manifold is then naturally called a* ***pseudo-Riemannian manifold****.*

In a specific chart (x, U) on M it is possible to write the Riemannian metric with the help of the metric tensor as

$$\langle\,,\rangle = g_{ij} dx^i \otimes dx^j.$$

So the scalar product of two tangent vectors $X, Y \in T_pM$ will have the form

$$\langle X, Y\rangle = g_{ij} X^i Y^j$$

at all points on M. Now we can see that the concept of length of a tangent vector is possible to define in a natural way, in accordance with our concept of length in $\mathbb{R}^n$, as

$$|X| = \sqrt{\langle X, X\rangle} = \sqrt{g_{ij} X^i X^j}.$$

This is a direct consequence of the introduction of a metric on the manifold.

Example 2.22 (Euclidean metric). The simplest example of a Riemannian metric is $\mathbb{R}^3$ equipped with the usual metric. The metric tensor is thus

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and from this we can construct the distance $d(x, y)$ between two points x, y

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

The signature for this metric is $(+, +, +)$

EXAMPLE 2.23 (The Minkowski metric). An example of a pseudo-Riemannian metric is the well known Minkowski metric used in special relativity. This is $\mathbb{R}^4$ equipped with a metric tensor of the form (with c=1)

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

which obviously has signature $(+,-,-,-)$. The distane between two points x, y will with this metric be

$$d(x,y) = \sqrt{(x_1-y_1)^2 - (x_2-y_2)^2 - (x_3-y_3)^2 - (x_4-y_4)^2}.$$

We note here that, in contrast to example 2.23, it is possible with the Minkowski metric to have $d(x,y) = 0$ without $x = y$, which makes this metric a pseudo-Riemannian metric. The points were $d(x,y) = 0$ forms a cone which sometimes in physics is referred to as the light cone.

DEFINITION 2.24. *For a Riemannian manifold, the* ***gradient vector*** ∇f *is defined as the contravariant vector associated to the covector* df *by*

$$df(v) = \langle \nabla f, v \rangle.$$

The gradient vector has the coordinates $(\nabla f)^i = g^{ij}\partial_j f$. We also see that if we have an orthogonal basis we will have the ordinary partial derivatives as components, since g^{ij} will be the identity matrix.

8. A Note on Cartesian-tensors

It should be mentioned here that the objects encountered in vector analysis in flat Euclidean space, usually called Cartesian-tensors, are not tensors in the general sense. This is also why we write “Cartesian-tensors” and not “Cartesian tensors”. The reason behind this is that in standard vector analysis, Cartesian-tensors are defined as objects which “are the same” under transformations between orthogonal coordinate systems (charts). General tensors on the other hand “are the same” under transformations between all coordinate systems. Thus tensors are Cartesian-tensors but Cartesian-tensors are not necessarily tensors. Another example of these special type of less restrictive tensors are Lorentz-tensors.

CHAPTER 3

Differential Forms

We now turn to an interesting part of the more modern formulations of geometry, and that is the concept of an exterior algebra. Grassmann introduced this new algebra in the middle of the 19th century as a vast generalization of the scalar and vector products in use today in vector analysis. This is also closely connected to the objects known as exterior forms, which much of modern differential geometry and theoretical physics is formulated in. Although one can view exterior forms as "objects one integrates on manifolds", we will only cover the differential part, not the integral part to avoid a too large detour.

1. The Exterior Algebra

In the same way as we introduced multilinearity as an fundamental property of tensors, the notion of antisymmetry is the main new property we need to define exterior forms. A tensor $T \in \otimes^r V^*$ is called **antisymmetric** if

$$T(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = -T(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_r)$$

for each pair of entries.

Definition 3.1. *An* ***(exterior) k-form*** *is an antisymmetric covariant k-tensor and the set of all k-forms is denoted by $\wedge^k V^* \subset \otimes^k V^*$.*

From now on we will drop the "exterior" part when referring to k-forms. In this section, where we construct the machinery behind the exterior algebra, will unfortunately like the section on tensors be very technical. Continuing, the important property of these k-forms is thus that they are antisymmetric. Now that we have these objects, we must define the operations used on them to get the full exterior algebra. Obviously $\wedge^k V^*$ is closed under addition of the k-forms and multiplication of a scalar function, yet for the tensor product this is not the case. That is for $\alpha, \beta \in \wedge^k V^*$, the tensor $\alpha \otimes \beta$ does not in general lie in $\wedge^k V^*$, and hence is not a form. Thus we must define a "product" that keeps this antisymmetry of the forms.

Let S_n denote the symmetric group of all permutations σ of $\{1, \cdots, n\}$. Then we define the "alternation" of a tensor to

$$\mathrm{Alt}\, T(v_1, \cdots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}\, \sigma \cdot T(v_{\sigma(1)}, \cdots, v_{\sigma(k)}).$$

Thus Alt is a projection operator $\otimes^k V^* \mapsto \wedge^k V^*$.

Definition 3.2. *The* ***exterior product (wedge product*** *or* ***Grassmann product)*** *of two forms $\omega \in \wedge^k V^*$ and $\eta \in \wedge^n V^*$ is defined by*

$$\omega \wedge \eta = \frac{(k+n)!}{k!n!}\,\mathrm{Alt}(\omega \otimes \eta)$$

The "funny" coefficient here is mainly to get nicer looking calculations later on, it is the alternation operator that is the important part of the definition.

Proposition 3.3. *The properties of the exterior product is*

1. *bilinear.*
2. *$f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta$.*
3. *anticommutative: $\omega \wedge \eta = (-1)^{kn}\eta \wedge \omega$.*

Since forms are also tensors we can easily see that the basis for $\wedge^k V^*$ is

$$\sigma^{i_1} \wedge \cdots \wedge \sigma^{i_k} \quad 1 \leq i_1 < \cdots < i_k \leq n,$$

where $\sigma^1, \ldots, \sigma^n$ is the basis for V^*. The dimension of $\wedge^k V^*$ is thus

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Hence we can now define the direct sum, $\wedge^* V^* = \wedge^0 V^* \oplus \cdots \oplus \wedge^n V^*$, of all these forms as the **Exterior** or **Grassmann algebra**. Here we define the space of 0-forms as just $\mathbb{R}$. Now from this structure we can derive some results.

With some algebra one can show [**Fra04**] that if $\alpha^1, \ldots, \alpha^k$ are 1-forms and $v_1, \ldots, v_k$ is an k-tuple of vectors, then

$$\alpha^1 \wedge \cdots \wedge \alpha^k(v_1, \ldots, v_n) = \det[\alpha^j(v_i)]. \tag{22}$$

An elegant theorem with which one can discuss linear independence in a nice coordinate-free way is the following.

Theorem 3.4. *The 1-forms $\alpha^1, \ldots, \alpha^r$ are linearly dependent if and only if*

$$\alpha^1 \wedge \cdots \wedge \alpha^r = 0.$$

Proof. If $\alpha^k = a_i\alpha^i$, then $\alpha^1 \wedge \cdots \wedge \alpha^k \wedge \cdots \wedge \alpha^r$ will be a sum of terms, each having a repeated α^i, and thus the product will be zero. If on the other hand the α's are linearly independent, then we can complete them to a basis $\alpha^1, \ldots, \alpha^n$. Letting $f_1, \ldots, f_n$ be the dual basis we get from equation (22) the expression $\alpha^1 \wedge \cdots \wedge \alpha^r \wedge \cdots \wedge \alpha^n(f_1, \ldots, f_n) = 1$. Hence $\alpha^1 \wedge \cdots \wedge \alpha^r \neq 0$, $\square$

With the general structure complete we turn our attention to the specific case of forms on manifolds and its (co)tangent space. First we define the form bundle in an analog way to the previous bundles as the disjoint union

$$\wedge^k TM^* = \bigcup_{p \in M} \wedge^k T_pM^*$$

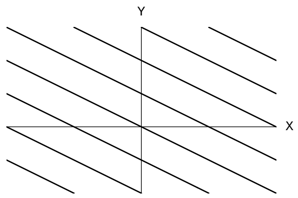


FIGURE 1. The 1-form $dx + 2dy$ in $\mathbb{R}^2$

for a manifold M. A field of k-forms on M is thus a section $\alpha : M \to \wedge^k TM^*$. If $dx^1(p), \dots, dx^n(p)$ is a basis for T_pM then

$$dx^{i_1}(p) \wedge \cdots \wedge dx^{i_k}(p) \qquad i_1 < \cdots < i_k,$$

is a basis for $\wedge^k T_pM^*$. So on U, every k-form field α can be written

$$\alpha = \alpha_{i_1 \dots i_k}\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

To make things simpler we use I to denote $(i_1, \cdots, i_k)$, where $i_1 < \cdots < i_k$, and use it to write $\omega_I = \omega_{i_1} \wedge \cdots \wedge \omega_{i_k}$. Thus we can write k-form fields as

$$\alpha = \alpha_I dx^I,$$

which we call a field of differential forms of order k, if the coefficients α_I are differentiable. We will actually restrict ourselves to differential forms from now on, even though we usually just say “forms”.

EXAMPLE 3.5. In $\mathbb{R}^3$ we have the following types of forms (where a_i are real functions in $\mathbb{R}^3$).

0-forms: functions in $\mathbb{R}^3$
1-forms: $a_1 dx^1 + a_2 dx^2 + a_3 dx^3$
2-forms: $a_{12} dx^1 \wedge dx^2 + a_{13} dx^1 \wedge dx^3 + a_{23} dx^2 \wedge dx^3$
3-forms: $a_{123} dx^1 \wedge dx^2 \wedge dx^3$

EXAMPLE 3.6. Let $\alpha = x_1 dx^1 + x_2 dx^2 + x_3 dx^3 \in (\mathbb{R}^3)^*$ be a 1-form in $\mathbb{R}^3$ and

$$\beta = x_1 dx^1 \wedge dx^2 + dx^1 \wedge dx^3 \in \wedge^2 (\mathbb{R}^3)^*$$

a 2-form in $\mathbb{R}^3$. Then, since $dx^i \wedge dx^i = 0$ and

$$dx^i \wedge dx^j = -dx^j \wedge dx^i, \quad i \neq j,$$

we get

$$\begin{aligned}\alpha \wedge \beta &= x_2 dx^2 \wedge dx^1 \wedge dx^3 + x_3 x_1 dx^3 \wedge dx^1 \wedge dx^2 \\ &= (x_1 x_3 - x_2) dx^1 \wedge dx^2 \wedge dx^3.\end{aligned}$$

2. Exterior Differentiation

From a 0-form (a regular function f) we can get an 1-form df which in a coordinate system (x, U) is given by

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

This we can generalize further to k-forms.

Theorem 3.7. *There is a unique operator, called the* ***exterior derivative***

$$d : \wedge^k M \to \wedge^{k+1} M,$$

such that

(1) $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$,
(2) $d(d\omega) = d^2 = 0$,
(3) $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2)$,

which given a k-form

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

we can write as

$$d\omega = d\omega_I \wedge dx^I = \frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^I.$$

Theorem 3.8. *Let α be a differential form, then the pull-back and the exterior derivative commutes such that we have*

$$f^*(d\alpha) = d(f^*\alpha)$$

Example 3.9 (Electromagnetism in Minkowski space). The electromagnetic field in vacuum, described by Maxwell's equations

$$\begin{array}{ll} \nabla \cdot \mathbf{E} = \rho & \nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j}, \\ \nabla \cdot \mathbf{B} = 0 & \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \end{array}$$

can be written in terms of the two potentials V and $\mathbf{A}$ using the equations

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

In Minkowski space with $c = 1$ and using a metric with the signature $(+, -, -, -)$, we can combine this to form a covariant 4-vector $A_i = (V, -\mathbf{A})$. That is, the electromagnetic field is described by the covector field

$$A = A_i(x) dx^i. \tag{23}$$

One can show that this potential can be used to describe the field strength in a compact way by forming the **electromagnetic field strength tensor**

$$F_{ij} = \partial_i A_j - \partial_j A_i.$$

These are the components of a covariant second-order tensor field that is clearly antisymmetric. Thus this field is actually an field of 2-forms

$$F = (\partial_i A_j - \partial_j A_i) dx^i \wedge dx^j.$$

From classical physics we know that the field strength is the derivative of the potential, e.g

$$\mathbf{E} = -\nabla V, \qquad \mathbf{F} = -\nabla U.$$

This is also the case with the electromagnetic field strength tensor F, as it is actually the exterior derivative of the covector field A [**Fra04**]

$$F = dA.$$

Although it would take the full machinery of integrating differential forms on manifolds, formulating the entire Maxwellian dynamics is a beautiful way to illustrate the power and elegance of the exterior algebra. For a deepened discussion on this, [**Fra04**] is an excellent resource.

EXAMPLE 3.10 (Gauge transformations in electromagnetism). As mentioned earlier the first example of a Gauge theory is Maxwell's electromagnetism. This can be illustrated by showing that we can transform the 4-vector potential A_i by a so-called *Gauge transformation*

$$A_i \to A_i - \partial_i \Lambda,$$

which leaves the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ unchanged. The reason for this will be investigated more deeply in the final chapter.

CHAPTER 4

Lie Theory

Originally Lie groups were introduced in 1870 by S. Lie to study symmetries of differential equations. His approach to groups was to study only elements close to the identity of the group and then use generators of the group, hence restricting the important part of the group to a neighborhood of the identity. The use of generators meant that the group needed to be equipped with differentiable maps adding some extra structure to the group. The area of Lie groups was further broadened by E. Cartan who contributed to the classifications of Lie algebras and applications of Lie groups in theoretical physics and differential geometry. Further developments over the years have shown that Lie groups are also a useful tool to study symmetries of structures in many other branches of mathematics and physics.

In order to have a Lie group we start with a manifold with group properties. The manifold G then needs to have the usual properties of a group, meaning that the manifold G must be equipped with a associative group operation $G \times G \to G$, an identity element and for all elements in the group there must exist an inverse. Furthermore for the group to be a Lie group we require the group operation and the inverse map to be smooth. We now state the formal definition of a Lie group.

Definition 4.1. *A* ***Lie group*** *is a group G which is also a manifold with a C^∞ structure such that*

$$\begin{array}{rcl}(x,y) & \mapsto & xy \\ x & \mapsto & x^{-1}\end{array}$$

are C^∞ functions

Lie groups can be classified according to their properties. Some examples are *simple*, *semi-simple*, *nilpotent* and *abelian* Lie groups. It is also of importance if the Lie group is classified as *connected* and/or *compact*.

1. Examples of Lie Groups

One trivial example of a Lie group is the Euclidean space $\mathbb{R}^n$ with the normal vector addition as the group operation. Obviously this operation and the inverse map both are C^∞ maps. Another important example is the following matrices.

Definition 4.2. *The* ***general linear group*** *of degree n over the field $\mathbb{R}$ is the set*

$$GL(n, \mathbb{R}) = \{A \in M(n, \mathbb{R}) \mid \det(A) \neq 0\},$$

where $M(n, \mathbb{R})$ denotes the set of all real $n \times n$ matrices.

Equipped with ordinary matrix multiplication it is not difficult to show that $GL(n, \mathbb{R})$ is a Lie group. In $GL(n, \mathbb{R})$ the inverse is always defined since the determinant is nonzero. Moreover the determinant of $A \in GL(n, \mathbb{R})$ is a continous function and obviously $\det(A) \neq 0$ does not change that it is an open set. Then it is clear that $GL(n, \mathbb{R})$ is a open subset of $\mathbb{R}^{n^2}$ and since the inverse is a function of the determinant it is clear that the inverse is smooth. Hence $GL(n, \mathbb{R})$ is a manifold with dimension n^2, However $GL(n, \mathbb{R})$ is not a *connected* Lie group. This can be seen by noting that for every $A \in GL(n, \mathbb{R})$ we have $\det(A) \neq 0$, so the group consists of two separate parts, one with $\det(A) > 0$ and one with $\det(A) < 0$. Another important class of matrices are the orthogonal group of matrices. We shall here show that this is a subgroup of $GL(n, \mathbb{R})$.

Definition 4.3. *The* ***orthogonal group*** *over the field $\mathbb{R}$ is the subgroup of $GL(n, \mathbb{R})$:*

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

To show that $O(n, \mathbb{R})$ really is a subgroup of $GL(n, \mathbb{R})$ we need to show that the set contains the group identity element, is closed under the group operation and that the set is closed under the inverse operation. Associativity is inherited from $GL(n, \mathbb{R})$.

(1) The identity element is in $O(n, \mathbb{R})$ since $II^T = I$
(2) We show that $O(n, \mathbb{R})$ is closed under the group operation by taking two elements in $O(n, \mathbb{R})$, A_1 and A_2, then A_1A_2 is in $O(n, \mathbb{R})$ since
$$(A_1A_2)^T(A_1A_2) = A_2^T A_1^T A_1 A_2 = A_2^T A_2 = I.$$
(3) We want to show that if A is in $O(n, \mathbb{R})$ so is A^{-1}. Since $A^{-1} = A^T$, we need to show that A^T is in $O(n, \mathbb{R})$. But this must be the case since
$$A^T(A^T)^T = A^T A = I$$

Furthermore it is possible to show that $O(n, \mathbb{R})$ is a closed and bounded subset of $M(n, \mathbb{R})$. Since $M(n, \mathbb{R})$ is homeomorphic to $\mathbb{R}^{n^2}$ we can note that $O(n, \mathbb{R})$ is in fact compact which we shall see is of importance later. From $O(n, \mathbb{R})$ we can form yet another subgroup of matrices by noting that if $A \in O(n, \mathbb{R})$ then $(\det A)^2 = 1$ since $AA^T = I$. So there are two disjoint sets one with $\det A = -1$ and one set with $\det A = 1$. The set with $\det A = -1$ can not form a group since it does not contain the identity element but the set of matrices with $\det A = 1$ forms a subgroup of $O(n, \mathbb{R})$. This subgroup is called the special orthogonal group $SO(n, \mathbb{R})$. [**Fra04**]

Definition 4.4. *The* ***special orthogonal group*** *over the field* $\mathbb{R}$ *is the group*

$$SO(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) \mid \det A = 1\}.$$

So far all groups has been over the field $\mathbb{R}$, but if we replace $\mathbb{R}$ with the field $\mathbb{C}$ in the definitions of the orthogonal and special orthogonal group we will get similar groups of the same form. These are the unitary group $U(n, \mathbb{C})$ and the special unitary group $SU(n, \mathbb{C})$. Alternatively we can use $O(n)$, $SO(n)$, $U(n)$ and $SU(n)$ since there is no risk of confusion over which field the groups are.

Definition 4.5. *The* ***unitary group*** $U(n)$ *is the group*

$$U(n) = \{A \in GL(n, \mathbb{C}) \mid A^\dagger = A^{-1}\}.$$

where $A^\dagger = \bar{A}^T$ *is the (hermitian) adjoint.*

Note that $U(1)$ is the group of complex numbers $z = e^{i\theta}$, that is, we can think of it as the group of all rotations in the plane. $U(1)$ is also the only abelian unitary group.

Definition 4.6. *The* ***special unitary group*** $SU(n)$ *is the group*

$$SU(n) = \{A \in U(n) \mid \det A = 1\}.$$

2. Invariant Vector Fields

Lie groups are equipped with two diffeomorphisms. These are called the left translations

$$L_g : G \to G \qquad L_g(h) = gh,$$

and the right translations

$$R_g : G \to G \qquad R_g(h) = hg.$$

By use of these two translations we can define two vector fields on the manifold as we can translate vectors to any point. So for a tangent vector at the identity e of G, we left or right translate the tangent vector to any point on G by the differentials (1.15) with

$$X_g = L_{g*} X_e$$

and

$$X_g = R_{g*} X_e.$$

We can now pay special attention to the vector fields which does not change under this transformation. We call a vector field X **left-invariant** on G if it is invariant under the left translation

$$L_{g*} X_h = X_{gh} \qquad \forall g, h \in G,$$

and of course similar for a right-invariant vector field.

3. One Parameter Subgroups

To each vector field X in $\mathbb{R}^n$, it is possible to associate a time independent flow ϕ_i having X as its vector field [**Fra04**]. As the flow is time-independent, it implies

$$\phi_s(\phi_t(p)) = \phi_t(\phi_s(p))$$

as well as

$$\phi_{-t}(\phi_t(p)) = p,$$

hence

$$\phi_{-t} = \phi_t^{-1}.$$

This makes it possible to define a 1-parameter group of maps. The integral curves of the system can then be established as $\phi_i(p)$, where $\phi_i(p)$ simply means that we shall move along the integral curve through the point p. With this in mind we can define the concept of a 1-parameter subgroup.

DEFINITION 4.7. *A* ***1-parameter subgroup*** *of the group G, is a differentiable group homomorphism*

$$\phi : \mathbb{R} \to G$$
$$t \to \phi(t) \in G$$

of the additive group of the reals into the group G. That is, we have

$$\phi(s+t) = \phi(s)\phi(t) = \phi(t)\phi(s).$$

For a 1-parameter subgroup of a matrix group G, we thus have $\phi(t+s) = \phi(t)\phi(s)$ with the normal matrix multiplication as the group operation. Now differentiating and evaluating at s=0 gives

$$\phi'(t) = \phi(t)\phi'(0). \tag{24}$$

Solving this equation we get

$$\phi(t) = \phi(0)e^{t\phi'(0)},$$

where the exponential function for matrices is defined as

$$\exp A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

If we now generalize equation (24) to the case where it is not a matrix group we get

$$\phi'(t) = L_{\phi(t)*}\phi'(0),$$

that is, the tangent vector X to the 1-parameter subgroup is left translated along the subgroup. So the 1-parameter subgroup of G with tangent vector X_e at e, is the integral curve through e of the vector field X on G after left-translation of X_e over all of G. The vector X_e is called the **infinitesimal generator** of the 1-parameter subgroup. We are now in a position to define the exponential map.

Definition 4.8. *The* ***exponential map***

$$\exp : T_eG \to G,$$

is the map such that when given $X \in T_eG$ with the diffeomorphism $\phi : \mathbb{R} \to G$ satisfying

$$\frac{d\phi}{dt}(0) = X,$$

then

$$\exp X = \phi(1).$$

That is, the exponential map is a local diffeomorphism from the tangent space T_eG at the identity element to G. Thus close to the identity, the exponential map and its inverse are both smooth functions. For any Lie group G, we shall also denote the 1-parameter subgroup whose generator at the identity is X, by

$$g(t) = \exp tX.$$

For all matrices we have $\det e^A = e^{\operatorname{tr} A}$, where $\operatorname{tr} A$ denotes the trace of A. This relation will be usefull a little later when we are dealing with the exponential map in an example.

4. Lie Algebra

The tangent space T_eG of G at the identity plays an important part, as we shall soon see. First we consider two left-invariant vector fields X_1 and X_2 on G, for which we can use (1.19) to show

$$L_{g*}[X_1, X_2] = [X_1, X_2], \tag{25}$$

that is, the Lie bracket of X_1 and X_2 are again left-invariant. This leads us to the definition of a Lie algebra for a Lie group.

Definition 4.9. *The vector space of all left-invariant vector fields on a Lie group G with the Lie bracket as composition, is called the* ***Lie algebra*** *of the group G and denoted $\mathfrak{g}$.*

Thus we can think of the Lie algebra as the tangent vector space T_eG of the Lie group G, and then the exponential map is obviously the map

$$\exp : \mathfrak{g} \to G.$$

Example 4.10 (The Lie algebra of $\mathbb{R}^n$). From the previous discussion we can conclude that the Lie algebra for the Euclidean space $\mathbb{R}^n$, is the set of all left-invariant vector fields. Since for $\mathbb{R}^n$ the tangent space at the identity is $\mathbb{R}^n$, and this obviously also is the set of left-translated tangent spaces, we can conclude that the Lie algebra of $\mathbb{R}^n$ is $\mathbb{R}^n$.

EXAMPLE 4.11 (The Lie algebra of $GL(n, \mathbb{R})$). For the general linear group $GL(n, \mathbb{R})$ we can say the following. First note that for every $A \in M(n, \mathbb{R})$ we have

$$\det e^A = e^{\operatorname{tr} A} > 0,$$

so it is clear that

$$\exp : M(n, \mathbb{R}) \to GL(n, \mathbb{R}).$$

Since $\exp_*(X) = X$ for a tangent vector X we can thus say that the exponential map is left-invariant. From dim $M(n, \mathbb{R}) = n^2 =$ dim $GL(n, \mathbb{R})$ we can see that the Lie algebra of $GL(n, \mathbb{R})$ is

$$\mathfrak{gl}(n, \mathbb{R}) = M(n, \mathbb{R}),$$

with the commutator $AB - BA$ as Lie bracket.

EXAMPLE 4.12 (The Lie algebra of $U(n)$). Since $U(n)$ is the group of matrices such that $A^\dagger = A^{-1}$, we note that if A is skew hermitian,

$$A^\dagger = -A,$$

then $\exp A \in U(n)$ and thus the Lie algebra of $U(n)$ is the vector space of all skew hermitian matrices denoted $\mathfrak{u}(n)$, with the commutator as Lie bracket.

DEFINITION 4.13. *A Lie algebra is* ***simple*** *if its only ideal are* $\{0\}$ *and the whole Lie algebra itself.*

DEFINITION 4.14. *A* ***semi-simple Lie group*** *is a Lie group with a* ***semi-simple Lie algebra****. The Lie algebra is a semi-simple Lie algebra if it is a direct sum of simple Lie algebras.*

5. Group action and Representations

Representation theory may be defined as the study of the ways in which a given group may act on vector spaces. There is, however, an interesting more general theory of representations in category theory which contain representations on vector spaces as a special case. Here, though, we will only consider representations on vector spaces, but an easily accessible account of the more general theory may be found in [**Ger85**].
Now, the aim of this section is not to give a detailed account of the theory of group actions and representations but rather define and exemplify the components needed in subsequent chapters. With this in mind we now proceed to find out, in particular, what a representation is and we will learn that it is a special kind of group action so we begin by defining this notion.

DEFINITION 4.15. *A* ***left action*** *of a group G on a set S is a map*

$$\pi : G \times S \to S$$

such that

$$\pi(id_G, s) = s$$

and

$$\pi(g, \pi(\tilde{g}, s)) = \pi(g\tilde{g}, s)$$

for $g, \tilde{g} \in G$ *and* $s \in S$*, where* id_G *denotes the identity element.*

If we denote $\pi(g, s)$ by gs the conditions in the definition become

$$g(\tilde{g}s) = (g\tilde{g})s$$

and

$$id_G s = s.$$

It is immediate from the definition that each π have an inverse $\pi(g)^{-1}$ and that the map $g \mapsto \pi(g, s)$ is a group homomorphism $G \to GL(S)$. Furthermore, any group homomorphism $G \to GL(S)$, reciprocally, defines a left action of G on S. We illustrate these notions with an example that takes us closer to what we are primarily interested in, that is, representations on vector spaces.

EXAMPLE 4.16. Let V be a vector space over a field $\mathbb{F}$, e.g. $\mathbb{R}$ or $\mathbb{C}$, and $G = GL(V)$. For $g \in G$ and $v \in V$ the map

$$(g, v) \mapsto gv$$

then defines an action of G on V. Usually, in this report, we have that $V = \mathbb{F}^n$ and then take G to be the group of all invertible $n \times n$-matrices over the field $\mathbb{F}$ so that, for $M \in G$ and $x \in \mathbb{F}^n$, an action of G on $\mathbb{F}^n$ is defined by

$$(M, x) \mapsto Mx.$$

Generally when we refer to an action we mean a left action, however, in the subsequent chapter on bundles we will, in connection to the so called principal G-bundle, find it useful to have a right action as well. A **right action** of G on a set S is defined as a map

$$\tilde{\pi} : S \times G \to S$$

satisfying

$$\tilde{\pi}(s, id_G) = s$$

$$\tilde{\pi}(\tilde{\pi}(s, g), \tilde{g}) = \tilde{\pi}(s, g\tilde{g})$$

for $g, \tilde{g} \in G$ and $s \in S$. Similarly as with the left action we will write $\tilde{\pi}(s, g)$ as sg. For both the left- and right action we have a natural action of a group G on itself given by

$$\pi(g, \tilde{g}) = g\tilde{g}$$

and

$$\tilde{\pi}(\tilde{g}, g) = \tilde{g}g,$$

respectively. If we have a left- and a right action on a set S by groups G and H simultaneously then if

$$(gs)h = g(sh), \quad \forall s \in S,\ g \in G,\ h \in H$$

the two actions are said to **commute**. A particular type of action, referred to as effective, will be used in subsequent chapters and we now examine this notion. If we have a left action

$$\pi : G \times S \to S$$

then associated to this is a subgroup

$$G^0 = \{g \in G \mid \forall s \in S,\ gs = s\}$$

of G, which is normal in G since it is the kernel of the homomorphism

$$G \to GL(S).$$

Now, an action is called **effective** if $G^0 = id_G$. Specifically for Lie groups we have seen earlier in this chapter that the tangent space T_gG at any point other then the identity can be canonically identified with the Lie algebra $\mathfrak{g}$ of a Lie group G in two ways. Considering right action and the fact that any $v \in T_gG$ can be written as

$$v = Ag,$$

according to section 2, with $A \in \mathfrak{g}$ the so called **adjoint map**

$$Ad_g(A) = gAg^{-1}$$

which maps $\mathfrak{g}$ into itself is obtained. We now know enough about group action to define representations of groups on vector spaces as follows.

Definition 4.17. *A* ***representation*** *of a group G on a vector space V over a field $\mathbb{F}$ is a left action of G on V such that $\forall v \in V$, $g \in G$ the map*

$$v \to gv$$

is $\mathbb{F}$-linear.

Usually one writes $gv = R(g)v$ with $R(g) \in GL(V)$ and from the definition of action it follows that $R(id_G) = id_V$, and for $g, \tilde{g} \in G$ we have that

$$R(g\tilde{g}) = R(g)R(\tilde{g}),$$

that is, we have a group homomorphism $G \to GL(V)$. Given a set of such maps we also, reciprocally, have a group representation, and if V is finite dimensional of dimension n we may fix a basis and thus represent each $R(g)$ as an $n \times n$-matrix $M(g) \in GL(n, F)$. These matrix representations are therefore simply group homomorphism

$$G \to GL(n, F)$$

and are particularly useful, especially in physics, as the following shows.

Example 4.18 (Matrix representations). The equations of modern physics are often eigenvalue equations of the form $A\psi = a\psi$, where A is a Hermitian matrix and a the corresponding (real) eigenvalues. The matrix representations arise naturally from symmetries of these equations as follows. Assume that the eigenvalue equation above is invariant under a group G of transformations T, e.g. coordinate rotations. Then we have

$$A\psi = a\psi,$$

which by assumption is invariant under the transformations T in G so that

$$A = TAT^{-1}.$$

Applying a transformation T to a solution ψ so that $\psi \to T\psi$ gives together with the two equations directly above

$$A\psi = a\psi$$
$$\Leftrightarrow \quad aT\psi = Ta\psi = TA\psi = TAT^{-1}T\psi = AT\psi.$$

Assuming the vector space V of solutions has finite dimension n and a basis $B = \{\psi_1, ..., \psi_n\}$ we may expand each member of V in terms of B as

$$T\psi_i = \sum_k t_{ik}\psi_k.$$

So to all transformations T in G we may associate a matrix (t_{ik}) and the map $T \to (r_{ik})$ is the matrix representation of G.

CHAPTER 5

Bundles and Gauge Theory

A key element in defining gauge theory is the notion of bundles, or fibre bundles. Bundles as part of mathematics was founded and recognized in the 1930's with, in particular the works of H. Whitney, H. Hopf and E. Steifel. A general interest in bundles followed fairly soon after their conception because of the theory's ability to supply great applications of topology to other fields. Their use in physics, or more specifically field theory, became a necessity when physicists tried to describe fundamental particle interactions and realized that the representation of physical fields as functions on space-time was obsolete and that topologically more complex objects was needed. A suitable choice turned out to be sections, or cross-sections, of bundles. However, the use of bundles in physics and mathematics is by no means restricted to the above mentioned, they have in fact turned out to be very useful in many areas, for instance in dynamical systems.

In this chapter the material from the theory of bundles necessary to present the elements of gauge theory is introduced and examples concerning gauge theories will be given, specifically the ones used to describe interactions in modern particle physics. The exposition will proceed from the general case, i.e. the fibre bundle, to the special case of a vector bundle, and in this general scheme the relevant bundle concepts, e.g. section and group of a bundle, will be defined, explained, exemplified, and related to previous theory. After this familiarization with the relevant types of bundles and concepts, notions such as connections and curvature is treated and, generally, considered in relation to one specific type of bundle where its significance, with respect to gauge theory, is the greatest.

The dependence of this chapter on the previous ones is as follows. The theory of Lie groups enters almost at once since the structure group of each of the bundles is generally of this type. The examples given in the bundle part has been chosen as to make full use of the previously introduced concepts such as that of tangent- and cotangent bundles. The sections concerning connections, curvature and covariant derivatives are basically applications of the theory of differential forms and tensors and may be seen as examples of the notions and techniques introduced in these chapters.

1. Fibre Bundles

In mathematics some objects are fortunate enough to have names that in some sense, without a great deal of imagination, conveys their structure through the meaning of the words with respect to the real world. This is the case with bundles and they are so called because they consist of objects, the fibres, which are bundled together to form the bundle, like a bundle of hay. The strength of the bundle concept and its importance lies in the fact that it generalizes the notion of product space and, in turn the notion of function, which allows one to treat situations where the range of the function is only well defined locally. These facts, and many more, will become evident in what follows and we now begin to introduce the mathematical object called fibre bundle.

In order to indicate the notions mathematical nature and origin, and for clarification we first point out that; generally the three spaces $\mathcal{T}$, $\mathcal{B}$ and $\mathcal{F}$, in the definition of fibre bundle to be given shortly, are only required to be topological spaces and the defined mappings homeomorphisms. However, for our purposes it is more efficient to consider them to be differentiable manifolds, objects which we have familiarized ourselves with earlier as opposed to topological spaces. Furthermore, what we call a fibre bundle will sometimes be referred to as a coordinate bundle because a fibre bundle in the proper sense should not be dependent on a particular covering of the base space, and is therefore defined as an equivalence class of coordinate bundles. In this sense one can think of a fibre bundle as a maximal coordinate bundle but this will, possibly without gain, obscure the path we are currently on, i.e. to gauge theory. Because the usual general definitions does not supply enough for us to work with we will also equip the bundle with a structure group $\mathcal{G}$ and also a class of transition functions. Thus, we define a fibre bundle as follows.

Definition 5.1. *A* ***fibre bundle*** $\mathfrak{F}$ *consist of and are subject to the following:*

- $\mathfrak{F}1$ *Three differentiable manifolds: The total- (or bundle-) space $\mathcal{T}$, the base space $\mathcal{B}$, and the (typical) fibre $\mathcal{F}$.*
- $\mathfrak{F}2$ *A surjective map, the projection,*

$$p : \mathcal{T} \to \mathcal{B}$$

- $\mathfrak{F}3$ *A covering C of $\mathcal{B}$ by a family of open sets $\{U_i\}$ and to each U_i a corresponding diffeomorphism, the local trivialization,*

$$\phi_i : U_i \times \mathcal{F} \to p^{-1}(U_i)$$

with the property that

$$p \circ \phi_i(u, f) = u, \quad \forall u \in U,\ f \in \mathcal{F}.$$

$\mathfrak{F}4$ *A group $\mathcal{G}$, the structure group (of the bundle), which acts effectively on the fibre $\mathcal{F}$ from the left.*

$\mathfrak{F}5$ *Define the map $\phi_{i,u} : \mathcal{F} \to p^{-1}(u)$ by*

$$\phi_{i,u}(f) = \phi_i(u, f)$$

so that $\forall i, j$ and $\forall u \in U_i \cap U_j$ the diffeomorphism

$$\phi_{j,u}^{-1} \circ \phi_{i,u} : \mathcal{F} \to \mathcal{F}$$

may be identified with a unique element of $\mathcal{G}$ and the map

$$g_{ji} : U_i \cap U_j \to \mathcal{G}$$

defined by

$$g_{ji}(u) = \phi_{j,u}^{-1} \circ \phi_{i,u}$$

is differentiable.

As for the all important question of equivalence, we say that two fibre bundles $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are **equivalent** if $\mathcal{T}_1 = \mathcal{T}_2$, $\mathcal{B}_1 = \mathcal{B}_2$, $\mathcal{F}_1 = \mathcal{F}_2$, $p_1 = p_2$, $\mathcal{G}_1 = \mathcal{G}_2$ and if $\{U_{1_i}\} \cup \{U_{2_j}\}$ is another covering of $\mathcal{B}$ and $\{\phi_{1_i}\} \cup \{\phi_{2_j}\}$ is another corresponding trivialization.

The definition above is the most general bundle definition that will appear in this report and its main purpose is to provide a unifying frame for the special types of bundles we will be examining. These bundles will essentially only differ from the characterization above by an additional structure on the fibre $\mathcal{F}$, i.e. the fibre may in addition of being a differentiable manifold, have e.g. a vector space- or group structure.

Furthermore, there are a few specifics worth noting concerning the definition. First of all it follows from $\mathfrak{F}3$ that $\forall u \in U$ the fibre over the point u,

$$F_u = p^{-1}(u),$$

is diffeomorphic to the (typical) fibre $\mathcal{F}$ of the bundle, since $\forall u \in U$ the assignment

$$f \in \mathcal{F} \to \phi_j(u, f) \in p^{-1}(u)$$

establishes a diffeomorphism. The defining criteria $\mathfrak{F}4$ and $\mathfrak{F}5$ are those who incorporate the group and its structure, e.g. topology, into the bundle structure. In the sequel the structure groups will generally be compact semi-simple Lie groups.

The maps g_{ij} defined in $\mathfrak{F}5$ are called **transition functions** and they are the entities that allows one to paste the local pieces together to form a bundle. Directly from the definition we can then derive the following relations, sometimes called **consistency conditions**, since if they are not satisfied the bundle cannot be assembled in a consistent way.

$$g_{kj}(u)g_{ji}(u) = g_{ki}(u)$$

$$g_{ii}(u) = id_{\mathcal{G}}$$
$$g_{ij}(u) = g_{ji}^{-1}(u)$$

If the only necessary transition function is the identity $id_{\mathcal{G}}$ then we have that the total space

$$\mathcal{T} = \mathcal{B} \times \mathcal{F}$$

and in this case the bundle is called **trivial**. Examining the consistency conditions in this case, where only one patch U_i is necessary to cover the base space so that $i = j = k$, it becomes evident that the group of the bundle consist of only the identity.

We now know what a fibre bundle is and have an idea about how it generalizes the concept of product space and it would be interesting to find out how this setting may be used to generalize the notion of differentiable function. This is shown in the next definition where we define the essential notion of a section, or cross-section, of a bundle.

DEFINITION 5.2. *A* ***section*** *on a fibre bundle $\mathfrak{F}$ is a differentiable map $s : \mathcal{B} \to \mathcal{T}$ such that $p \circ s = id_{\mathcal{B}}$.*

In general we do not have sections defined on the whole of $\mathcal{B}$ but rater on some $U \subset \mathcal{B}$ and we call these

$$s_U : U \to \mathcal{T}$$

local sections. Now, in a bundle $\mathfrak{F}$ we have for $u \in U \subset \mathcal{B}$ that $s(u) \in \mathcal{F}_u$ and we label the class of all sections defined on $U \subset \mathcal{B}$ by $\Gamma(U, \mathcal{F})$. If we do have sections defined on all of $\mathcal{B}$ we call them **global sections** and denote the class of them by $\Gamma(\mathcal{B}, \mathcal{F})$.

To familiarize ourselves with the concepts above and add some context we examine the trivial bundle and then we revisit the slightly more interesting case of the tangent bundle in the following two examples.

EXAMPLE 5.3 (The Trivial Bundle). This is the simplest bundle there is and it is the closest relative to the object that bundles are supposed to generalize, i.e. product spaces. However, the trivial bundle is just a relative, it is not a product space as some, especially in the physics oriented, literature make it out to be. It is indeed true that the total space $\mathcal{T}$ in this case is of the form

$$\mathcal{T} = \mathcal{B} \times \mathcal{F},$$

however, by construction of the bundle the natural projection

$$\pi : X \times Y \to Y,$$

in the case of topological or manifold product spaces $X \times Y$, defined by

$$\pi(x, y) = y, \quad \forall x \in X,\ y \in Y$$

is lost. Now, let us go through each point of definition 5.1 for the trivial bundle and see how it works.

$\mathfrak{F}1$ The three manifolds are $\mathcal{B}$, $\mathcal{F}$ and

$$\mathcal{T} = \mathcal{B} \times \mathcal{F}.$$

$\mathfrak{F}2$ The projection $p : \mathcal{B} \times \mathcal{F} \to \mathcal{B}$ is the natural one defined by

$$p(b, f) = b.$$

$\mathfrak{F}3$ Covering $\mathcal{B}$ by $U = \mathcal{B}$ we may choose the global trivialization $\phi = id$ which clearly is a diffeomorphism and it make

$$p \circ \phi_i(u, f) = u, \quad \forall u \in U, f \in \mathcal{F}$$

trivially satisfied by definition of the projection.

$\mathfrak{F}4 - 5$ $\mathcal{G}$ will consist of only the identity since the elements of this group are the transitions maps, denoted g in definition 5.1, between the fibres and these are in the case of the trivial bundle identical and hence only a map mapping one fibre onto itself, i.e. the identity, is necessary. Actually, it follows directly from $\mathfrak{F}5$ that if the group of a bundle consist of the identity alone, then the bundle is equivalent to a trivial bundle.

Furthermore, this bundle have one global section

$$s : \mathcal{B} \to \mathcal{T} = \mathcal{B} \times \mathcal{F},$$

and this is of course the inverse, p^{-1}, of the projection p.

The reason for calling the map

$$\phi_i : U_i \times \mathcal{F} \to p^{-1}(U_i),$$

in the definition of the fibre bundle, a *local trivialization* now seem very sensible since what the map actually does is to make the part of the bundle over each U_i into a trivial bundle.

Example 5.4 (The Tangent Bundle). The base space $\mathcal{B}$ in this case is an n-dimensional differentiable manifold, and the total space $\mathcal{T}$ is the set of all tangent vectors at all points of $\mathcal{B}$. The projection p is the map that sends each tangent vector, in $\mathcal{T}$, to its initial point, in $\mathcal{B}$. Clearly

$$p : \mathcal{T} \to \mathcal{B}$$

is surjective. Now the fibre over the point $b \in \mathcal{B}$, $\mathcal{F}_b$, is the tangent plane at b, and this is a linear space. We saw earlier that each fibre over a point in the base space was diffeomorphic to the (typical) fibre, and we have linear spaces so a vector space isomorphism $\mathcal{F}_b \to \mathcal{F}$ can be constructed, thus in this case the structure group is

$$\mathcal{G} = GL(n, \mathbb{R}).$$

A section of the tangent bundle is a vector field and thus the set of all sections is the set of all vector fields over the base space $\mathcal{B}$.

This example may be somewhat un-pedagogical since the tangent bundle carries additional structure on its fibre, namely a linear structure, and therefore it is not a general fibre bundle but rather a special case called a vector bundle, which we shall examine in a while.

As usual when one embarks on new ground in mathematics two questions are of special interest, namely, what are the objects and what kind of maps (or morphism) do we have between these objects. For example in the first part of this report the objects were manifolds and the mappings homeomorphism, which soon became differentiable manifolds and diffeomorphisms and even later we encountered the pair vector space and (vector space) isomorphism. This type of pairing, object-map, is a general trait of mathematics, which a subdiscipline of mathematics called category theory exploits, so we now ask what the maps between our newly found objects, the bundles, are.
The main property of these maps is, quite naturally, that they do not disturb the structure on the fibre, i.e. they are fibre preserving, as the following definition shows.

Definition 5.5. *Let $\mathfrak{F}_1$ and $\mathfrak{F}_1$ be two fibre bundles with the same fibre $\mathcal{F}$ and structure group $\mathcal{G}$. A* ***bundle map***

$$\beta : \mathfrak{F}_1 \to \mathfrak{F}_2$$

is then defined as a differentiable map $\beta : \mathcal{T}_1 \to \mathcal{T}_2$ subject to the following:

(1) *$\forall b_1 \in \mathcal{B}_1$ β takes every $\mathcal{F}_{b_1} \in \mathcal{T}_1$ diffeomorphically onto a $\mathcal{F}_{b_2} \in \mathcal{T}_2$. This induces, through the projections in each bundle, a differentiable map $\tilde{\beta} : \mathcal{B}_1 \to \mathcal{B}_2$ such that*

$$p_2 \circ \beta = \tilde{\beta} \circ p$$

(2) *If $u \in U_{1_i} \cap \tilde{\beta}^{-1}(U_{2_j})$ and*

$$\beta_u : \mathfrak{F}_1 \to \mathfrak{F}_2$$

is the map induced by $\beta(\tilde{\beta}(b_1))$ then the map $\tilde{g} : \mathcal{F} \to \mathcal{F}$ defined by

$$\tilde{g}_{ji}(u) = \phi^{-1}_{j,b_2} \circ \beta_x \circ \phi_{i,u} = p_{2_j} \circ \beta_x \circ \phi_{i,u}$$

coincide with the operation of an element in $\mathcal{G}$ and the map

$$\tilde{g}_{ji}(u) : U_{1_i} \cap \tilde{\beta}^{-1}(U_{2_j})$$

obtained in this way is differentiable.

We collect a few facts regarding these bundle maps in the following theorem.

Theorem 5.6. *Properties of bundle maps are*

(1) *the identity map $id : \mathcal{T} \to \mathcal{T}$ is a bundle map $\mathfrak{F} \longrightarrow \mathfrak{F}$*
(2) *the composition of two bundle maps $\mathfrak{F}_1 \longrightarrow \mathfrak{F}_2 \longrightarrow \mathfrak{F}_3$ is also a bundle map $\mathfrak{F}_1 \longrightarrow \mathfrak{F}_3$*

Proofs are relatively easy obtained from the definition and are not given in detail here, however, they can be found in [**Ste51**]

The $\tilde{g}_{ji}$:s above are obviously related to the transition functions introduced earlier and as can be seen in the definition they are mapping transformations analogous to the role of coordinate transformations that the transformation functions have. These two types of transformation functions are also related by the following relations which the $\tilde{g}_{ji}$:s are required to satisfy

$$\tilde{g}_{1_{kj}}(u)g_{1_{ji}}(u) = \tilde{g}_{1_{ki}}(u),$$

for $u \in U_{1_i} \cap U_{1_j} \cap \tilde{\beta}^{-1}(U_{2_k})$, and

$$g_{2_{lk}}(\tilde{\beta}(u))\tilde{g}_{1_{kj}} = \tilde{g}_{1_{lj}},$$

for $u \in U_{1_i} \cap \tilde{\beta}^{-1}(U_{2_k} \cap U_{2_l})$. Similarly to the related consistency conditions for transition functions these relations follow directly from the definition.

This concludes our treatment of the general fibre bundle and we now examine some important special cases.

2. Vector Bundles

A vector bundle is a fibre bundle $\mathfrak{F}$ whose fibre $\mathcal{F}$ is equipped with a linear structure, i.e. the vector bundle fibre $\mathcal{F}$ is a differentiable manifold and a vector space, and thus the fibre over every point b in the base space $\mathcal{B}$ is of this type also. To get a more intuitive idea of the vector bundle one may think of it as a family of vector spaces parameterized by points in the base space $\mathcal{B}$, through the projection. Furthermore, in our exposition the vector spaces will be finite dimensional over the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. For information about the cases where one considers vector spaces over other number fields, e.g. the quaternions and the p-adic numbers, or infinite dimensional spaces in the bundle setting we refer to other sources.
Now we formally define the notion of vector bundle using the definition of fibre bundle with appropriate changes to incorporate the structured fibre.

Definition 5.7. *A* ***vector bundle*** $\mathfrak{V}$ *consist of and are subject to the following:*

- $\mathfrak{V}1$ *Two differentiable manifolds: The total- (or bundle-) space $\mathcal{T}$, the base space $\mathcal{B}$.*
- $\mathfrak{V}2$ *A differentiable manifold equipped with a vector space structure called the (typical) fibre and denoted by $\mathcal{F}$.*
- $\mathfrak{V}3$ *A surjective map, the projection,*

$$p : \mathcal{T} \to \mathcal{B}$$

- $\mathfrak{V}4$ *A covering C of $\mathcal{B}$ by a family of open sets $\{U_i\}$ and to each U_i a corresponding diffeomorphism, the local trivialization,*

$$\phi_i : U_i \times \mathcal{F} \to p^{-1}(U_i)$$

with the property that

$$p \circ \phi_i(u, f) = u, \quad \forall u \in U,\ f \in \mathcal{F}.$$

$\mathfrak{V}5$ *A group $\mathcal{G} \subset GL(n, \mathbb{F})$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ and n is the (vector space) dimension of $\mathcal{F}$, called the structure group (of the vector bundle), which acts linearly on the fibre $\mathcal{F}$ from the left.*

$\mathfrak{V}6$ *Define the map $\phi_{i,u} : \mathcal{F} \to p^{-1}(u)$ by*

$$\phi_{i,u}(f) = \phi_i(u, f)$$

so that $\forall i, j$ and $\forall u \in U_i \cap U_j$ the vector space isomorphism

$$\phi_{j,u}^{-1} \circ \phi_{i,u} : \mathcal{F} \to \mathcal{F}$$

may be identified with a unique element of $\mathcal{G}$ and the map

$$g_{ji} : U_i \cap U_j \to GL(n, \mathbb{F})$$

defined by

$$g_{ji}(u) = g_{j,u}^{-1} \circ g_{i,u}$$

is differentiable.

If the fibre $\mathcal{F}$ is one-dimensional the vector bundle is called a **line bundle**. Note that the transition functions g_{ij} in the vector bundle are elements of $GL(n, \mathbb{F})$ and can thus be thought of as invertible $n \times n$-matrices and we emphasize this by calling the g_{ij}:s transition matrices. Generally, the larger the group is the more complicated it is to have as a group of a bundle and therefore one often have smaller subgroups of $GL(n, \mathbb{F})$ such as $SU(n, \mathbb{F})$, $U(n, \mathbb{F})$, $O(n, \mathbb{F})$ etc. as structure groups instead.

The sections of a vector bundle are of importance in what follows and some particularities that may be difficult to see from just generalizing the definition of section given previously are useful, and therefore we briefly survey the notion here. The basic definition is the same as in the case of fibre bundles.

DEFINITION 5.8. *a **section** on a vector bundle $\mathfrak{V}$ is a differentiable map*

$$s : \mathcal{B} \to \mathcal{T}$$

such that $p \circ s = id_{\mathcal{B}}$.

Now, for two sections s_1 and s_2 of a vector bundle pointwise vector addition and scalar multiplication is defined as one expects by

$$(s_1 + s_2)(b) = s_1(b) + s_2(b), b \in \mathcal{B}$$

$$(fs_1)(p) = f(p)s_1(p).$$

In every vector bundle $\mathfrak{V}$ there is also a global section, the null section s_0, which in any trivialization satisfies

$$\phi^{-1}(s_0(b)) = (b, 0).$$

3. Principal G-Bundles

In the previous section we found out that the essential difference between a vector- and a fibre bundle is that the typical fibre $\mathcal{F}$ of the former is a vector space whereas it in the latter only is a manifold. When we now examine the principal G-bundle we will find similarly that what essentially distinguishes it from the fibre bundle is that its fibre has a group structure, more specifically its typical fibre is equal to the structure group of the bundle and we have the following definition.

Definition 5.9. *A **principal** G**-bundle**, or G**-bundle** for short, $\mathfrak{G}$ consist of and are subject to the following:*

$\mathfrak{G}1$ *Two differentiable manifolds: The total- (or bundle-) space $\mathcal{T}$, the base space $\mathcal{B}$.*

$\mathfrak{G}2$ *A differentiable manifold equipped with a group structure called the (typical) fibre, denoted by $\mathcal{F}$, and this will be identical to the structure group to be defined in $\mathfrak{G}4$, thus we have $\mathcal{F} = \mathcal{G}$.*

$\mathfrak{G}3$ *A surjective map, the projection,*

$$p : \mathcal{T} \to \mathcal{B}$$

$\mathfrak{G}4$ *A covering C of B by a family of open sets $\{U_i\}$ and to each U_i a corresponding diffeomorphism, the local trivialization,*

$$\phi_i : U_i \times \mathcal{F} \to p^{-1}(U_i)$$

with the property that

$$p \circ \phi_i(u, f) = u, \quad \forall u \in U, \ f \in \mathcal{F}.$$

$\mathfrak{G}5$ *A group $\mathcal{G} \subset GL(n, \mathbb{F})$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ and n is the dimension of $\mathcal{F}$, called the structure group (of the principal bundle), which acts linearly on the fibre $\mathcal{F}$, i.e. itself, from the left.*

$\mathfrak{G}6$ *Define the map $\phi_{i,u} : \mathcal{G} \to p^{-1}(u)$ by*

$$\phi_{i,u}(g) = \phi_i(u, g)$$

so that $\forall i, j$ and $\forall u \in U_i \cap U_j$ the isomorphism

$$\phi_{j,u}^{-1} \circ \phi_{i,u} : \mathcal{G} \to \mathcal{G}$$

may be identified with a unique element of $\mathcal{G}$ and the map

$$g_{ji} : U_i \cap U_j \to \mathcal{G}$$

defined by

$$g_{ji}(u) = g_{j,u}^{-1} \circ g_{i,u}$$

is differentiable.

What is remarkable about the G-bundles is that one also have a natural right action of the structure group on the total space of the bundle itself, i.e. the natural right action commutes with the left action, something which is not shared by our other types of bundles.

4. Associated Bundles

In this section we will not get to know a new type of bundle in the sense that the typical fibre has some new type of structure. Instead we associate, in particular, the special types of bundles we have encountered above, i.e. the vector- and principal G- bundles. The general idea underlying what follows is that one begin with one special kind of bundle and then find a way to, from in particular the typical fibre of this bundle, generate a new bundle, and this is called an **associated bundle** of the former. We will mainly consider the case when one starts out with a principal G-bundle and to it associates a vector bundle by means of a representation.
The reason for this associating of bundles is generally that some property or structure of the former bundle naturally induces a desirable property or structure in the associated bundle. In our case the connection on the G-bundle will induce a covariant derivative on the associated vector bundle.

To figure out how we may construct associated bundles it is of interest to us to know what information we necessarily must have in order to reconstruct a fibre bundle $\mathfrak{F}$, if we for now consider the general case. It turns out that if $\mathcal{B}$, $\mathcal{F}$, $\mathcal{G}$, $\{U_i\}$ and $g_{ij}(u)$ are known then the remaining definitional components p, $\mathcal{T}$ and ϕ_i can be determined uniquely and thus the fibre bundle may be reconstructed. To show this we proceed as follows.

Define a space X by

$$X = \bigcup_i U_i \times \mathcal{F}.$$

Then for $(u_1, f_1) \in U_i \times \mathcal{F}$ and $(u_2, f_2) \in U_j \times \mathcal{F}$ define an equivalence relation $\sim$ by

$$(u_1, f_1)\ (u_2, f_2) \quad \Leftrightarrow \quad u_1 = u_2$$

and $g_{ij}(u) f_1 = f_2$. The total space of $\mathfrak{F}$ may then be defined as

$$\mathcal{T} = X/\sim .$$

If we now denote an element of $\mathcal{T}$ by $[(u, f)]$, the projection is given by

$$p([(u, f)]) = u.$$

Finally the local trivialization $\phi_i : U_i \times \mathcal{F} \to p^{-1}(U_i)$ is given by

$$\phi_i((u, f)) = [(u, f)].$$

The above tells us not only if and how we can recover a partially lost fibre bundle, but more importantly how we may construct an associated bundle to a given bundle. Since we are mainly interested in vector bundles associated to principal G-bundles what is done will sometimes seem devoid from the above construct, however, the reader will benefit from keeping it in mind.

Now consider a principal G-bundle $\mathfrak{G}$ with fibre $\mathcal{F} = \mathcal{G}$ and local transition matrices

$$g_{ij} : U_i \cap U_j \to \mathcal{G}.$$

If we have that

$$\rho : \mathcal{G} \to GL(n, \mathbb{C})$$

is some representation of the structure group $\mathcal{G}$, we may define a new vector bundle $\mathfrak{V}_{\mathfrak{G},\rho}$ associated to the principal bundle $\mathfrak{G}$ by the representation ρ using as transition matrices the $\rho(g_{ij})$:s instead of the g_{ij}:s. The associated vector bundle will, in this case, have $\mathcal{F} = \mathbb{C}^n$ and as for the change of transition matrices we define an equivalence relation $\sim$ for $u \in U_i \cap U_j$, $(u, \varphi_{U_i}) \in U_i \times \mathbb{C}^n$ and $(u, \varphi_{U_j}) \in U_j \times \mathbb{C}^n$ by

$$(u, \varphi_{U_i}) \ (u, \varphi_{U_j}) \Leftrightarrow \varphi_{U_i} = \rho(g_{ij}(u))\varphi_{U_j}$$

The direction of association is not really important for the terminology and we may refer to an associated principal G-bundle of a vector bundle, and it is perfectly in order to associate a bundle of a given type to another bundle of the same type, e.g. an associated vector bundle of a vector bundle. This last paragraph and the reasoning above is enlightened in the following example which defines the concept of duality of vector bundles.

Example 5.10 (Dual vector bundles: the tangent- and cotangent bundles). From the previous chapters of this report re-collect that if TM is the tangent bundle over the manifold M, we know that $g_{12} = \frac{\partial u_1}{\partial u_2}$, and if we define the representation

$$\rho^* : GL(n, \mathbb{C}) \to GL(n, \mathbb{C})$$

by

$$\rho^*(h) = (h^{-1})^T = h^*,$$

the associated transition matrices are

$$\rho^*(g_{12}) = [g_{12}^{-1}]^T = g_{12}^*.$$

Identifying

$$t_a^{U_1} = \left[\frac{\partial u_1}{\partial u_2}\right]_a^T b t_b^{U_2} = t_b^{U_2} \left[\frac{\partial u_1^b}{\partial u_2^a}\right]^T$$

shows that $\mathcal{TM}_\rho$ is indeed the cotangent bundle. As indicated above this applies to a general vector bundle $\mathfrak{V}$ and we call its, by the representation ρ^*, associated vector bundle the **dual bundle** of $\mathfrak{V}$ and denote it by $\mathfrak{V}^*$.

5. Connections on Principal G-Bundles

A connection is basically a covariant differentiation operation, in that it specifies how tensors are transported along a curve. In this section we will examine, in particular, connections on G-bundles. The fact that we have G-bundle connections as our primary objective we will not exhaustively examine the notion of connections but rather collect a few, for us, interesting definitions and facts. The reader is referred to [**Fra04**] or [**Nak03**] for a more detailed treatment of bundle connections, in particular their relation to parallel transport etc.

The general idea is to separate the tangent space $T_t\mathcal{T}$ over the total space $\mathcal{T}$ of a G-bundle $\mathfrak{G}$ over each point $t \in \mathcal{T}$ into vertical $V_t\mathcal{T}$ and horizontal $H_t\mathcal{T}$ parts, somewhat analogous to the resolving of an arbitrary vector, e.g. in the plane, into orthogonal parts, usually as multiples of orthonormal base vectors. Generally this is done by projecting the vector onto each vector in the base and this is essentially what we do for the tangent space to $\mathcal{T}$ in terms of differential forms, and we will refer to a connection as a $\mathfrak{g}$-valued 1-form (on $\mathcal{T}$). Before we give a precise definition of connection we need to specify what we mean by vertical- and horizontal parts, i.e. subspaces, of a tangent space to a G-bundle. Throughout all of this it is essential to remember that we are working with G-bundles since we will frequently use the natural right action which they, but not the other bundles, are endowed with.

The vertical part $V_t\mathcal{T}$ is a subspace of $T_t\mathcal{T}$ which is tangent to the fibre $\mathcal{F}_u$ where $u = p(t)$ and is constructed as follows. Choose an $M \in \mathfrak{g}$ so that by the right action

$$R_{\exp(xM)}t = t\,\exp(xM)$$

a curve through t is defined in $\mathcal{T}$ and this curve is confined to $\mathcal{F}_u$ since

$$p(t) = p(t\,\exp(xM)) = u.$$

Let $\phi : \mathcal{T} \to \mathbb{R}$ be an arbitrary smooth function and define a vector $\tilde{M} \in T_t\mathcal{T}$ by

$$\tilde{M}\phi(t) = \frac{d}{dt}\phi(t\,\exp(xM))|_{x=0}.$$

This vector $\tilde{M}$ is thereby tangent to $\mathcal{T}$ at t and thus $\tilde{M} \in V_t\mathcal{T}$. If this is done for all points of $\mathcal{T}$ a vector field $\tilde{M}$ has been constructed and there is a vector space isomorphism

$$i : \mathfrak{g} \to V_t\mathcal{T}$$

defined by

$$i(M) = \tilde{M}.$$

Now as soon as we have a connection in $\mathfrak{G}$ the horizontal part $H_t\mathcal{T}$ is uniquely specified as the complement of $V_t\mathcal{T}$ in $T_t\mathcal{T}$ and we proceed to define the notion of a connection in a G bundle.

Definition 5.11 (The Ehresmann Connection 1-form ω). *A* ***connection 1-form*** $\omega \in \mathfrak{g} \otimes T^*P$ *is a projection of* $T_t\mathcal{T}$ *onto* $V_t\mathcal{T}$ *such that*

$$\omega(\tilde{M}) = M, M \in \mathfrak{g},$$

$$R_g^*\omega = Ad_{g^{-1}}\omega, g \in \mathcal{G},$$

i.e. for $X \in T_t\mathcal{T}$ *we have that*

$$R_g^*\omega_{tg} = \omega_{tg}(R_{g^*}X) = g^{-1}\omega_t(x)g.$$

The horizontal part is then defined by

$$H_t\mathcal{T} = Ker(\omega) = \{X \in T_t\mathcal{T} \mid \omega(X) = 0\}$$

What we really want is to localize this connection and this may be done in different ways depending on if one chooses to use the frame bundle explicitly or not, and we choose not to do this and we go about it as follows. Let $\mathfrak{G}$ be a G-bundle with a covering $\{U_i\}$ of $\mathcal{B}$ and let $\{s_i\}$ be a corresponding set of local sections, i.e. for all i the section s_i is defined on the patch U_i. Now we localize the connection by defining a $\mathfrak{g}$-valued 1-form $\mathcal{A}_i$ through

$$\mathcal{A}_i = s_i^*\omega \in \mathfrak{g} \otimes \Omega^1(U_i).$$

Thus the pull-back $\mathcal{A}_i$ is defined locally, but in general not globally since a non-trivial G-bundle cannot have a global section, as discussed previously. We now have our first gauge theory component and it is the $\mathcal{A}_i$ defined above, which in gauge theory is identified with a **gauge**- or **Yang-Mills potential**. Generally we will, of course, need to have more than one gauge potential $\mathcal{A}_i$ for full coverage since the bundles of relevant gauge theories are rarely trivial, the exception being the $U(1)$ gauge theory for electromagnetic interaction where the bundle is in fact trivial. Now, if we need several potentials these cannot be completely arbitrary, and they must satisfy the following **compatibility condition**

$$\mathcal{A}_j = g_{ij}^{-1}\mathcal{A}_i g_{ij} + g_{ij}^{-1}dg_{ij},$$

where the d denotes the exterior derivative on $\mathfrak{G}$ and the g_{ij}:s are transition functions. This is obtained by using the properties of the connection 1-form and an expression describing how sections act on an element of the tangent space over a point, and a derivation of this and the details of the following may be found in [**Nak03**]. Furthermore, this compatibility condition lead to a relation between two corresponding gauge potentials $\mathcal{A}_1$ and $\mathcal{A}_2$ given by

$$\mathcal{A}_2 = c^{-1}\mathcal{A}_1 c + c^{-1}dc,$$

where the c is the **canonical local trivialization** defined, generally, by

$$\phi_i^{-1}(t) = (u, c_i)$$

for

$$t = s_i(u)c_i.$$

Written in component form this relation becomes

$$\mathcal{A}_{2\mu}(u) = c^{-1}(u)\mathcal{A}_{1\mu}(u)c(u) + c^{-1}(u)\partial_\mu c(u),$$

which is the form of a **gauge transformation**.

6. Curvature on Principal G-Bundles

As in the above section we will here only take a brief look at curvature and find a localized form of it which, similar to the way the section based local form of a connection gives a gauge field, provide the (gauge or Yang-Mills) **field strength**. To start with we define the usual non-localized curvature 2-form Ω in a principal G-bundle $\mathfrak{G}$ as the covariant derivative of the connection 1-form ω, that is

$$\Omega = \nabla\omega \in \Omega^2(\mathcal{T} \in \mathfrak{G}) \otimes \mathfrak{g},$$

then we localize it analogously to the case of the connection and we have the following definition.

DEFINITION 5.12. *For a G bundle $\mathfrak{G}$ the **local form $\mathcal{C}$ of the curvature** Ω is*

$$\mathcal{C} = s^*\Omega,$$

where s is a local section defined on a chart U of $\mathcal{T}$ of $\mathfrak{G}$.

The localized curvature $\mathcal{C}$ may be expressed in terms of the gauge potential $\mathcal{A}$ as

$$\mathcal{C} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A},$$

where d is, as above, the exterior derivative on $\mathcal{T}$. To prove this one makes use of Cartan's structure equation and a proof can be found in [**Nak03**]. The localized curvature $\mathcal{C}$ act on elements X and Y in the tangent space $T\mathcal{T}$ according to

$$\mathcal{C}(X,Y) = d\mathcal{A}(x,Y) + [\mathcal{A}(X), \mathcal{A}(Y)].$$

Now if U is a chart with coordinates $t^\mu = \varphi(u)$, $\mathcal{A} = \mathcal{A}_{mu}dt^{mu}$ and we write

$$\mathcal{C} = \frac{1}{2}\mathcal{C}_{\mu\nu}dt^\mu \wedge dt^\nu$$

we find that $\mathcal{C}$ in component form becomes

$$\mathcal{C}_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]$$

Now we have an idea of what the localized curvature $\mathcal{C}$ is and as we remarked in the beginning this corresponds to the field strength in gauge theories.

7. Covariant Derivative on Associated Vector Bundles

In this section we will briefly outline how a connection $\mathcal{A}$ in a G-bundle $\mathfrak{G}$ specifies a covariant derivative ∇ in an associated vector bundle $\mathfrak{V}_{\mathfrak{G},\rho}$. The full details may be found in [**Fra04**]. This result is very useful since being able to covariantly differentiate sections in a vector bundle associated to a principal G-bundle is a necessity in gauge theory because this, for instance, enables one to construct gauge invariant actions. The reason why this induced covariant derivative is better than others one may define is that it is directly related to the connection on the G-bundle and is therefore by construction compatible with this. Which is essential since information extracted from both the (localized) connection in $\mathfrak{G}$ and the (localized) covariant derivative in $\mathfrak{V}_{\mathfrak{G},\rho}$ is used in the same theory. Our aim is, as in the previous section, to obtain a localized form of the object we are dealing with, which here is the covariant derivative, and we proceed as follows.

Let $\mathfrak{G}$ be a principal G-bundle and $\mathfrak{V}_{\mathfrak{G},\rho}$ be an, through the representation ρ, associated vector bundle. Choose a local section $s_i \in \Gamma(U_i, \mathcal{T}_{\mathfrak{G}})$ which by the canonical trivialization give

$$s_i(u) = \phi(u, id_{\mathcal{G}_{\mathfrak{G}}})$$

for $u \in U_i$ and let $\gamma : [0,1] \to \mathcal{B}_{\mathfrak{G}}$ be a curve in U_i and $\tilde{\gamma}$ its horizontal lift, i.e.

$$\tilde{\gamma}(t) = s_i(t)g_i(t), g_i(t) = g_i(\gamma(t)) \in \mathcal{G}_{\mathfrak{G}}$$

Now take a section

$$e_n(u) = \left[\left(s_i(u), e_n^0\right)\right] \in (T)_{\mathfrak{V}_{\mathfrak{G},\rho}}$$

where e_n^0 is the n:th basis vector of $\mathcal{F}_{\mathfrak{V}_{\mathfrak{G},\rho}}$. Thus we have that

$$e_n(t) = \left[\left(\tilde{\gamma}(t)g_i(t)^{-1}, e_n^0\right)\right] = \left[\left(\tilde{\gamma}(t), g_i(t)^{-1}e_n^0\right)\right].$$

Denote $\gamma(0)$ by u_0 and let $X \in T_u\mathcal{B}_{\mathfrak{G}}$ be a tangent vector to the curve γ at u_0. Then the covariant derivative of e_n along $\gamma(t)$ at u_0 is given by

$$\begin{aligned}\nabla_X e_n &= \left[\left(\tilde{\gamma}(0), \frac{d}{dt}\left(g_i(t)^{-1}e_n^0\right)|_{t=0}\right)\right] \\ &= \left[\left(\tilde{\gamma}(0)g_i(0)^{-1}\right), \mathcal{A}_i(X)e_n^0\right] \\ &= \left[s_i(0), \mathcal{A}_i(X)e_n^0\right],\end{aligned} \tag{26}$$

where $\mathcal{A}_i$ is the local connection form the previous section. Now if we express the $\mathcal{A}_i$ in local coordinates x^μ similarly to the localization of curvature in the previous section and combining this with the last equation above we obtain

$$\nabla e_n = \mathcal{A}_{in}^m e_m.$$

This is now the local form of the covariant derivative on an associated vector bundle induced from a connection $\mathcal{A}$ on the corresponding principal G-bundle.

It should be noted that the connection induces a covariant derivative unique up to representations.

8. Gauge Theories

In the language of bundles introduced in this chapter one may summarize gauge theory as the study of principal G-bundle connections. The gauge fields are identified with localized counterparts of these connections, which in physics is identified with physical fields, e.g. the electromagnetic field. The structure (Lie-)group of the bundle, now called the gauge group, represents the symmetries of the physical system under consideration, and from this group one can via a representation get an associated vector bundle on which the principal bundle connection induces a covariant derivative which provides a way to differentiate sections of the vector bundle. Furthermore, the field strength is identified with the localized curvature corresponding to the connection. In this section we will give an example containing the simplest of the common gauge theories, namely the $U(1)$ gauge theory for the electromagnetic interaction and then make some remarks on gauge theory in general.

Example 5.13 (Maxwell's Electromagnetism as a Gauge Theory). Maxwell's electromagnetism is described by the unitary group $U(1)$. This is a Lie group and a general element may be represented in the form $e^{i\theta}$ with $\theta \in \mathbb{R}$. Thus $U(1)$ may be thought of as the complex numbers of modulus one or the circle

$$x^2 + y^2 = 1$$

in $\mathbb{R}^2$ so the manifold part of $U(1)$ is the 1-sphere S^1. Considering the principal G-bundle $\mathfrak{G}$ with

$$\mathcal{F} = \mathcal{G} = U(1)$$

and the base space $\mathcal{B}$ being e.g. Minkowiski space-time, it is clear that $\mathfrak{G}$ is trivial so we have that the total space

$$\mathcal{T} = \mathbb{R}^n \times U(1)$$

and we need only one global trivialization. The other gauge concepts, in this case, also come in their simplest form with gauge potential

$$\mathcal{A} = \mathcal{A}_\mu dx^\mu$$

and field strength

$$\mathcal{C} = d\mathcal{A}.$$

At present the theory that provide the most accurate description of the forces of nature, with the exception of gravity, is the so called Standard model. In terms of the contents of this report the symmetries of the standard model may be expressed as the direct product

$$U(1) \times SU(2) \times SU(3),$$

where the components $U(1) \times SU(2)$ and $SU(3)$ are the symmetry groups of the electroweak- and strong interactions, respectively. The standard model is by no means a complete theory, rather an intermediate step in the quest for a unified theory of interaction. The most obvious reason for this conclusion is, of course, that gravity is not included. A key concept in understanding the relation between the building blocks of the standard model and the current attempts to extend the theory by enlarging the symmetry group, to e.g. $SU(5)$, is that of *spontaneous symmetry breaking* and this is something the reader interested in these matters should look into. Another notion of great importance is that of *renormalization*. The gauge theories are endowed with a property called renormalizability, which G. t'Hooft has shown, and this is one of the main reasons for their use as models of interactions. The problems with unifying gravity and the other interactions also has to do with renormalization, so the reader may find it worth while to investigate this very technical subject.
Among the most important things to do in gauge theory in the future is to put the modern theory on a solid mathematical foundation. At present there are several things that are unclear of which the existence of a, so called, mass gap, in quantum Yang-Mills theories are the most important. The mathematical community has recognized the importance of this and the Clay Mathematics Institute has included this as one of their Millennium Prize Problems. More specifically they offer a prize of $1,000,000 to anyone who solves the following problem.

Prove that for any compact simple gauge group G, quantum Yang-Mills theory on $\mathbb{R}^4$ exists and has mass gap $\Delta > 0$.

For the details concerning the problem and the prize visit the Clay Mathematics Institute on the Internet at `www.claymath.org`.

This constitutes a suitable ending to our report, and we do not include any exercises so that the reader may fully devote him- or herself to the above problem, good luck!

APPENDIX A

Prerequisites

1. Topology

Definition A.1. *A class $\mathcal{T}$ of subsets of a set $X \neq \emptyset$ is a* ***topology*** *on X if*

(1) $\emptyset, X \in \mathcal{T}$.
(2) *The union of any number of sets in $\mathcal{T}$ belongs to $\mathcal{T}$.*
(3) *The intersection of a finite number of sets in $\mathcal{T}$ belongs to $\mathcal{T}$.*

The elements of $\mathcal{T}$ are called **open sets**, or $\mathcal{T}$-open to indicate that they are open with respect to the topology $\mathcal{T}$.

The pair $(X, \mathcal{T})$ is called a **topological space**. Usually one denotes the space by X instead of $(X, \mathcal{T})$.

Definition A.2. *Let $(X, \mathcal{T})$ be a topological space. A class $\mathcal{B} \subset \mathcal{T}$ is a* ***basis*** *for $\mathcal{T}$ if every open set $S \in \mathcal{T}$ is the union of elements in $\mathcal{B}$.*

Definition A.3. *Let X and Y be topological spaces. The* ***product topology*** *on $X \times Y$ is the topology having as basis the collection $\mathcal{B}$ of all sets of the form $U \times V$ where U is an open subset of X and V an open subset of Y. Furthermore, let $\pi_1 : X \times Y \to X$ be defined by*

$$\pi_1(x, y) = x$$

and $\pi_2 : X \times Y \to Y$ be defined by

$$\pi_2(x, y) = y.$$

These maps π_1 and π_2 are called ***the projections*** *of $X \times Y$ onto its first and second component, respectively.*

Definition A.4. *A topological space X is said to be* ***disconnected*** *if it is the union of two nonempty disjoint open sets. Otherwise the topological space is said to be* ***connected.***

Definition A.5. *A topological space X is said to be* ***compact*** *if every covering of X with open subsets admits a finite sub-cover.*

Definition A.6. *Let X be a topological space. If for any two distinct points u and v in X there exist neighborhoods U of u and V of v such that*

$$(U \cap V) \neq \emptyset$$

the space X is called a ***Hausdorff space****.*

Definition A.7. *A topological space X is said to be* ***second countable*** *if it has a countable base. That a topological space X has a base means that every open set can be constructed from a specific union of open sets. This union of open sets are then said to be defining the base. A countable base is then a base where the base elements are countable.*

2. Analysis

Definition A.8. *A function is said to be* ***smooth*** *or C^∞ if it is differentiable infinitely many times.*

Definition A.9. *A* ***metric*** *on a set S is a map $m : S \times S \to \mathbb{R}$ such that $\forall\, x_1, x_2, x_3 \in S$ satisfies*

(1) $m(x_1, x_1) = 0$,
(2) $m(x_1, x_2) > 0$ *if* $x_1 \neq x_2$,
(3) $m(x_1, x_2) = m(x_2, x_1)$,
(4) $m(x_1, x_2) + m(x_2, x_3) \geq m(x_1, x_3)$ *(the Triangle inequality).*

A set equipped with such a map is said to be a ***metric space****. If the second condition is replaced with $m(x_1, x_2) \geq 0$ for $x_1 \neq x_2$, then the map m is called a* ***pseudo metric****.*

Definition A.10. *The set of* ***germs of functions*** *is defined as*

$$C_p^\infty(M) = \{f : M \to \mathbb{R} \mid f \text{ differentiable}\}/\sim,$$

where the equivalence relation $\sim$ is defined by $f \sim \tilde{f}$ if and only if f and $\tilde{f}$ coincide in a neighborhood of p.

3. Algebra

Definition A.11. *A* ***group*** *G is a nonempty set together with a binary operation $G \times G \to G$ with the properties:*

(1) *if* $g_1, g_2 \in G$ *then* $g_1 g_2 \in G$
(2) $g_1(g_2 g_3) = (g_1 g_2) g_3 \quad \forall g_1, g_2, g_3 \in G$
(3) *there is a element* $e \in G$ *such that* $ge = eg = e\ \forall g \in G$
(4) $\forall g \in G$ *there is an element* $g^{-1} \in G$ *such that* $gg^{-1} = e$.

Definition A.12. *A group G is called* ***commutative****, or* ***abelian****, if*

$$g_1 g_2 = g_2 g_1 \quad \forall g_1, g_2 \in G.$$

Definition A.13. *A* ***normal subgroup*** *H of a group G is a group such that the left- and right cosets of H in G coincide.*

Definition A.14. *A* ***simple group*** *G is a group where the only normal subgroups of G are the* ***trivial group*** *and G.*

DEFINITION A.15. *An **ideal** is a subset I of a ring R, if it together with the addition operator of the ring $(I,+)$ forms a subgroup of the abelian group of the ring $(R,+)$, and also if the multiplicative operation is closed in the subset for all elements in the ring R .*

$$\forall i \in I\,, r \in R \quad ri \in I \; or \; ir \in I$$

References

[AW95] G.B. Arfken and H.J. Weber. *Mathematical Methods for Physicists*. Academic Press, 4th edition, 1995.

[Car94] M. P. do Carmo. *Differential forms and applications*. Springer-Vlg, New York, 1994.

[Che99] S.S. Et al Chern. *Lectures on Differential Geometry*. World Scientific Publishing, 1999.

[D'I92] R. A. D'Inverno. *Introducing Einstein's relativity*. Clarendon Press, Oxford, 1992.

[Eul00] N. Euler. *A first course in nonlinear mathematical physics : nonlinear differential equations, Lie algebras, continuous transformations and applied differential geometry. Part 1*. Luleå university of technology, Luleå, 2000.

[FH91] W. Fulton and J. Harris. *Representation Theory : A First Course*. Springer-verlag, New York, 1991.

[Fra04] T. Frankel. *The geometry of physics : an introduction*. Cambridge University Press, Cambridge, 2 edition, 2004.

[GA74] V. Guillemin and Pollack A. *Differential Topology*. Prentice Hall, Englewood Cliffs, 1974.

[Ger85] R. Geroch. *Mathematical Physics*. University of Chicago Press, Chicago, 1985.

[Küh02] W. Kühnel. *Differential geometry : curves, surfaces, manifolds*. American Mathematical Society, Providence, R.I., 2002.

[Lan03] S. Lang. *Algebra*. Springer-verlag, New York, rev. 3rd edition, 2003.

[Mil97] J. W. Milnor. *Topology from the differentiable viewpoint*. Princeton Univ. Press, Princeton, N.J., 1997.

[Mun00] J.R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, 2nd international edition, 2000.

[Nak03] M. Nakahara. *Geometry, Topology and Physics*. IOP Publishing, London, 3rd edition, 2003.

[Opr97] J. Oprea. *Differential geometry and its application*. Prentice Hall, Upper Saddle River, N.J., 1 edition, 1997.

[Pra96] V.V. Prasolov. *Problems and theorems in linear algebra*. American Mathematical Society, Providence, R.I., 1996.

[Spi65] M. Spivak. *Calculus on manifolds : a modern approach to classical theorems of advanced calculus*. AddisonWesley, Redwood City, 1965.

[Spi99] M. Spivak. *A comprehensive introduction to differential geometry. Vol. 1*. Publish or perish, Houston, Tex., 3 edition, 1999.

[Ste51] N. Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, Princeton, 1951.

Index