university of groningen / faculty of science and engineering

University of Groningen
Faculty of Science and Engineering
**BACHELOR THESIS**

Tim Eerke Folgering

# Chern-Simons theory and its relation to inflation

Van Swinderen Institute and Bernoulli Institute

Supervisors: Ema Dimastrogiovanni (VSI)
Tamás Görbe (BI)

Study programme: Physics and Mathematics

**Abstract**

This thesis gives an introduction to Gauge theory, first by considering the Yang-Mills approach, and then by using connections. Then it will explain the origin of the Chern-Simons term, and give an introduction into inflation, more specifically into slow-roll inflation. Finally, it adds the Chern-Simons term to the inflation Lagrangian and examines its effects.

# Contents

**Introduction** **3**

**1 Preliminary** **4**
1.1 Differential Geometry . . . . . 4
1.2 Lie theory . . . . . 7

**2 Motivation** **10**
2.1 Differential geometry in electromagnetism . . . . . 10
2.1.1 Maxwell's equations . . . . . 10
2.2 Introduction to Gauge theory . . . . . 12
2.2.1 Lagrangian approach . . . . . 12
2.2.2 Yang-Mills . . . . . 16

**3 Connection** **18**
3.1 Principal fibre bundles . . . . . 18
3.2 Connection . . . . . 19
3.2.1 Splitting the bundle . . . . . 20
3.2.2 Connection 1-form . . . . . 21
3.2.3 Maurer-Cartan form . . . . . 22
3.3 Curvature . . . . . 22

**4 Gauge theory** **24**
4.1 Relating Gauge theory to connection 1-forms . . . . . 24
4.2 Connection 1-form on a trivial section . . . . . 26
4.3 Connection 1-form on any section . . . . . 29
4.4 Relating back to Electromagnetism and Yang Mills . . . . . 30

**5 Chern-Simons theory** **32**
5.1 Koszul Connection . . . . . 32
5.2 Chern-Weil theory . . . . . 35
5.3 Chern-Simons theory . . . . . 37

**6 Inflation** **38**
6.1 Introduction to inflation . . . . . 38
6.1.1 Christoffel symbols . . . . . 39
6.1.2 Energy momentum tensor . . . . . 41
6.1.3 Motion of our universe . . . . . 42
6.2 Space Lagrangian . . . . . 44
6.3 Inflaton matter field . . . . . 45

6.4 Rewriting the inflation conditions . . . . . 45
6.5 Slow-roll inflation . . . . . 46

**7 Chern-Simons term in inflation** **48**
7.1 Adding the Chern-Simons term . . . . . 48
7.1.1 How is this term related to Chern-Simons theory? . . . . . 49
7.1.2 What does the term do for inflation? . . . . . 49

**8 Conclusion** **51**

**Bibliography** **52**

**A Intuition with Schrödinger equation** **53**

**B Trace identities** **57**

# Introduction

To describe all laws of physics in one theory is one of the big questions of physics. Currently the most promising way to do this is by using Gauge theory. Gauge theories are extremely important in physics, solving some of the most fundamental questions we have about physics. It is used to formulate electromagnetism, the weak and the strong forces, as well as combining these three into the standard model (1). Gauge theory defines a Lagrangian which is invariant under operations of a Lie group, an object which is both a manifold and a group (2).

Chern-Simons theory is such a Gauge theory, originating from differential geometry. It is used in a multitude of concepts in physics and mathematics such as cosmology and knot theory (3). The part that this thesis focuses on is gravity and inflation. Chern-Simons theory is also a topological theory, meaning that it is a more general way to describe objects than the metric theory we use in general relativity (4).

The goal of this thesis is to explain Gauge theory from both a physics perspective and a mathematics perspective. It does so by looking one of the more influential Gauge theories, namely the Yang-Mills theory for the physics perspective, and using the language of differential geometry for the mathematics perspective. Then we look at how these two points of view relate to one another. Using what was learned about connections we look at Chern-Weil theory, which gives rise to the Chern-Simons term. In the second part inflation is explained, looking at the Lagrangian of the inflaton field we derive the conditions for slow-roll inflation. And finally add the Chern-Simons term to the inflaton field and see how it changes inflation.

Here a short distinction of the physics and mathematics parts of this thesis. Chapter 1 gives a short background on the essential mathematical concepts used throughout the thesis (background). Chapter two motivates this notation for physicists and introduces Gauge theory (physics). Chapter 3 explains the concept of a fibre bundle and what a connection on such a fibre bundle is (mathematics). Chapter 4 explains Gauge theory using these connections and relates it back to physics (mathematics and physics). Chapter 5 derives the Chern-Simons term using a new (related) kind of connection (mathematics). Then the final two chapters 6 and 7 explain inflation and how the Chern-Simons term is used in inflation (physics).

# Chapter 1

# Preliminary

In order to introduce Chern-Simons theory, a strong understanding of Gauge theory is required. Therefore this chapter will collect some essential concepts (e.g. Lie groups and Fibre bundles) that are used in Gauge theory. This chapter is based on the books Bleecker - Gauge Theory and Variational Principles (2) and Sontz - Principal bundles: the classical case (5).

## 1.1 Differential Geometry

The very first idea to understand is that of a manifold. A (smooth) manifold is a topological space (examples of a topological space include, spheres, Euclidean spaces, and a cross +) which is not too scarce (Hausdorff) and not too dense (second countable) and can locally be described as a Euclidean space. Hausdorff being that there is always space between any two points. Second countable meaning that we can have a countable basis of the manifold. The last part (locally Euclidean) is important since this means that it is possible to define continuity and differentiation on such a space in a "natural" way. The simplest example of a manifold is the $n$-dimensional euclidean space $\mathbb{E}^n$ itself. Another simple example is the circle, which can locally be approximated by a line. So any smooth shape. from previous examples, the sphere and euclidean spaces are manifolds. The cross is not, because at the intersection point it cannot be described as an euclidean space.

The way we describe a manifold as Euclidean space is by using Charts, those are maps from the manifold to euclidean space which stitch bits of Euclidean space together. Two charts are related by a transition map which is a smooth diffeomorphism (an invertible map which is $C^\infty$-differentiable and whose inverse is also $C^\infty$-differentiable).

**Definition 1.1.1** (Vector field)**.** *Let $TM = \cup_{x\in M} T_xM$. A **vector field** on the manifold $M$ is a function $Y : M \to TM$ s.t. $Y_x \in T_xM$. The set of all vector fields on $M$ is often denoted by $\Gamma(TM)$ or by $\mathcal{X}(M)$.*

The vector field basically assigns a tangent vector at each point of the manifold $M$. So for $\mathbb{R}^2$ it assigns a direction (vector) at every point in $\mathbb{R}^2$. For a sphere, it would give a vector on the planes which are tangent to the sphere at each point (see figure 1.1).

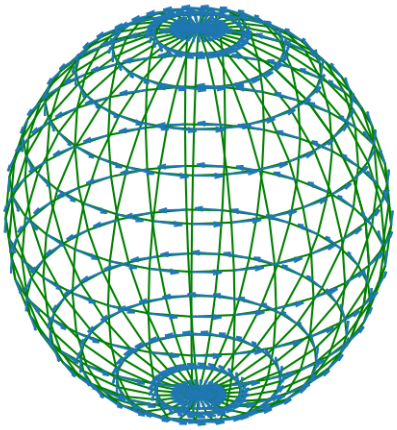

Figure 1.1: Example of a vector field on the 2-sphere.

$T_xM$ is the tangent space at a specific point, for the two-dimensional sphere this would be the tangent plane.

**Definition 1.1.2** (Tensor of a vector space)**.** *Let $E$ and $F$ be vector spaces and $E^*$ the dual space of $E$. Then $T^{p,q}(E,F)$ is the set of multilinear functions (**F-valued Tensors**) $f : \underbrace{E^* \times \cdots \times E^*}_{p \text{ copies}} \times \underbrace{E \times \cdots \times E}_{q \text{ copies}} \to F$. $T^{0,0}(E,F) = F$ and $T^{p,q}(E,\mathbb{R})$ is denoted by $T^{p,q}(E)$.*

A $k$-form is a set of very specific tensors, namely the anti-symmetric $T^{0,k}(E,F)$ tensors. Formally that means the following

**Definition 1.1.3** (k-forms)**.** *The **k-form** on $M$ is given by*

$$\omega : M \times \underbrace{TM \times \cdots \times TM}_{k \text{ copies}} \to \mathbb{R}. \tag{1.1}$$

*We often consider (where the $x$ could be implicit)*

$$\omega_x : \underbrace{T_xM \times \cdots \times T_xM}_{k \text{ copies}} \to \mathbb{R} \quad x \in M, \tag{1.2}$$

*such that it is anti-symmetric.*

$$\omega(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma)\omega(v_1, \cdots, v_k), \tag{1.3}$$

*for any permutation $\sigma$ of the index set $\{1, \ldots, k\}$. We denote the set of $k$-forms on $M$ by $\Omega^k(M)$.*

The basis of $T_xM$ is the set $\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\}$ , which defines the basis of its dual by, $\{dx_1, \cdots, dx_n\}$, using $dx_i(\frac{\partial}{\partial x_j}) = \delta_i^j$.We therefore write the $k$-form

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} f_{i_1, \ldots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \tag{1.4}$$

Where this $\wedge$ is called the **wedge product** and has the following property: $dx_i \wedge dx_j = -dx_j \wedge dx_i$. Which therefore respects the earlier definition given of the k-form.
Now that we know what a k-form is we want to do something with this, the interesting properties that we will mostly be using are the exterior derivative and the Hodge star.

**Definition 1.1.4** (Exterior derivatives)**.** *The* ***exterior derivative*** *is a map* $d : \Omega^k(M) \to \Omega^{k+1}(M)$ *in the following way. Take some k-form* $\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} f dx_{i_1} \wedge \cdots \wedge dx_{i_k}$*. Then*

$$d\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\partial f}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \tag{1.5}$$

The exterior derivative has a very nice property, namely

$$\begin{aligned} dd\omega &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\partial^2 f}{\partial x_j \partial x_k} f dx_j \wedge dx_k \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= -\left( \sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\partial^2 f}{\partial x_k \partial x_j} f dx_k \wedge dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \right) \\ &= -dd\omega. \end{aligned}$$

Showing that

$$d^2\omega = 0. \tag{1.6}$$

**Definition 1.1.5** (Hodge star)**.** *Given a k-form* $\omega$ *on an n-dimensional manifold M, then the* ***Hodge star operation*** $\star$ *acts as:*

$$\star\omega = \star \sum_{\sigma \in S_n} f dx_{\sigma(1)} \wedge \cdots \wedge dx_{\sigma(k)} = \operatorname{sgn}(\sigma) \sum_{\sigma \in S_n} f dx_{\sigma(k+1)} \wedge \cdots \wedge dx_{\sigma(n)} \tag{1.7}$$

*So* $\star\omega$ *is a* $n - k$ *form on M.*

Let us look at what happens when we take the hodge star twice, we get back the original $k$-form, but with what sign? We get a plus sign if parity of $n$ and $k$ match, and -1 if $n$ and $k$ have different parity.This can be summarised as the following:

$$\star \star \omega = (-1)^{k(n-k)} \omega \tag{1.8}$$

where $\omega$ is a $k$-form in an $n$-dimensional manifold.

We can now find the adjoint of the exterior derivative,

$$\delta = (-1)^{nk+n+1} \star d \star . \tag{1.9}$$

We can immediately see from $dd\omega = 0$ and $\star\star\,\omega = (-1)^{k(n-k)}\omega$ that

$$\delta^2\omega = 0 \tag{1.10}$$

for $\omega$ a $k$-form. We also see that $\delta$ maps $k$-forms in the following way

$$\begin{aligned}\delta(k\text{-form}) &= (-1)^{nk+n+1} \star d \star k\text{-form} = (-1)^{nk+n+1} \star d(n-k)\text{-form} \\ &= (-1)^{nk+n+1} \star (n-k+1)\text{-form} = (-1)^{nk+n+1}(n-n+k-1)\text{-form} \\ &= (-1)^{nk+n+1}(k-1)\text{-form}.\end{aligned}$$

So it sends $k$-forms to $(k-1)$-forms.

Let's now look at some examples of these maps. Consider we are in $\mathbb{R}^3$ with (local) coordinates $x, y, z$. Then we define the 2-form $\omega = xdx \wedge dy + ydx \wedge dz$. First the exterior derivative

$$\begin{aligned}d\omega &= \frac{\partial x}{\partial z}dz \wedge dx \wedge dy + \frac{\partial y}{\partial y}dy \wedge dx \wedge dz \\ &= dy \wedge dx \wedge dz = -dx \wedge dy \wedge dz.\end{aligned}$$

Now let's look at the Hodge star operation with the same $\omega$

$$\star\omega = xdz - ydy.$$

The adjoint exterior derivative of $\omega$ is

$$\begin{aligned}\delta\omega = \star d \star \omega &= \star d(xdz - ydy) \\ &= \star(\frac{\partial x}{\partial x}dx \wedge dz + \frac{\partial x}{\partial y}dy \wedge dz - \frac{\partial y}{\partial x}dx \wedge dy - \frac{\partial y}{\partial z}dz \wedge dy) \\ &= -dy.\end{aligned}$$

where we see that $nk + n + 1 = 3 \cdot 2 + 3 + 1$ is even, meaning we get a $+$ sign in front of the $\delta$.

## 1.2 Lie theory

The goal of Lie theory is to combine manifolds and groups, to define symmetries on (physical) spaces.
First some intuition into what groups are. A group is a set of operations that leave an object unchanged. For example consider an equilateral triangle, which can be rotated by 120 and 240 degrees about its centre and it still looks the same. A reflection along any of the altitudes also leaves the triangle looking the same. Another example is a circle, which can be rotated by any degree about the centre, and reflect it along any diameter.

A Lie group is a group and a manifold. The following definition formalises this.

**Definition 1.2.1** (Lie group)**.** *Let $G$ be a manifold, and also a group such that the group multiplication $\mu : G \times G \to G$ given by $\mu(g_1, g_2) \to g_1 g_2$ and the inverse $i : G \to G$ given by $i(g) \to g^{-1}$ are $C^\infty$ maps. Then $G$ is called a* ***Lie group****.*

One of the simplest examples is the set of rotations in the complex plane, that is $\{e^{i\theta} : \theta \in [0, 2\pi)\}$, with complex number multiplication. We can consider a rotation by an angle $\phi$ applied to any element in this group as $e^{i\theta}e^{i\phi} = e^{i(\theta+\phi)}$ which is again an element in the group that rotates by an angle of $\theta + \phi$. In the complex plane when tracing out $e^{i\theta}$ the $\theta$ ranges from 0 to $2\pi$ you will see a circle. From this we can see that this group is also a manifold.

Though Lie groups are much broader, we will focus on the ones that can be represented by matrices. This will make the following definitions easier to state and grasp and will not be a limitation in further chapters. The most important examples are unitary matrices i.e. matrices whose inverse is the conjugate transpose. This gives rise to the Lie group $U(n)$, the set of $n \times n$ unitary matrices whose determinant is modulus 1. $SU(n)$ is the set of $n \times n$ special unitary matrices, where special means that the determinant is $+1$.

Now we recall some relevant ways in which Lie groups and manifolds interact.

*Note* (Lie group and manifold relations). Let $G$ be a Lie group (with identity element $e$) and $M$ be a smooth manifold.
A **right action** of $G$ on $M$ is a smooth map $M \times G \to M$ given by $(x, g) \mapsto x \cdot g = xg$, s.t. it satisfies $xe = x$ and $(xg)h = x(gh)$ $\forall x \in M, g, h \in G$.
If a right action exists, then $\forall g \in G$ define

$$R_g : M \to M \text{ by } R_g(x) := xg \quad \forall x \in M, \tag{1.11}$$

then this is a diffeomorphism of $M$ whose inverse is $R_{g^{-1}}$. This is called the **right action map**. The **left action** and the **left action map** $L_g$ are defined analogously.
Given the right actions of $G$ on smooth manifolds $M$ and $N$, and a smooth function $f : M \to N$ is ***G*-Equivariant/*G*-morphism** if $f(xg) = f(x)g$.
An **action** of $G$ on $M$ is **free** if $xg = x$ for any $x \in M$ implies that $g = e$.
Let $M$ and $N$ be smooth manifolds and $f : M \to N$ a diffeomorphism. Define the **pushforward function** as

$$f_* : \mathcal{X}(M) \to \mathcal{X}(N) \quad X \mapsto f_*X = df \circ X \circ f^{-1}. \tag{1.12}$$

**Definition 1.2.2** (Left-invariant vector field)**.** *Let $G$ be a Lie group and $X : G \to TG$ be a vector field on $G$. $X$ is called the **left-invariant vector field** if $L_{g*}X = XL_g$ $\forall g \in G$. $L_{g*}$ being the push forward of $L_g$.*

**Definition 1.2.3** (Lie algebra)**.** *Let $G$ be a Lie group. We define $\mathcal{G} = T_eG$ where $e$ is the identity element of $G$, assume the elements of $\mathcal{G}$ are left-invariant vectors. Take $X, Y \in \mathcal{G}$, define the map $[\cdot, \cdot] : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ as the matrix commutator $[X, Y] = XY - YX$.*

For an example of the Lie algebra we look at the trusted example of $U(1)$ the circle given by $e^{i\theta}$, at the identity $e^{i0} = 1$ the tangent space is given by $\mathcal{U}(1) = i\mathbb{R}$ see figure 1.2. Since elements in $(\mathbb{R}, +)$ commute, so do elements in $(i\mathbb{R}, +)$. If all elements in $\mathcal{G}$ commute, then such a Lie algebra is called **abelian**.

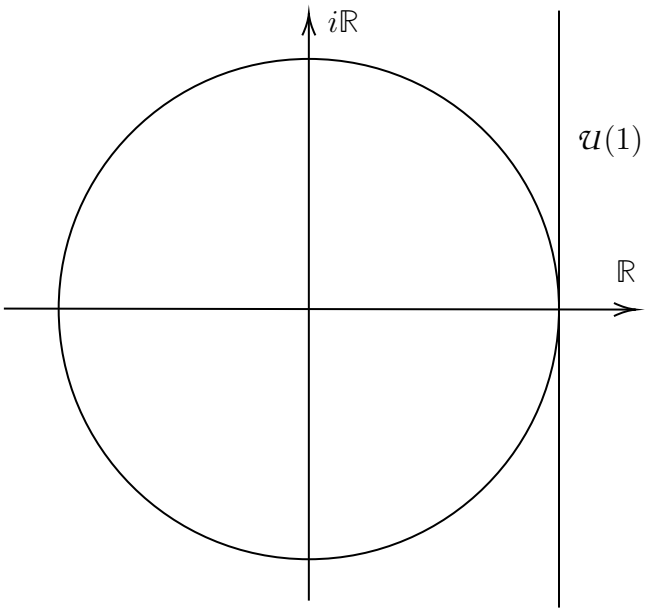


Figure 1.2: Visual representation of the Lie group $U(1)$ with Lie algebra $\mathcal{U}(1)$

.

# Chapter 2

# Motivation

This chapter will show the applications of differential forms in physics, to give intuition on how to work with them. As well as introducing Gauge theory through the Lagrangian of a field. The material in this chapter is based on the books of Sontz (5) and of Rubakov (6).

## 2.1 Differential geometry in electromagnetism

### 2.1.1 Maxwell's equations

First recall Maxwell's equations:

$$\nabla \cdot \vec{E} = 4\pi\rho, \qquad \nabla \cdot \vec{B} = 0,$$
$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \qquad \nabla \times \vec{B} = 4\pi\vec{J} + \frac{\partial \vec{E}}{\partial t}.$$

(we take units such that $c = 1$, and $\mu_0 = 4\pi$ meaning that $\varepsilon_0 = \frac{1}{4\pi}$). Consider 3-dimensional space on which we have defined $\vec{E}$ and $\vec{B}$. This means that, for now, we are in a 3-dimensional manifold. The differential forms of a 3-dimensional manifold have, for 0-forms 1 dimension, 1-forms 3 dimensions, for 2-forms 3 dimensions and for 3-forms 1 dimension, as shown in table 2.1.

| form | Dimensions | Basis elements |
|---|---|---|
| 0 | 1 | scalar |
| 1 | 3 | $dx_1, dx_2, dx_3$ |
| 2 | 3 | $dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2$ |
| 3 | 1 | $dx_1 \wedge dx_2 \wedge dx_3$ |

Table 2.1: 3-forms explained

Since $\vec{E}$ is 3-dimensional it must be a 1- or a 2-form, mathematically it would not change anything about the physical properties, but physically it makes more sense for $\vec{E}$ to be a 1-form since having forces as a 1-form makes consistent definitions with energies.
So define $E = E_1 dx_1 + E_2 dx_2 + E_3 dx_3 = E_i dx_i$, where we assume summation

over repeated indices. Let us take the exterior derivative of $E$,

$$\begin{aligned}dE =&\frac{\partial E_1}{\partial x_1}dx_1\wedge dx_1+\frac{\partial E_1}{\partial x_2}dx_2\wedge dx_1+\frac{\partial E_1}{\partial x_3}dx_3\wedge dx_1+\\&\frac{\partial E_2}{\partial x_1}dx_1\wedge dx_2+\frac{\partial E_2}{\partial x_2}dx_2\wedge dx_2+\frac{\partial E_2}{\partial x_3}dx_3\wedge dx_2+\\&\frac{\partial E_3}{\partial x_1}dx_1\wedge dx_3+\frac{\partial E_3}{\partial x_2}dx_2\wedge dx_3+\frac{\partial E_3}{\partial x_3}dx_3\wedge dx_3\\=&(\frac{\partial E_3}{\partial x_2}-\frac{\partial E_2}{\partial x_3})dx_2\wedge dx_3+(\frac{\partial E_1}{\partial x_3}-\frac{\partial E_3}{\partial x_1})dx_3\wedge dx_1+(\frac{\partial E_2}{\partial x_1}-\frac{\partial E_1}{\partial x_2})dx_1\wedge dx_2.\end{aligned}$$

We see that these terms are the same as for $\nabla\times\vec{E}$, motivating us to write $\vec{B}$ as a 2-form (also 3 dimensional) as: $B=B_1dx_2\wedge dx_3+B_2dx_3\wedge dx_1+B_3dx_1\wedge dx_2=\frac{1}{2}\epsilon_{ijk}B_idx_j\wedge dx_k$. So apparently the exterior derivative on a 1-form in a 3-dimensional manifold is equivalent to the curl. This is not true when we apply an exterior derivative to a 2-form, since exterior derivatives take $k$-forms to $(k+1)$-forms. So what does happen when we apply an exterior derivative to a 2-form $B$? We get

$$\begin{aligned}dB&=\frac{\partial B_1}{\partial x_1}dx_1\wedge dx_2\wedge dx_3+\frac{\partial B_2}{\partial x_2}dx_2\wedge dx_3\wedge dx_1+\frac{\partial B_3}{\partial x_3}dx_3\wedge dx_1\wedge dx_2\\&=(\frac{\partial B_1}{\partial x_1}+\frac{\partial B_2}{\partial x_2}+\frac{\partial B_3}{\partial x_3})dx_1\wedge dx_2\wedge dx_3.\end{aligned}$$

And this is the same as the divergence of $B$! In a very similar way we have the codifferential $\delta=\star d\star$, let's apply this to $E$ and see what happens. $\delta E=\star d\star E$, $\star E$ takes $E$ to a 2-form, and the exterior derivative on a 2-form gives the divergence, applying the final Hodge star brings $\delta E$ to a 0-form. So the codifferential on a 1-form is the divergence of a 1-form. A good exercise is to see what happens when we take $\delta B$. The answer is that this gives the curl of B as a 1-form!

Now we get to why we decide to introduce this new way to describe divergences and curls. We know that we can write the fields using potentials as $\vec{E}=-\nabla\phi-\frac{\partial\vec{A}}{dt}$, and $\vec{B}=\nabla\times\vec{A}$ and we found that a curl on a 1-form gives a 2-form, meaning that $\vec{A}$ is a 1-form. Since $\phi$ is a scalar, it will be either a 0-form or a 3-form. Since $E$ is a 1-form and that there is no "nice" map from 3-forms to 1-forms, it is natural to assume $\phi$ is a 0-form. Applying the exterior derivative on a 0-form gives a 1-form exactly like the gradient does. Though this works, it would really be nice to be able to write $\phi$ also as a 1-form and combine it with $\vec{A}$. So let's increase the dimension of our manifold to 4 dimensions and introduce the only logical coordinate t (since we are really already using it) where the minus sign appears due to Minkowski space as in special relativity. We find $A=-\phi dt+A_1dx_1+A_2dx_2+A_3dx_3$, which is the potential 1-form. Let's look at the partial derivative of $A$, summing over repeated indices and removing the zero terms we get

$$dA=-\frac{\partial\phi}{\partial x_i}dx_i\wedge dx_0+\frac{\partial A_i}{\partial x_0}dx_0\wedge dx_i+\frac{\partial A_i}{\partial x_j}dx_j\wedge dx_i,$$

the first two terms together make $\vec{E}$ the last term is the curl of $\vec{A}$, therefore $B$.

$$dA=E\wedge dt+B. \tag{2.1}$$

Now you might see the uses of the introduction of forms. Define $F = dA$ the field strength of the potential $A$. From equation 1.6 we know that $dF = ddA = 0$. Look at how the 2-form $F$ acts under the exterior derivative:

$$\begin{aligned} dF &= d(E \wedge dt + B) = d(E \wedge dx_0) + dB, \\ &= (dE \wedge dt) + \frac{\partial B}{\partial t} \wedge dt + \frac{1}{2}\epsilon_{ijk}\frac{\partial B_i}{\partial x_i} dx_i \wedge dx_j \wedge dx_k = 0. \end{aligned}$$

The first two terms yield $\nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0$, and the final term is $\nabla \cdot \vec{B} = 0$. Reclaiming the two homogeneous Maxwell equations!

If we take the codifferential of $F$ we get the following:

$$\delta F = \frac{\partial E_i}{\partial x_i} dt - \frac{\partial E_k}{\partial t} dx_k + \epsilon_{ijk}\frac{\partial B_i}{\partial x_j} dx_k.$$

This is in vector notation $\delta F = (\nabla \cdot E)dt + (\nabla \times B - \frac{\partial E}{\partial t})_k dx_k$ Showing that the non-homogeneous Maxwell's equations can be written as $\delta F = 4\pi(\rho dt + j_k dx_k)$. We will look further, since $\delta\delta F = 0$ (equation 1.10). We should consider what that tells us

$$\begin{aligned} \delta\delta F &= \delta[4\pi(\rho dt + j_k dx_k)] \\ &= 4\pi(\frac{\partial \rho}{\partial t} + \epsilon_{ijk}\frac{\partial j_k}{\partial x_k}). \end{aligned}$$

Which is in vector notation: $0 = \frac{\partial \rho}{\partial t} + \nabla \cdot j$ exactly the continuity equation. Implying the conservation of electric charge.

| | |
|---|---|
| Homogeneous | $dF = 0$ |
| Non-Homogeneous | $\delta F = 4\pi J$ |
| Continuity equation | $\delta^2 F = 0$ |

Table 2.2: Maxwell's equations in differential geometry language

## 2.2 Introduction to Gauge theory

Now moving on to Gauge theory! A very rudimentary way to look at Gauge theory is to say that $A$ a potential 1-form as defined earlier, can be transformed by $A' = A + df$ where $f$ is some (differentiable) function. Then $F' = dA' = d(A+df) = dA+ddf = dA = F$. Then we say that $F$ is gauge invariant by gauge transformation $A' = A + df$. In this section this concept will be expanded on.

### 2.2.1 Lagrangian approach

Consider a complex-valued function

$$\psi : \mathbb{R}^4 \to \mathbb{C}^n \tag{2.2}$$

and the complex vector field Lagrangian:

$$\mathcal{L} = \partial_\mu \psi^* \partial_\mu \psi - m^2 \psi^* \psi, \tag{2.3}$$

where $^*$ indicates a complex conjugate.
We would like to transform $\psi \mapsto \psi_1$ in a way where the Lagrangian stays invariant (the same). Looking at the second term, $m^2\psi^*\psi$, it is reasonable to consider the following transformation: $\psi \mapsto \psi_1 = U_1\psi$, where $U_1$ is an $n \times n$ unitary matrix. $\psi_1^*\psi_1 = (U_1\psi)^* U_1\psi = \psi^* U_1^* U_1 \psi = \psi^*\psi$. Since for now we have that $\partial_\mu U_1 = 0$, we find a very similar relation: $\partial_\mu \psi_1^* \partial_\mu \psi_1 = \partial_\mu (U_1\psi)^* \partial_\mu U_1 \psi = \partial_\mu \psi^* U_1^* U_1 \partial_\mu \psi = \partial_\mu \psi^* \partial_\mu \psi$. So this would be invariant under this transformation. But what if $\partial_\mu U_1 \neq 0$? This does not matter for the second term, since in the same way as before we still have $\psi_1^*\psi_1 = \psi^*\psi$, but for the first term we do find a change,

$$\begin{aligned}\partial_\mu \psi_1^* \partial_\mu \psi_1 &= \partial_\mu(\psi^* U^*)\partial_\mu(U\psi) = (\partial_\mu(\psi^*)U^* + \psi^*\partial_\mu(U^*))(\partial_\mu(U)\psi + U\partial_\mu(\psi)) \\ &= \partial_\mu(\psi^*)U^*\partial_\mu(U)\psi + \partial_\mu\psi^*\partial_\mu\psi + \psi^*\partial_\mu(U^*)\partial_\mu(U)\psi + \psi^*\partial_\mu(U^*)U\partial_\mu(\psi).\end{aligned}$$

Since $\partial_\mu(\psi^*)U^*\partial_\mu(U)\psi + \psi^*\partial_\mu(U^*)\partial_\mu(U)\psi + \psi^*\partial_\mu(U^*)U\partial_\mu(\psi)$ is not in general 0. Which means that this is not invariant, therefore we should consider the partial derivatives to change during the transformation as well, therefore we look at the following. To have the solution follow convention we take $U_1 = e^{iH}$, where $H$ is a hermitian matrix. (Any unitary matrix can be written as the exponential of a skew hermitian matrix, which $iH$ is). This also means that $H$ is dependent on the partials.

$$\begin{aligned}\partial_\mu\psi_1 &= \partial_\mu(U_1\psi) \\ &= \partial_\mu(e^{iH}\psi) \\ &= i\partial_\mu(H)e^{iH}\psi + e^{iH}\partial_\mu\psi.\end{aligned}$$

Taking $B = -i\partial_\mu H$, and returning $U_1 = e^{iH}$

$$\begin{aligned}\partial_\mu\psi_1 &= -BU_1\psi + U_1\partial_\mu\psi \\ \partial_\mu\psi_1 &= -B\psi_1 + U_1\partial_\mu\psi.\end{aligned}$$

Which makes

$$\partial_\mu\psi = U_1^{-1}(\partial_\mu\psi_1 + B\psi_1). \tag{2.4}$$

We could have also taken $\psi \mapsto \psi_2 = U_2\psi$, then this should also satisfy the same equation meaning

$$\begin{aligned}U_1^{-1}(\partial_\mu\psi_1 + B_1\psi_1) &= \partial_\mu\psi \\ U_1^{-1}(\partial_\mu\psi_1 + B_1\psi_1) &= U_2^{-1}(\partial_\mu\psi_2 + B_2\psi_2) \\ U_2U_1^{-1}(\partial_\mu\psi_1 + B_1\psi_1) &= (\partial_\mu\psi_2 + B_2\psi_2).\end{aligned}$$

First we see that $\psi_2 = U_2\psi = U_2U_1^{-1}\psi_1$. Now introducing general notation, $\psi_1 = \psi$, $\psi_2 = \psi'$, $B_1 = B_\mu$, $B_2 = B'_\mu$, $U_2U_1^{-1} = S^{-1}$, then we can therefore write $\psi' = S^{-1}\psi$

$$S^{-1}(\partial_\mu\psi + B_\mu\psi) = (\partial_\mu\psi' + B'_\mu\psi').$$

We can do a very similar computation with the conjugate, this gives

$$(\partial_\mu \psi^* + B\psi^*)S = (\partial_\mu \psi'^* + B'\psi'^*).$$

We introduce the covariant derivative $D_\mu = (\partial_\mu + B_\mu)$, where $D'_\mu = (\partial_\mu + B'_\mu)$ we have

$$S^{-1}D_\mu\psi = D'_\mu\psi' = (D_\mu\psi)', \tag{2.5}$$
$$D_\mu(\psi^*)S = D'_\mu\psi'^* = (D_\mu\psi^*)' \tag{2.6}$$

(finding the $\psi^*$ case can be done by following the same steps). With this we rewrite the Lagrangian in the following way:

$$\mathcal{L} = D_\mu\psi^* D_\mu\psi - m^2\psi^*\psi. \tag{2.7}$$

Where we have the transformations from before

$$\psi \to \psi' = S^{-1}\psi$$
$$D_\mu\psi \to (D_\mu\psi)' = S^{-1}D_\mu\psi,$$

where $S$ is a unitary matrix. We can immediately see that this Lagrangian is in fact invariant.
We could also write $B'$ in terms of $B$ in the following way

$$\begin{aligned}
S^{-1}(\partial_\mu\psi + B_\mu\psi) &= (\partial_\mu\psi' + B'_\mu\psi') \\
S^{-1}(\partial_\mu\psi + B_\mu\psi) &= (\partial_\mu S^{-1}\psi + B'_\mu S^{-1}\psi) \\
S^{-1}\partial_\mu\psi + S^{-1}B_\mu\psi &= \partial_\mu(S^{-1})\psi + S^{-1}\partial_\mu\psi + B'_\mu S^{-1}\psi \\
S^{-1}B_\mu\psi &= \partial_\mu(S^{-1})\psi + B'_\mu S^{-1}\psi \\
S^{-1}B_\mu &= \partial_\mu(S^{-1}) + B'_\mu S^{-1} \\
B'_\mu &= S^{-1}B_\mu S - \partial_\mu(S^{-1})S \\
B'_\mu &= S^{-1}B_\mu S + S^{-1}\partial_\mu(S),
\end{aligned}$$

where in the last equality we use that $0 = \partial_\mu I = \partial_\mu(S^{-1}S) = \partial(S^{-1})S + S^{-1}\partial_\mu(S)$. Meaning that $-\partial_\mu(S^{-1})S = S^{-1}\partial_\mu(S)$.
We will consider this property ($B'_\mu = S^{-1}B_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}$) the gauge invariance property. Here $B = B_\mu dx_\mu$ is the potential 1-form. Similarly to before it would be good if there is an $F$ which is gauge invariant. Let us check what the previous choice for $F$ was:

$$F = dB = \frac{\partial B_\mu}{\partial x_\nu} dx_\nu \wedge dx_\mu. \tag{2.8}$$

Now see how $F' = dB'$ transforms

$$\begin{aligned}
\frac{\partial B'_\mu}{\partial x_\nu} dx_\nu \wedge dx_\mu &= \frac{\partial}{\partial x_\nu}\left(S^{-1}B_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}\right) dx_\nu \wedge dx_\mu \\
&= \left(\frac{\partial S^{-1}}{\partial x_\nu}B_\mu S + S^{-1}\frac{\partial B_\mu}{\partial x_\nu}S + S^{-1}B_\mu\frac{\partial S}{\partial x_\nu} + \frac{\partial S^{-1}}{\partial x_\nu}\frac{\partial S}{\partial x_\mu} + S^{-1}\frac{\partial}{\partial x_\nu}\frac{\partial S}{\partial x_\mu}\right) dx_\nu \wedge dx_\mu.
\end{aligned}$$

The final term is symmetric, and since the wedge product is anti-symmetric, it cancels that term out, the second term is equal to $S^{-1}FS$. Seeing where this is going also take out $S^{-1}$ on the left and $S$ on the right. Leaving us with:

$$dB' = S^{-1}dBS + S^{-1}\left(S\frac{\partial S^{-1}}{\partial x_\nu}B_\mu + B_\mu\frac{\partial S}{\partial x_\nu}S^{-1} + S\frac{\partial S^{-1}}{\partial x_\nu}\frac{\partial S}{\partial x_\mu}S^{-1}\right)Sdx_\nu \wedge dx_\mu. \tag{2.9}$$

Here we see some kind of symmetry forming, to show this more clearly we restrict to the case where $\nu < \mu$, we have all terms again with switched indices

$$dB' = S^{-1}dBS + S^{-1}\Bigg(\Big(-\frac{\partial S}{\partial x_\nu}S^{-1}B_\mu + B_\mu\frac{\partial S}{\partial x_\nu}S^{-1} - \frac{\partial S}{\partial x_\nu}S^{-1}\frac{\partial S}{\partial x_\mu}S^{-1}\Big) - \Big(-\frac{\partial S}{\partial x_\mu}S^{-1}B_\nu + B_\nu\frac{\partial S}{\partial x_\mu}S^{-1} - \frac{\partial S}{\partial x_\mu}S^{-1}\frac{\partial S}{\partial x_\nu}S^{-1}\Big)\Bigg)Sdx_\nu \wedge dx_\mu.$$

Reordering the terms to show the result clearly:

$$dB' = S^{-1}dBS + S^{-1}\Bigg(\Big(B_\mu\frac{\partial S}{\partial x_\nu}S^{-1} + \frac{\partial S}{\partial x_\mu}S^{-1}B_\nu + \frac{\partial S}{\partial x_\mu}S^{-1}\frac{\partial S}{\partial x_\nu}S^{-1}\Big) - \Big(B_\nu\frac{\partial S}{\partial x_\mu}S^{-1} + \frac{\partial S}{\partial x_\nu}S^{-1}B_\mu + \frac{\partial S}{\partial x_\nu}S^{-1}\frac{\partial S}{\partial x_\mu}S^{-1}\Big)\Bigg)Sdx_\nu \wedge dx_\mu.$$

Then we see that we can factor these into:

$$dB' = S^{-1}dBS + S^{-1}\Bigg(\Big(B_\mu + \frac{\partial S}{\partial x_\mu}S^{-1}\Big)\Big(B_\nu + \frac{\partial S}{\partial x_\nu}S^{-1}\Big) - B_\mu B_\nu - \Big(B_\nu + \frac{\partial S}{\partial x_\nu}S^{-1}\Big)\Big(B_\mu + \frac{\partial S}{\partial x_\mu}S^{-1}\Big) + B_\nu B_\mu\Bigg)Sdx_\nu \wedge dx_\mu.$$

Bringing back the $S, S^{-1}$ in the brackets we find a familiar equation:

$$\begin{aligned} dB' &= S^{-1}dBS + \Bigg(S^{-1}\Big(B_\mu + \frac{\partial S}{\partial x_\mu}S^{-1}\Big)SS^{-1}\Big(B_\nu + \frac{\partial S}{\partial x_\nu}S^{-1}\Big)S - S^{-1}B_\mu B_\nu S \\ &\quad - S^{-1}\Big(B_\nu + \frac{\partial S}{\partial x_\nu}S^{-1}\Big)SS^{-1}\Big(B_\mu + \frac{\partial S}{\partial x_\mu}S^{-1}\Big)S + S^{-1}B_\nu B_\mu S\Bigg)dx_\nu \wedge dx_\mu \\ &= S^{-1}dB_\mu S + \Bigg(\Big(S^{-1}B_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}\Big)\Big(S^{-1}B_\nu S + S^{-1}\frac{\partial S}{\partial x_\nu}\Big) - S^{-1}B_\mu B_\nu S \\ &\quad - \Big(S^{-1}B_\nu S + S^{-1}\frac{\partial S}{\partial x_\nu}\Big)\Big(S^{-1}B_\mu S + S^{-1}\frac{\partial S}{\partial x_\mu}\Big) + S^{-1}B_\nu B_\mu S\Bigg)dx_\nu \wedge dx_\mu. \end{aligned}$$

In the brackets there is a term that we have seen earlier, namely $B'$, also note that we have inserted an identity matrix in between the brackets

$$dB' = S^{-1}dBS + \big(B'_\mu B'_\nu - S^{-1}B_\mu B_\nu S - B'_\nu B'_\mu + S^{-1}B_\nu B_\mu S\big)dx_\nu \wedge dx_\mu.$$

Now we see the commutators of $B$ and $B'$

$$dB' = S^{-1}dBS + ([B'_\mu, B'_\nu] - S^{-1}[B_\mu, B_\nu]S)dx_\nu \wedge dx_\mu.$$

Swapping the wedge product we find

$$dB' + [B'_\mu, B'_\nu]dx_\mu \wedge dx_\nu = S^{-1}\big(dB + [B_\mu, B_\nu]dx_\mu \wedge dx_\nu\big)S. \tag{2.10}$$

Redefining

$$F := dB + [B_\mu, B_\nu]dx_\mu \wedge dx_\nu \qquad \nu < \mu, \tag{2.11}$$

we immediately get the result $F' = S^{-1}FS$. This $F$ is called the field strength. Important here is to know that this is ordered. So there would be $1/2$ if we decide not to order it.

In the electromagnetic case we have $B = A$, for this potential the elements commute, therefore we get $F = dA + [A_\mu, A_\nu]dx_\mu \wedge dx_\nu = dA$, exactly as earlier.

We also find the following motivating why we defined the covariant derivative earlier:

$$\begin{aligned}[D_\mu, D_\nu] &= [\partial_\mu + B_\mu, \partial_\nu + B_\nu] \\ &= (\partial_\mu + B_\mu)(\partial_\nu + B_\nu) - (\partial_\nu + B_\nu)(\partial_\mu + B_\mu) \\ &= \partial_\mu\partial_\nu + \partial_\mu B_\nu + B_\mu\partial_\nu + B_\mu B_\nu - \partial_\nu\partial_\mu - \partial_\nu B_\mu - B_\nu\partial_\mu - B_\nu B_\mu.\end{aligned}$$

Since we are acting on a constant function, the derivatives on the right dissapear and leave us with

$$[D_\mu, D_\nu] = \partial_\mu B_\nu + B_\mu B_\nu - \partial_\nu B_\mu - B_\nu B_\mu.$$

This leads to

$$\begin{aligned}[D_\mu, D_\nu]dx_\mu \wedge dx_\nu &= (\partial_\mu B_\nu - \partial_\nu B_\mu)dx_\mu \wedge dx_\nu + (B_\mu B_\nu - B_\nu B_\mu)dx_\mu \wedge dx_\nu \\ &= dB + [B_\mu, B_\nu]dx_\mu \wedge dx_\nu \\ &= F,\end{aligned}$$

where we have ordered $\mu$ and $\nu$. This also shows us to write $F = F_{\mu\nu}dx_\mu \wedge dx_\nu$, where we have that

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu + [B_\mu, B_\nu]. \tag{2.12}$$

### 2.2.2 Yang-Mills

The idea of Gauge theory is that any term that is gauge invariant should be a term in the Lagrangian, as it is a physical property. Therefore an important part of Gauge theory is finding such terms. We have seen that $F_{\mu\nu} \mapsto F'_{\mu\nu} = S^{-1}F_{\mu\nu}S$, in the Lagrangian we want things to be Lorentz invariant as well, this naturally leads us to consider $F_{\mu\nu}F_{\mu\nu} \mapsto F'_{\mu\nu}F'_{\mu\nu} = S^{-1}F_{\mu\nu}SS^{-1}F_{\mu\nu}S = S^{-1}F_{\mu\nu}F_{\mu\nu}S$, this is nearly gauge invariant. We would need a property that allows us to cancel $S$ and $S^{-1}$. A function that has this exact property is the trace: $\mathrm{tr}(F_{\mu\nu}F_{\mu\nu}) \mapsto \mathrm{tr}\big(F'_{\mu\nu}F'_{\mu\nu}\big) = \mathrm{tr}(S^{-1}F_{\mu\nu}F_{\mu\nu}S) = \mathrm{tr}(F_{\mu\nu}F_{\mu\nu}SS^{-1}) = \mathrm{tr}(F_{\mu\nu}F_{\mu\nu})$ So this is gauge invariant! Since any $\mathrm{tr}(F_{\mu\nu}F_{\mu\nu})$ multiplied with a constant would

then be gauge invariant we will call these equivalent and therefore choose one specific multiplication namely

$$\mathcal{L} = \frac{1}{2g^2}\,\mathrm{tr}(F_{\mu\nu}F_{\mu\nu}), \tag{2.13}$$

where $g$ is any positive constant.

Now assume $A_\mu$ takes values in the Lie algebra of $SU(2)$, as is the case for the original Yang-Mills theory, the following can easily be generalised to any Lie group. We can write $A_\mu = -ig\frac{\tau^a}{2}A^a_\mu$, where $\frac{\tau^a}{2}$ are the generators of the $SU(2)$ algebra, and $A^a_\mu$ scalar fields. $a = 1, 2, 3$ since the $SU(2)$ algebra has 3 generators. $-g$ is added by convention, $i$ to make $\frac{\tau^a}{2}$ hermitian (instead of skew-hermitian, since the generators of $SU(2)$ algebra are skew-hermitian). With this $F_{\mu\nu}$ becomes

$$\begin{aligned}
F_{\mu\nu} &= \partial_\nu A_\mu - \partial_\mu A_\nu + [A_\mu, A_\nu] \\
&= -ig\frac{\tau^a}{2}(\partial_\nu A^a_\mu - \partial_\mu A^a_\nu) + (-ig)^2 A^a_\mu A^b_\nu [\frac{\tau^a}{2}, \frac{\tau^b}{2}] \\
&= -ig\frac{\tau^a}{2}(\partial_\nu A^a_\mu - \partial_\mu A^a_\nu) - g^2 A^a_\mu A^b_\nu i\epsilon^{abc}\frac{\tau^c}{2} \\
&= -ig\frac{\tau^a}{2}(\partial_\nu A^a_\mu - \partial_\mu A^a_\nu + g\epsilon^{abc}A^b_\mu A^c_\nu) \\
&= -ig\frac{\tau^a}{2}F^a_{\mu\nu}.
\end{aligned}$$

Where we have defined $F^a_{\mu\nu} = (\partial_\nu A^a_\mu - \partial_\mu A^a_\nu + g\epsilon^{abc}A^b_\mu A^c_\nu)$. Note that we use the property $[\frac{\tau^a}{2}, \frac{\tau^b}{2}] = i\epsilon^{abc}\frac{\tau^c}{2}$ Then substituting this into the Lagrangian, we obtain

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2g^2}tr(F_{\mu\nu}F_{\mu\nu}) \\
&= \frac{1}{2g^2}tr[(-ig\frac{\tau^a}{2}F^a_{\mu\nu})(-ig\frac{\tau^b}{2}F^b_{\mu\nu})] \\
&= -\frac{1}{2}F^a_{\mu\nu}F^b_{\mu\nu}tr(\frac{\tau^a}{2}\frac{\tau^b}{2}) \\
&= -\frac{1}{2}F^a_{\mu\nu}F^a_{\mu\nu}tr(\frac{\tau^a}{2}\frac{\tau^a}{2}) \\
&= -\frac{1}{4}F^a_{\mu\nu}F^a_{\mu\nu}.
\end{aligned} \tag{2.14}$$

# Chapter 3

# Connection

Having introduced a physics motivation for Gauge theory we will now look at the mathematics approach. This chapter gives the necessary information to understand the complete Gauge theory explained in chapter 4.

## 3.1 Principal fibre bundles

The given definitions are taken from the book of Eguchi, Gilkey and Hanson (7).

**Definition 3.1.1** (Fibre bundle)**.** *The topological space $E$ over a base manifold $M$ with fibre $F$ and topological projection $\pi : E \to M$ s.t. $\forall x \in M$ $\exists$ neighbourhood $U_i$ of $x$ and $\phi_i : U_i \times F \to \pi^{-1}(U_i) \subset E$ an isomorphism is called a **fibre bundle**. We require that $\pi(\phi_i(x, f)) = x$ where $(x, f) \in U_i \times F$.*
*Transition functions are defined on $U_i \cap U_j$ as $\phi_{ij} = \phi_i^{-1}\phi_j$. These transition functions are required to belong to a group $G$ of transformations of the fibre space. $G$ is called the structure group of the fibre bundle.*

Now a more intuitive explanation of a fibre bundle, at every position in space (the base manifold) there are properties that space can have there, for example the probability of a particle being at that location, or the strength of the magnetic field (or any field), it could also give the flow of a system. Basically any physical property that something can have can be put in the fibre, these are all the values that the space can take. What the transition functions do is communicate between two regions in space. To make this useful in a physical setting is by use of a section, a section picks the values of the fibre at every point in space. And therefore defines a physical situation. An example of how the transition functions work: $\phi_{ij}(x, f) = \phi_i^{-1}\phi_j(x, f) = (x, f')$.
Now some examples of fibre bundles: Vector bundle is a fibre bundle where the structure group belongs to $GL(k, \mathbb{R})$. Which gives a vector field as the fibre, for example the electric field. The tangent bundle on $M$ has fibres $T_xM$ with transition functions $\frac{\partial x'}{\partial x}$. Which gives a flow/direction field as the fibre, this could then be the flow of a particle in some fluid. Finally the cotangent bundle on $M$ has fibres $T_x^*M$. Like we said, these fibre can be seen as for example an electric field. Then it would be logical to also consider assigning a specific value in the fibre at each point. This makes us able to describe what the electric field looks like at a specific point. This is what a section does.

**Definition 3.1.2** (Section)**.** *A* ***section*** $\sigma_i : U_i \to F$ *is a map that gives a preference for each element of the base manifold on the fibre. The set of sections of a fibre bundle* $E$ *is usually denoted as* $\Gamma(E)$*.*

**Definition 3.1.3** (Principal bundle)**.** *The* ***principal bundle*** $P$ *is a fibre bundle whose structure group is a Lie group.*

It is also possible to give a complete definition of a principal fibre bundle without first defining fibre bundles and sections. But this is completely analogous to the previous statement. See Bleeckers (2)

**Definition 3.1.4** (Principal fibre bundle (self-contained))**.** *A* ***principal fibre bundle*** *consists of a total manifold* $P$*, a Lie group* $G$*, a base manifold* $M$ *and a projection* $\pi : P \to M$ *s.t. the following statements hold:*

1. *For each* $g \in G$ *there is a homeomorphism* $R_g : P \to P$ *where* $R_g(p) = pg$ *s.t.* $p(g_1 g_2) = (pg_1)g_2 \ \forall g_1, g_2 \in G$*;* $e \in G$ *is the identity element iff* $pe = p$ $\forall p$*. The function* $P \times G \to P$ *given by* $(p, g) \mapsto pg$ *is a map.*

2. *The map* $\pi : P \to M$ *is onto, and* $\pi^{-1}(\pi(p)) = \{pg : g \in G\}$*, this is called the* ***orbit*** *of* $G$ *through* $p$*. If* $x \in M$ *then* $\pi^{-1}(x)$ *is called the* ***fibre above*** $x$*.*

3. *For each* $x \in M$ *there exists an open neighbourhood* $U_i$ *and a diffeomorphism* $T_i : \pi^{-1}(U_i) \to U_i \times G$ *of the form* $T_i(p) = (\pi(p), s_i(p))$ *where* $s_i : \pi^{-1}(U_i) \to G$ *has the property that* $s_i(pg) = s_i(p)g \ \forall g \in G, p \in \pi^{-1}(U_i)$*.* $T_i$ *is called the* ***local trivialisation*** *or the* ***choice of gauge****.*

Relating this to the previous definition, $T_i = \phi_i^{-1}$, is well-defined since $T_i$ is a diffeomorphism. And $s_i(p) = \sigma_i(\pi(p))$. Similarly to before we will also need to define the transition functions

**Definition 3.1.5** (Transition function)**.** *Consider* $T_i : \pi^{-1}(U_i) \to U_i \times G$ *and* $T_j : \pi^{-1}(U_j) \to U_j \times G$ *two local trivialisations of a principle fibre bundle* $\pi : P \to M$ *with lie group* $G$*. The* ***transition function*** *from* $T_i$ *to* $T_j$ *is defined by the map* $g_{ij} : U_i \cap U_j$ *s.t. for* $x = \pi(p) \subset U_i \cap U_j$ $g_{ij}(x) = s_i(p)s_j^{-1}(p)$*. We have*

1. $g_{ii}(y) = e$ *for all* $y \in U_i$

2. $g_{ji}(y) = g_{ij}^{-1}(y)$ *for all* $y \in U_i \cap U_j$

3. $g_{ij}(y)g_{jk}(y)g_{ki}(y) = e$ *for all* $y \in U_i \cap U_j \cap U_k$

## 3.2 Connection

For this part the book of Sontz (5) was used.

### 3.2.1 Splitting the bundle

In this section we will first get an intuitive idea of what a connection is, after a formal definition will be given.

Consider the tangent space of principal bundle $P$, $T_pP$ to consist of two sets of directions. One in the direction of the base manifold $M$ called the horizontal direction, and one in the direction of the Lie group $G$ the vertical direction. This makes it possible to say how an element of $P$ changes over the manifold $M$ and over the group $G$ separately. The way $p \in P$ can change over $M$ is called the connection.
As an example consider $P = M \times G$, with $M = \mathbb{R}$ and $G = U(1)$, then we get a cylinder P, with are a lot of circles ($U(1)$) on top of the real line ($\mathbb{R}$) (or a lot of lines on a circle, pick which you like more). Then the vertical and horizontal space can be seen as in figure 3.1.

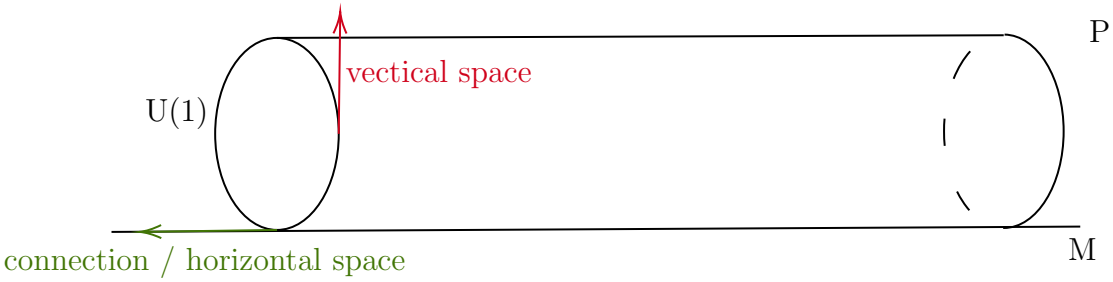


Figure 3.1: $P = \mathbb{R} \times U(1)$ to show the horizontal and vertical direction.

Now for the formal definition, consider the map

$$i_p : G \to P \quad \text{by} \quad i_p(g) = pg \quad \text{for each } p \in P, \tag{3.1}$$

where $P$ is the principal bundle and $G$ its fibre. This map is the unique $G$-equivariant map.

Then look at the push forward of $i_p$ evaluated at $e$,

$$i_{p*}|_e : T_eG \to T_{i_p(e)}P = T_pP. \tag{3.2}$$

This map therefore projects the directions of $T_eG$ onto $T_pP$, the image $\mathrm{Im}(i_{p*}|_e) =: \mathrm{Vert}(T_pP)$ is called the vertical direction. The horizontal direction($\mathrm{Hor}(T_pP)$) "completes" this vertical direction, to make the space $T_pP$ in the following way.

1. Any element in $T_pP$ can be written as a linear combination of $\mathrm{Vert}(T_pP)$ and $\mathrm{Hor}(T_pP)$. $\mathrm{Vert}(T_pP) \oplus \mathrm{Hor}(T_pP) = T_pP$.
2. $\mathrm{Vert}(T_pP) \cap \mathrm{Hor}(T_pP) = 0$.
3. $\mathrm{Hor}(T_pP)$ depends smoothly on $p \in P$.
4. $\mathrm{Hor}(T_pP)$ satisfies the $G$-invariance condition:

$$R_{g*}(\mathrm{Hor}(T_pP)) = \mathrm{Hor}(T_{pg}P) \quad \forall g \in G, p \in P. \tag{3.3}$$

### 3.2.2 Connection 1-form

It is often useful to describe this as a 1-form. To do this we need to define a few concepts.

**Definition 3.2.1** (W-valued 1-form)**.** *Let $M$ be a smooth manifold, $W$ a vector space and $\zeta : TM \to W$ smooth and linear when restricted to each fibre of $TM$. Then $\zeta$ is a* ***W-valued 1-form*** *on $M$. The set of all W-valued 1-forms on $M$ is $\Omega^1(M;W)$.*

First of a "normal" 1-form is an $\mathbb{R}$-valued 1-form, we often write 1-forms as $A_\mu dx_\mu$, where $A : M \to \mathbb{R}$. So in the case of a W-valued 1-form we have $A_\mu dx_\mu$, where $A : M \to W$. To make a connection in a more applicable form we need the following definitions.

**Definition 3.2.2** (Pullback function)**.** *Let $M$, $N$ be smooth manifolds, $\zeta$ a W-valued 1-form on $N$ and $f : M \to N$. Then the* ***pullback*** *of $\zeta$ by $f$ is*

$$f^*(\zeta) := \zeta \circ f_* : TM \to W. \tag{3.4}$$

**Definition 3.2.3** (Adjoint)**.** *Let $G$ be a Lie group, then define the* ***adjoint map****:*

$$ad_g : G \to G \quad where \quad ad_g(h) := ghg^{-1} \tag{3.5}$$

*for all $g \in G$. Note that this is equivalent to: $ad_g = L_g \circ R_{g^{-1}} = R_{g^{-1}} \circ L_g$, showing that the adjoint is a diffeomorphims.*
*Also define the adjoint representation of $G$ by the pushforward of $ad_g$ evaluated at $e$*

$$Ad_g := ad_{g*}|_e : T_eG \to T_eG. \tag{3.6}$$

**Definition 3.2.4** (Fundamental vector field)**.** *Take $A \in T_eG$, then we define $A^\#(p) := i_{p*}(A) \in Vert(T_pP)$ the pushforward of $i_p$ at $A$. Then*

$$A^\# : P \to TP \quad p \mapsto A^\#(p), \tag{3.7}$$

*Then $A^\#$ i- the* ***fundamental vector field*** *on $P$ associated to $A$.*

Now we move to the motivation. Since we can write any element in $v \in T_pP$ as a horizontal part $v_H$ and a vertical part $v_V$ s.t. $v_H + v_V = v$. We can have a linear map $pr_p : T_pP \to \text{Vert}(T_pP)$ that maps $v = v_H + v_V \mapsto v_V$. We can also make the map $i_{g*} : T_eG \to \text{Vert}(T_pP)$, which is an isomorphism. Then this means that there exist a map

$$\omega_p := (i_{p*})^{-1} \circ pr_p : T_pP \to T_eG \quad \text{for } p \in P \tag{3.8}$$

Then we also have

$$\omega : TP \to T_eG \quad s.t. \quad \omega|T_pP := \omega_p. \tag{3.9}$$

So $\omega$ is a $T_eG$ valued 1-form on $P$.
Then one can show that

$$R_g^*\omega = Ad_{g^{-1}} \circ \omega. \tag{3.10}$$

holds for every $g \in G$. (the proof is in the book of (5) theorem 10.1).
This motivates the following definition, showing the relation between the connection and the connection 1-form.

**Definition 3.2.5** (connection 1-form)**.** *The* ***connection 1-form*** *on principal bundle $P$, is a $T_eG$ valued 1-form $\omega$ defined as before on $P$, that satisfies the equivariance condition*

$$R_g^*\omega = Ad_{g^{-1}} \circ \omega \quad \forall g \in G \tag{3.11}$$

*as well as*

$$\omega(A^{\#}) = A \quad \forall A \in T_eG. \tag{3.12}$$

We will want consider the 1-form connection on different principal fibre bundles. Since the elements in a $T_eG$ valued 1-form are in the tangent space of the principal fibre bundle $TG$ and the basis elements are in $TM$, then it makes sense for us to move from one form to another by use of the pullback function which maps the tangent bundle of one principal bundle to the tangent bundle of another principal bundle. This motivates the following theorem from the book of Sontz Theorem 10.2 (5) which also states the proof.

**Theorem 3.2.1** (Pullback connection)**.** *Take $P_1$ and $P_2$ two principal bundles with the same structure Lie group $G$ and base manifolds $M_1$ and $M_2$, respectively. Let $F : P_1 \to P_2$ a morphism of principal $G$-bundles covering a smooth map $f : M_1 \to M_2$. If $\omega$ is a connection 1-form on $P_2$ then $F^*\omega$ is a connection 1-form on $P_1$.*

### 3.2.3 Maurer-Cartan form

Now we define a very specific connection 1-form,

**Definition 3.2.6** (Maurer-Cartan form)**.** *Let $G$ be a Lie group, and $\mathcal{G}$ its Lie algebra, then there is a $\mathcal{G}$-valued 1-form on $G$, denoted by:*

$$\Omega : TG \to g \quad \Omega_g := (L_{g^{-1}})_* : T_gG \to g, \quad \forall g \in G. \tag{3.13}$$

*$\Omega$ is called the* ***Maurer-Cartan*** *form on $G$.*

We do this because really every connection 1-form is such a form, as shown by Sontz in Proposition 10.4 and Theorem 10.3 (5).

**Theorem 3.2.2** (Maurer-Cartan Uniqueness)**.** *Let $G$ be a Lie group. Then its Maurer-Cartan form $\Omega$ is the only connection 1-form in the principal bundle $\{x\} \times G$.*

## 3.3 Curvature

To relate our findings back to the physics part we define the curvature, which in physics actually has another name, namely the Field strength,

**Definition 3.3.1** (Curvature)**.** *Let $\omega$ be a connection 1-form, then the curvature $D\omega : TP \times TP \to g$ is*

$$\Omega^\omega := D\omega(v_1, v_2) := d\omega(v_1^H, v_2^H) \tag{3.14}$$

*$v_i^H$ is the horizontal part of the vector $v_i$.*
*If $\omega$ is a potential, then we call $\Omega^\omega$ the field strength.*

This definition however is not very practical, and since we want to relate this to physics it should be. Luckily there is a nice result relating this curvature 2-form to a 1-form without considering the horizontal space, it uses the fact that $\omega(v^H) = 0$, and that $D\omega(v_1^V, v_2^V) = 0$. The complete proof can be found in Sontz Theorem 11.2 (5). Here the statement

**Theorem 3.3.1** (The structure theorem)**.**

$$D\omega = d\omega + \frac{1}{2}\omega \wedge \omega. \tag{3.15}$$

This equation will motivate the relation between connections and gauge fields.

# Chapter 4

# Gauge theory

In this chapter we study Gauge theory from a more mathematical approach, first we consider the physics motivation. Seeing that our field strength has the same form as the structure theorem. Which motivates the search for a relation between our physical $\mathcal{G}$-valued 1-forms and connections. Ending the chapter by confirming that these two concepts are indeed related using the example of Yang-Mills. Most of this chapter was studied from Sontz book (5).

## 4.1 Relating Gauge theory to connection 1-forms

Consider a gauge field $B = B_\mu dx_\mu$ a 1-form, with a vector field $V = V_\mu dx_\mu$. The $B_\mu$ are complex $n \times n$ matrices, and $V_\mu$ are smooth real-valued functions on $\mathbb{R}^n$. Then we have $B(V) = V_\mu B_\mu$ which is an $n \times n$ matrix. Meaning

$$B : T\mathbb{R}^n \to M(n, \mathbb{C}), \quad B(V) := V_\mu B_\mu(x), \quad V \in T_x\mathbb{R}^n$$

This motivates the following definition:

**Definition 4.1.1** (Vector valued 1-form)**.** *Let $M$ be a manifold and $W$ a vector space over $\mathbb{R}$. Then a smooth function $A : TM \to W$ linear on every fibre of $TM$ is called a* ***vector-valued 1-form*** *on $M$. If $W$ is a Lie algebra, then $A$ is called* ***Lie algebra-valued 1-form*** *on $M$.*

But wait, we have seen this definition before! It is the same as the $W$-valued 1-form (definition 3.2.1). These are 1-forms, since we have previously defined $k$-forms, we also want to define $W$-valued $k$-forms. This is done as follows

**Definition 4.1.2** (Vector valued $k$-form)**.** *Let $M$ be a manifold and $W$ a vector space over $\mathbb{R}$. Then we call*

$$A : \underbrace{TM \times \cdots \times TM}_{k \text{ copies}} \to W, \tag{4.1}$$

*smooth, antisymmetric and $k$-multilinear on every fibre of $\underbrace{TM \times \cdots \times TM}_{k \text{ copies}}$. This is called a* ***vector valued $k$-form on*** *$M$. Moreover if $W$ is a Lie algebra, then $A$ is called a* ***Lie algebra valued k-form on M****.*

If $x_\mu$ are local coordinates on the open subset $V \subset M$, then in that chart we have,

$$\begin{aligned} A &= A_I dx_I \\ A_I &:= A \circ \partial_I : V \to W, \\ \partial_I &:= \partial_{i_1,\cdots,i_k} : V \to \underbrace{TV \times \cdots \times TV}_{k \text{ copies}} \subset \underbrace{TM \times \cdots \times TM}_{k \text{ copies}} \\ \partial_I(p) &:= \partial_{i_1,\cdots,i_k}(p) := (\partial_{i_1}(p), \cdots, \partial_{i_k}(p)) \\ dx_I &= dx_{i_1,\cdots,i_k} := dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \end{aligned}$$

Here $I$ is $k$-indexed and strictly increasing. Also defining the exterior derivative for a $W$-valued $k$-form as,

$$\begin{aligned} dA =& \partial_\alpha A_I dx_\alpha \wedge dx_I \\ & \partial_\alpha A_I : V \to W. \end{aligned}$$

Using all the definitions above we look at the following:

$$dB = (\partial_\nu B_\mu - \partial_\mu B_\nu) dx_\nu \wedge dx_\mu,$$

where $B$ is a Lie algebra valued 1-form. We also find

$$(B \wedge B)(v, w) = (B_\mu(v) B_\nu(w) - B_\nu(v) B_\mu(w)) dx_\mu \wedge dx_\nu.$$

Ordering this and seeing that this is a commutator we find:

$$B \wedge B = 2[B_\mu, B_\nu] dx_\mu \wedge dx_\nu.$$

(So $\mu < \nu$). We found that the field strength $F$ is (cf. equation 2.11)

$$\begin{aligned} F &= dB + [B_\mu, B_\nu] dx_\mu \wedge dx_\nu \\ &= dB + \frac{1}{2}(B \wedge B). \end{aligned} \tag{4.2}$$

When seeing this formula and looking back at chapter 3 we see that this is exactly the same term as in The structure theorem 3.3.1! So is $B$ a connection 1-form? First off, $B$ is a $\mathcal{G}$-valued 1-form, but we are not yet considering the setting in a principal bundle. So it is time to do so now. Since we are already in the manifold $M$, we consider this to be the base manifold. Similarly we already have a Lie algebra, to which a Lie group is related, calling this $G$ we can take $P := M \times G$ to be our principal bundle, this is a trivial principal bundle. Then $\pi : P \to M$ the projection map $\pi_1 : P \to M$, $\pi(x, h) = \pi_1(x, h) = x$, for $(x, h) \in P, x \in M$ and $h \in G$. We will also write $(x, h_1)h_2 = (x, h_1 h_2)$, for $h_1, h_2 \in G$. Given that $B$ is a $\mathcal{G}$-valued 1-form on M then $\pi_1^* B$ is a $\mathcal{G}$-valued 1-form on P (that is by the theorem under connection 1-forms). So the question becomes is $\pi_1^* B$ a connection 1-form on $P$?
For that we need to check the two properties of a connection 1-form: For the first property recall from definition 3.2.5 we need to check that $R_g^* \omega = Ad_{g^{-1}} \omega$, we have that

$$R_g^* \pi_1^* B = (\pi_1 R_g)^* B = (id)^* B = 0.$$

Then we must have that $Ad_g^{-1}\pi^*B = 0$, so $g^{-1}\pi_1^*Bg$ is constant. But this is not true in general, so the first condition is not satisfied. Still let us also consider the second condition. Consider $A \in g$ then $A^\# \in \text{Vert}(TP)$, note that $\ker \pi_{1*} = TG = \text{Vert}(TP)$ (since $P$ is a principal bundle). That means

$$\pi_1^*B(A^\#) = B \circ \pi_{1*}(A^\#) = B(0) = 0 \neq A. \tag{4.3}$$

So then, is there any way the two concepts are related? To see this we consider the following.
We want to find a connection 1-form $\omega$ on a trivial principal bundle $\pi : P \to M$ with given section $\sigma : M \to P$ s.t. $\sigma^*\omega = B$, with $B$ a given $\mathcal{G}$-valued 1-form on $M$.

## 4.2 Connection 1-form on a trivial section

To simplify the problem we take a specific section $\sigma_1(x) = (x, e)$. Thus $\sigma_1$ embeds $M$ onto its range $M \times e$. Then we see

$$\pi_1^*B = \pi_1^*\sigma_1^*\omega = \omega. \tag{4.4}$$

Note here that

$$\pi_1^*\sigma_1^*\omega = \pi_1^*\omega \circ \sigma_{1*} = \omega \circ \sigma_{1*} \circ \pi_{1*} = \omega \circ (\sigma_1\pi_1)_* = (\sigma_1\pi_1)^*\omega = \omega$$

Since $\sigma_1\pi_1$ is the identity map.

If we want $\omega$ to be a connection 1-form then we must have:

$$Ad_{h^{-1}}\omega = R_h^*\omega = \omega R_{h*}. \tag{4.5}$$

We will look at slices of $P$, $M \times \{h\}$. Note that we have that $\pi_1 R_h = \pi_1$. So we will rewrite this condition to define $\omega$

$$\begin{aligned}
\omega &= Ad_{h^{-1}}\pi_1^*BR_{h^{-1}*} \\
&= Ad_{h^{-1}}B\pi_{1*}R_{h^{-1}*} \\
&= Ad_{h^{-1}}B\pi_{1*} \\
&= Ad_{h^{-1}}\pi_1^*B.
\end{aligned}$$

We would like this to hold over all of $P = M \times G$, we define $g : P = M \times G \to G$ by $g(x, h) = h$. Following the derivation from earlier and using this new $g$ we define

$$\omega_1 = Ad_{g^{-1}}\pi_1^*(B). \tag{4.6}$$

We see that $\sigma_1^*\omega_1 = Ad_{g^{-1}(x,e)}B = B$, where $g^{-1}(x, h) = h^{-1}$ by convention.
Let's look if the first connection 1-form requirement is satisfied, $v \in T_pP$.

$$\begin{aligned}
(R_h^*\omega_1)(v) &= \omega(R_{h*}v) \\
&= Ad_{g^{-1}(ph)}(\pi_1^*B)(R_{h*}v) \\
&= Ad_{h^{-1}g^{-1}(p)}(B\pi_{1*}R_{h*}v) \\
&= Ad_{h^{-1}}(\pi_1^*Bv) \\
&= Ad_{h^{-1}}\omega_1(v).
\end{aligned}$$

So it satisfies the first condition for a connection 1-form. Now let's look at the second condition. Let $A \in g$ and $A^{\#}$ be its fundamental vector field, then we get

$$\omega_1(A^{\#}) = Ad_{g^{-1}}(\pi_1^* B)(A^{\#}) = Ad_{g^{-1}} B \pi_{1*} A^{\#} = 0. \tag{4.7}$$

Since $\pi_{1*} A^{\#} = 0$ as shown before.
This $\omega_1$ is not a connection 1-form, but it is close. It is actually called a connection 1-form displacement, which we will now formally define:

**Definition 4.2.1** (Connection 1-form displacement)**.** *A $\mathcal{G}$-valued 1-form $\eta$ is called a* ***connection 1-form displacement****, if it satisfies*

1. $Ad_{h^{-1}} \eta = R_h^* \eta$.
2. $\eta(A^{\#}) = 0 \quad \forall A \in g$.

This connection 1-form displacement has some properties related to the connection 1-form that are interesting for us to find the connection 1-form that we are looking for.

**Theorem 4.2.1.**

1. *Let $\omega_1, \omega_2$ be two connection 1-forms, then $\omega_1 - \omega_2 = \eta$ is a connection 1-form displacement.*
2. *Let $\omega_1$ be a connection 1-form and $\eta$ a connection 1-form displacement. Then $\omega_1 + \eta = \omega_2$ is a connection 1-form.*
3. *Take some $\omega_1$ a connection 1-form, then $\forall \omega_2$ a connection 1-form, there exists an $\eta$ a connection 1-form displacement s.t. $\omega_1 + \eta = \omega_2$.*

*Proof.* First proof of 1) Take $\omega_1$ and $\omega_2$ two connection 1-forms, we have that

$$Ad_{h^{-1}}(\omega_1 - \omega_2) = Ad_{h^{-1}}\omega_1 - Ad_{h^{-1}}\omega_2 = R_h^*\omega_1 - R_h^*\omega_2 = R_h^*(\omega_1 - \omega_2),$$

so $\omega_1 - \omega_2$ satisfies the first property of a connection 1-form displacement. Let's look at the second one take $A^{\#}$ a fundamental vector field on P associated to A. Then

$$(\omega_1 - \omega_2)(A^{\#}) = \omega_1(A^{\#}) - \omega_2(A^{\#}) = A - A = 0,$$

so $\eta = \omega_1 - \omega_2$ is a connection 1-form displacement.
Now let's look at statement 2:
Take $\omega_1$ a connection 1-form and $\eta$ a connection 1-form displacement, we have that

$$Ad_{h^{-1}}(\omega_1 + \eta) = Ad_{h^{-1}}\omega_1 + Ad_{h^{-1}}\eta = R_h^*\omega_1 + R_h^*\eta = R_h^*(\omega_1 + \eta),$$

so $\omega_1 + \eta$ satisfies the first property of a connection 1-form. Let's look at the second one take $A^{\#}$ a fundamental vector field on P associated to A. Then

$$(\omega_1 + \eta)(A^{\#}) = \omega_1(A^{\#}) + \eta(A^{\#}) = A + 0 = A,$$

so $\omega_2 = \omega_1 + \eta$ is a connection 1-form.
Now for the final statement:
Take some $\omega_1$ and some $\omega_2$, then from statement 1 we have that $\omega_2 - \omega_1 = \eta$ is a 1-form displacement. Reordering these terms we get $\omega_1 + \eta = \omega_2$, meaning that such an $\eta$ always exists and completing the proof. □

Knowing this, we can take any connection 1-form $\alpha$ and our original $\omega_1 = Ad_{g^{-1}}\pi_1^*(B)$, and define

$$\omega := \omega_1 + \alpha. \tag{4.8}$$

For $\omega$ we have the requirement

$$\sigma_1^*(\omega) = B.$$

Computing this gives

$$\begin{aligned} \sigma_1^*(\omega_1 + \alpha) &= B, \\ \sigma_1^*(\omega_1) + \sigma_1^*(\alpha) &= B, \\ B + \sigma_1^*(\alpha) &= B, \\ \sigma_1^*(\alpha) &= 0. \end{aligned}$$

Since $\sigma_1(x) = (x, e)$ we have that $\sigma_{1*}(v) = (v, 0)$ for $v \in TM$. That means that $\alpha(v, 0) = \alpha\sigma_{1*}(v) = 0$ by our requirement. So $\alpha$ is completely determined by $TG$. In chapter 3 we considered the map $i : G \to P = M \times G$, s.t. $i_x(h) = (x, h)$. Then the maps $i_{x*}(\omega) = (0, \omega)$ with $\omega \in TG$. Consider the map

$$\beta = \alpha i_{x*} = i_x^*\alpha.$$

Since $\alpha$ is a connection 1-form on P, it follows that $\beta$ is a connection 1-form on $G$ by theorem 3.2.1. But there is only one connection 1-form on $G$ by theorem 3.2.2 and that is the Maurer-carter form $\Omega$. So

$$\Omega = \beta = i_x^*\alpha. \tag{4.9}$$

But we really want to find an equation for $\alpha$, so let's try

$$\pi_2^*\Omega = \pi_2^* i_x^*\alpha = (i_x\pi_2)^*\alpha,$$

here $\pi_2$ is the projection onto G. We have $i_x\pi_2(y, h) = i_x(h) = (x, h)$, then $(i_x\pi_2)_*(v, w) = (0, w)$ for $v \in TM, w \in TG$.

$$(i_x\pi_2)^*\alpha(v, w) = \alpha(0, w) = \alpha(v, w).$$

Then

$$\alpha = \pi_2^*\Omega, \tag{4.10}$$

where we have that $\sigma_1^*(\pi_2^*\Omega) = (\pi_2\sigma_1)^*\Omega$, $\pi_2\sigma_1(x) = \pi_2(x, e) = e$, so $(\pi_2\sigma_1)^*$ is the zero map. Finally this gives us an equation that relates $B$ to an connection 1-form in the following way:

$$\omega := Ad_{g^{-1}}\pi_1^*(B) + \pi_2^*\Omega. \tag{4.11}$$

Note that this was for a trivial section. In the next part we will look at how this changes when not taking the trivial section.

## 4.3 Connection 1-form on any section

Motivated by the previous section we state the following theorem:

**Theorem 4.3.1** (Unique connection 1-form)**.** *Suppose $\pi : P \to M$ is a principal $G$-bundle, with Lie group $G$. Take $\sigma : M \to P$ a section on this bundle and $B$ a $\mathcal{G}$-valued 1-form on $M$, where $\mathcal{G} = T_eG$ the Lie algebra of $G$.*
*Then there exists a unique connection 1-form $\omega : TP \to \mathcal{G}$, s.t. $B = \sigma^*\omega$. Moreover the explicit formula for $\omega$ is*

$$\omega = Ad_{g^{-1}}\pi^*(B) + g^*(\Omega), \tag{4.12}$$

*where $\Omega$ is the Maurer-Carter form on $G$ and $g : P \to G$ a map defined in the proof.*

*Proof.* The idea of the proof is to end up at the section we had before. To do this we take $p \in P$, $x = \pi(p) \in M$. Then $\sigma(x)$ and $p$ lie in the same fibre $\pi^{-1}(x)$ of $P$. Then by definition of the principal bundle there exists a unique $g(p) \in G$ such that

$$p = \sigma(x)g(p). \tag{4.13}$$

This is how we would like to define $g : P \to G$ since it depends on the fibre a better notation would be $g_\sigma$ but $g$ is common notation so we use $g$.

$$g(p) = (\sigma(x))^{-1}p = (\sigma(\pi(p))^{-1}p. \tag{4.14}$$

Now define the trivialisation $\phi_\sigma : P \to M \times G$ as

$$\phi_\sigma(p) = (\pi(p), g(p)). \tag{4.15}$$

Consider $\omega_0$ as the connection 1-form constructed before on $M \times G$,

$$\omega_0 = Ad_{\pi_2^{-1}}\pi_1^*(B) + \pi_2^*(\Omega). \tag{4.16}$$

Connection 1-forms are conserved under the pull back, so we will pull this $\omega_0$ back using $\phi_\sigma^*$ which gives the connection 1-form on $P$

$$\omega := \phi_\sigma^*(\omega_0) = \phi_\sigma^*(Ad_{\pi_2^{-1}}\pi_1^*(B) + \pi_2^*(\Omega)). \tag{4.17}$$

Note the following: $\pi_1\phi_\sigma = \pi$, $\pi_2\phi_\sigma = g$ meaning that $\phi_\sigma^*\pi_1^* = \pi^*$ and $\phi_\sigma^*\pi_2^* = g^*$, we also see that $\phi_\sigma^*(Ad_{\pi_2^{-1}}) = Ad_{\pi_2^{-1}(\phi_\sigma)} = Ad_{g^{-1}}$. Then

$$\omega = Ad_{g^{-1}}\pi^*(B) + g^*(\Omega). \tag{4.18}$$

We also have that $\phi_\sigma\sigma = \sigma_1$,

$$\sigma^*(\omega) = \sigma^*\phi_\sigma^*(\omega_0) = \sigma_1^*(\omega_0) = B. \tag{4.19}$$

The uniqueness is because $\phi_\sigma$ is an isomorphism of principal bundles. □

Now let's look at the trivialisation $\phi_\sigma : P \to M \times G$ in more detail. The section related to this this trivialisation has the following important property, which follows immediately from the definition

$$(\phi_\sigma \circ \sigma)(x) = \phi_\sigma(\sigma(x)) = (\pi(\sigma(x)), g(\sigma(x)) = (x, (\sigma(x))^{-1}(\sigma(x))) = (x, e).$$

This holds for any section $\sigma$ with respect to its own trivialisation $\phi_\sigma$. And there is a trivialisation for all possible sections. So now we ask ourselves, how do two different trivialisations/sections relate?
Consider two sections $\sigma_1, \sigma_2 : M \to P$ on the same trivial principal bundle $\pi : P \to M$. $\forall x \in M$ we have $\sigma_1(x) \in \pi^{-1}(x)$ and $\sigma_2(x) \in \pi^{-1}(x)$, so both are in the same fibre. This means that

$$\sigma_2(x) = \sigma_1(x)g(x), \tag{4.20}$$

for a smooth function $g : M \to G$. Not only that, we also have that given a particular section $\sigma_1 : M \to P$, and $g : M \to G$ smooth, we have that $\sigma_2(x) = \sigma_1(x)g(x)$ defines a section. Therefore we can represent all sections of $P$ by

$$\{g : M \to G | g \text{ smooth}\}.$$

Such a function $g$ is called a gauge transformation, we will soon see that this is indeed a more general definition. For this to make sense you would need a "basis" section, which we will always take to be the trivial section $\sigma(x) = (x, e)$.
We have been considering trivial principal bundles, but with these transformations we can easily generalise to any principal fibre bundle. Since every bundle is locally trivial, the base manifold is covered by $\{V_\alpha | \alpha \in A\}$, then we define the local trivialisations as

$$\phi_\alpha : \pi^{-1}(V_\alpha) \to V_\alpha \times G. \tag{4.21}$$

The sections that are related to these local trivialisations are given by

$$\sigma_\alpha(x) := \phi_\alpha^{-1}(x, e) \quad x \in V_\alpha. \tag{4.22}$$

Moreover in each non-empty intersection $V_\alpha \cap V_\beta$ we have two sections $\sigma_\alpha, \sigma_\beta$, then we know that for all $x \in V_\alpha \cap V_\beta$ there exists a smooth function $g_{\alpha\beta}$ such that,

$$\sigma_\beta(x) = \sigma_\alpha(x)g_{\alpha\beta}. \tag{4.23}$$

This leads to the following theorem that relates our discussion back to chapter 2.

## 4.4 Relating back to Electromagnetism and Yang Mills

**Theorem 4.4.1.** *Let $\pi : P \to M$ be a Principal fibre bundle with Lie group $G$ and Lie algebra $\mathcal{G} = T_eG$, and $\omega$ a connection 1-form on $P$.*
*Suppose $\sigma_1, \sigma_2 : M \to P$ two sections of $P$ and define $B_1 = \sigma_1^*\omega$, $B_2 = \sigma_2^*\omega$.*

*So $B_1, B_2$ are $\mathcal{G}$-valued 1-forms on the base space (in physics, we call the terms $B_1, B_2$ are gauge fields on $M$). Also for all $x \in M$ there is a $g$ s.t.*

$$\sigma_2(x) = \sigma_1(x)g(x), \tag{4.24}$$

*where $g = g_{12} : M \to G$ a local gauge transformation.*
*Then we have the following relation between $B_1$ and $B_2$*

$$B_2 = Ad_{g^{-1}} B_1 + g^*\Omega, \tag{4.25}$$

*where $\Omega$ denotes the Maurer-Cartan form on $G$.*

*Proof.* By theorem 4.3.1 we must have that

$$\omega = Ad_{g_1^{-1}} \pi^* B_1 + g_1^*\Omega, \tag{4.26}$$

where $g_1 : P \to G$ is given by $g_1(p) = (\sigma_1 \circ \pi(p)^{-1})p$. Then we find

$$g(x) = (\sigma_1(x))^{-1}\sigma_2(x) = (\sigma_1(\pi\sigma_2(x)))^{-1}\sigma_2(x) = g_1(\sigma_2(x)). \tag{4.27}$$

So $g = g_1\sigma_2$ then

$$\begin{aligned} B_2 &= \sigma_2^*\omega \\ &= \sigma_2^*(Ad_{g_1^{-1}}\pi^* B_1 + g_1^*\Omega) \\ &= Ad_{\sigma_2^* g_1^{-1}}\sigma_2^*\pi^* B_1 + \sigma_2^* g_1^*\Omega \\ &= Ad_{g^{-1}} B_1 + g^*\Omega. \end{aligned}$$

□

When we consider matrix Lie groups we get

$$B_2 = g^{-1}B_1 g + g^{-1}dg. \tag{4.28}$$

When considering this over open cover of $M$ we find the same result

$$B_\beta = Ad_{g_{\alpha\beta}^{-1}} B_\alpha + g_{\alpha\beta}^*\Omega. \tag{4.29}$$

Now we move to the curvature of a connection 1-form

$$\Theta = D\omega \tag{4.30}$$

to relate this to the physical field strength similar to $B = \sigma^*\omega$ we take

$$F = \sigma^*(\Theta) = \sigma^*(D\omega). \tag{4.31}$$

Then using the structure theorem

$$\begin{aligned} F &= \sigma^*(d\omega + \frac{1}{2}\omega \wedge \omega) \\ &= d\sigma^*\omega + \frac{1}{2}\sigma^*\omega \wedge \sigma^*\omega \\ &= dB + \frac{1}{2}B \wedge B. \end{aligned}$$

Exactly the same as what we found before! Where we use that

$$\begin{aligned} \sigma^* d\omega = d\omega(\sigma_*) = d\sigma^*\omega \quad \text{and} \\ \sigma^*(\omega \wedge \omega) = (\omega \wedge \omega)\sigma_* = \omega\sigma_* \wedge \omega\sigma_* = \sigma^*\omega \wedge \sigma^*\omega. \end{aligned}$$

# Chapter 5

# Chern-Simons theory

In this chapter the connection will be related to a new way to consider a connection namely the Koszul connection or the connection derivative. Chern-Weil theorem uses this Koszul connection, which in turn defines the Chern-Simons term. In the first section the book of Tu was used (8), for the second two sections the first chapter of the lecture notes by Zhang (9).

## 5.1 Koszul Connection

To go to Chern-Simons theory we need to look at a new kind of connection, where again we find the relation between that connection and the connection 1-form.

We begin this definition by defining the vector bundle, a concept similar to the principal bundle, but distinct in the way that the fibre of a vector bundle is a vector space, where the fibre of a principal bundle was a Lie group.

**Definition 5.1.1.** *A smooth surjection $\pi : E \to M$ is a smooth vector bundle of rank $r$ if*

1. *$\forall x \in M$, $E_x := \pi^{-1}(x)$ is a real vector space of dim $r$. We call $E_x$ the fibre of $E$ above $x$.*

2. *$\forall x \in M$ $\exists$ an open subset $U \subset M$ such that there is a fibre preserving diffeomorphism*
$$\phi_U : \pi^{-1}(U) \to U \times \mathbb{R}^r$$
*that restricts to a linear isomorphism $E_x \to \{x\} \times \mathbb{R}^r$ on a fibre.*

Identically to what we did on the principal bundle, we will also define a section on the vector bundle.

**Definition 5.1.2.** *A section of a vector bundle $\pi : E \to M$ over an open subset $U \subset M$ is a function $\sigma : U \to E$ s.t. $\pi \circ \sigma = \mathbb{1}_U$ the identity map of $U$. The set of all smooth sections of $E$ over $U$ is denoted by $\Gamma(U, E)$, if $U$ is the manifold $M$ then we write $\Gamma(M, E) = \Gamma(E)$.*

What is happening is that for each $x \in U$ the function $\sigma$ picks out one element of the fibre $E_x$.

*Note.*
- $\Gamma(U, E)$ is a vector space over $\mathbb{R}$.

- $\sigma \in \Gamma(U, E)$, $f \in C^\infty(U)$ we have

$$(f\sigma)(x) := f(x)\sigma(x) \in E_x, \quad x \in U \tag{5.1}$$

that makes $f\sigma \in \Gamma(U, E)$.

Now we define something new, namely a frame which is the basis of a fibre.

**Definition 5.1.3.** *A frame of a vector bundle $E$of rank $r$ over an open $U \in M$ is a collection of sections $e_1, ..., e_r$ of $E$ over $U$ such that for each point $x \in U$ the elements $e_1(x), ..., e_r(x)$ form a basis for the fibre $E_x$.*

We finally get to the promised connection derivative gives as:

**Definition 5.1.4.** *Let $E \to M$ be a smooth vector bundle over $M$. A connection derivative on $E$ is a map*

$$\nabla^E : \mathcal{X}(M) \times \Gamma(E) \to \Gamma(E) \tag{5.2}$$

*such that for $X \in \mathcal{X}$ and $\sigma \in \Gamma(E)$ we have that*

1. *$\nabla^E_X(\sigma)$ is $C^\infty$-linear in $X$ and $\mathbb{R}$-linear in $\sigma$.*
2. *(Leibniz rule) if $f$ is a smooth function then*

$$\nabla^E_X(f\sigma) = (Xf)\sigma + f\nabla^E_X\sigma.$$

*Since $XF = df(X)$ we may write this as*

$$\nabla^E_X(f\sigma) = df(X)\sigma + f\nabla^E_X\sigma.$$

We might as well consider to not evaluate $X$ yet and make this the function:

$$\nabla^E : \Gamma(E) \to \Omega^1(M, E). \tag{5.3}$$

So we send sections to 1-forms. 1-forms are maps from $TM \to E$, and sections are maps $M \to E$ which is a 0-form. Then we could extend this to the following

$$\nabla^E : \Omega^k(M, E) \to \Omega^{k+1}(M, E). \tag{5.4}$$

But we will consider its significance later. For now we stay with the original definition. We are also able to restrict this connection to a subset of $M$, which makes for a slightly weird notation but we take this to be

$$\nabla^U : \mathcal{X}(U) \times \Gamma(U, E) \to \Gamma(U, E).$$

We have $e_1, ..., e_r$ a frame for $E$ over $U$, take $X \in \mathcal{X}(U)$. Then $\sigma \in \Gamma(U, E)$ is a linear combination $\sigma = a^i e_i$ and we have

$$\nabla^U_X\sigma = \nabla^U_X(a^i e_i) = a^i\nabla^U_X e_i.$$

So $\nabla^U_X\sigma$ is defined by $\nabla^U_X e_i$, this is also a section. So we can also write this in terms of the basis:

$$\nabla_X e_i = \omega^j_i(X)e_j.$$

Then $\omega^i_j : TM \to \mathbb{R}$, is a 1-form! We take $\omega^i_j$ and consider $\omega = [\omega^i_j]$ a matrix and $e = [e_1, ..., e_r]$, then we have $\nabla_X e = e\omega(X)$, taking out $X$ again we have $\nabla e = e\omega$.

*Note.* Say we change the basis of our section in the following way $\bar{e} = ea$, then we must also have that $\nabla\bar{e} = \bar{e}\bar{\omega}$, we see that

$$\nabla_X(ea) = (\nabla_X e)a + eX(a)$$

by Leibniz rule, then again taking it as a function of $X$

$$\nabla\bar{e} = \nabla(ea) = (\nabla e)a + eda = e\omega a + eda$$

combining gives:

$$\begin{aligned} e\omega a + eda &= ea\bar{\omega} \\ \omega a + da &= a\bar{\omega} \\ \bar{\omega} &= a^{-1}\omega a + a^{-1}da. \end{aligned}$$

Which is the gauge transformation we saw earlier, if $\omega$ is indeed a connection in the same way.

Moving on we define the Frame bundle of a vector space, this will be a principal bundle, and we can therefore use this to relate this new connection back to the one we previously used.

**Definition 5.1.5** (Frame bundle of a vector bundle)**.** *Let $V$ be a vector space and $Fr(V)$ be the set of all ordered bases in $V$. Suppose $V$ as dimension $r$, we will represent and ordered basis $v = [v_1, ..., v_r]$ s.t. the general linear group $GL(r, \mathbb{C})$ i.e. the group of all invertible complex matrices under matrix multiplication. Which acts on $Fr(V)$ by matrix multiplication on the right.*

$$vA \qquad v \in Fr(V), g \in GL(r, \mathbb{C}).$$

*If $v' \in F(V)$, then there exists an $g \in GL(r, \mathbb{C})$ s.t. $v' = vg$ We will consider the function:*

$$\phi_v : GL(r, \mathbb{C}) \to Fr(V) \quad g \mapsto vg$$

*a bijection.*
*Then $Fr(V)$ is called a frame manifold of the vector space $V$.*
*We can consider the map $\pi : Fr(V) \to M$, this is a smooth principal $GL(r, \mathbb{C})$-bundle, we call this the frame bundle.*

To prove the final result we will need the following lemma as stated in the book of Tu Lemma 29.7 (8).

**Lemma 5.1.1.** *Take $\tilde{X} \in H_p$ the horizontal subspace of $T_pP$ then then $\tilde{X}$ is called the horizontal lift of $X$ to the frame bundle $P$. Suppose $\nabla$ is a connection on vector bundle $E \to M$ and $\omega_\sigma$ is its connection matrix on a framed open set $(U, \sigma)$. For $x \in U$, $p = \sigma(x) \in Fr(E)$, and $X_x \in T_xM$, let $\tilde{X}_p$ be the horizontal lift of $X_x$ to $p$ in $Fr(E)$. Then*

$$\tilde{X}_p = \sigma_{*,x}(X_x) - (\omega_\sigma(X_x))^{\#}_p. \tag{5.5}$$

Then we see the link between the connection on a vector bundle and the connection 1-form of a principal bundle.

**Theorem 5.1.2.** *Let $\nabla$ be a connection on a vector bundle $E \to M$ and let $\omega$ be a connection 1-form on the frame bundle $Fr(E)$.*
*If $e = (e_1, ..., e_r)$ is a frame for $E$ over an open set $U$, viewed as a section $e : U \to Fr(E)|_U$ and $\omega_e$ is the connection matrix of $\nabla$ relative to the frame $e$, then $\omega_e = e^*\omega$.*

*Proof.* Let $x \in U$ and $p = e(x) \in Fr(E)$. Suppose $X_x \in T_xM$, write $\omega_{e,x}$ for the value of the connection matrix at $x \in U$, then $\omega_{e,x}$ is an $r \times r$ matrix of 1-forms at x and $\omega_{e,x}(X_x)$ is an $r \times r$ matrix of complex numbers, i.e. an element of the Lie algebra $gl(r, \mathbb{C})$. The corresponding fundamental vector field on $Fr(E)$ is $(\omega_{e,x}(X_x))^\#$. By lemma 5.1.1, the horizontal lift of $X_x$ to $p \in Fr(E)$ is

$$\tilde{X}_p = e_*(X_x) - (\omega_{e,x}(X_x))^\#_p. \tag{5.6}$$

Then applying the connection one form $\omega_p$ on both sides of this equation, we get

$$\begin{aligned} 0 = \omega_p(\tilde{X}_p) &= \omega_p(e_*(X_x)) - \omega_p((\omega_{e,x}(X_x))^\#_p) \\ &= e^*\omega_p(X_x) - \omega_{e,x}(X_x). \end{aligned}$$

Since this is true for all $X_x \in T_xM$, we write

$$e^*\omega_p = (e^*\omega)_x = \omega_{e,x}. \tag{5.7}$$

Which shows the relation between the connection 1-form and the connection. □

Also note that this is the same way we define Gauge theory for physics!

## 5.2 Chern-Weil theory

An endomorphism is a map from space $E$ to itself, denoted by $End(E)$. Then we have that the map

$$\big[\Gamma(E) \to \Omega^k(M; E)\big] \in \Omega^k(M, End(E)). \tag{5.8}$$

We will then define the following map called the trace

**Definition 5.2.1.** *The trace is a map*

$$\mathrm{tr} : \Omega^k(M, End(E)) \to \Omega^k(M)$$

*such that for all $\omega \in \Omega^k(M)$, $A \in \Gamma(End(E))$*

$$\mathrm{tr} : \omega A \mapsto \omega \,\mathrm{tr}(A).$$

*We define the Lie bracket as follows*

$$[\omega A, \eta B] = (\omega A)(\eta B) - (-1)^{deg(\omega)deg(\eta)}(\eta B)(\omega A).$$

Combining these two parts of the definition we see that:

**Lemma 5.2.1.** *If $A, B \in \Omega^*(M, End(E))$, then $\mathrm{tr}[A, B] = 0$.*

Here $\Omega^*(M, End(E))$ is the collection of $\Omega^k(M, End(E))$ for all $k$.

*Proof.* Take $A, B \in \Omega^*(M, End(E))$, then we can write $A, B$ as $k$, $l$-forms, $A'\omega_A, B'\omega_B$ respectively, where $A', B'$ are matrices. Then we have

$$\begin{aligned}\operatorname{tr}[A,B] &= \operatorname{tr}[A'\omega_A, B'\omega_B] = \omega_A \wedge \omega_B \operatorname{tr}[A', B'] \\ &= \omega_A \wedge \omega_B(\operatorname{tr}(A'B') - \operatorname{tr}(B'A')) = \omega_A \wedge \omega_B(\operatorname{tr}(A'B') - \operatorname{tr}(A'B')) \\ &= 0.\end{aligned}$$

□

This leads us to the following lemma:

**Lemma 5.2.2.** *If $\nabla^E$ is a connection on $E$, then for $A \in \Omega^k(M, End(E))$ one has*

$$d \operatorname{tr}[A] = \operatorname{tr}\big[\nabla^E, A\big]. \tag{5.9}$$

*Proof.* Take some other connection $\tilde{\nabla}^E$, then $\nabla^E - \tilde{\nabla}^E \in \Omega^1(M, End(E))$. That means that

$$\begin{aligned}\operatorname{tr}\Big[\nabla^E - \tilde{\nabla}^E, A\Big] = 0, \\ \operatorname{tr}\big[\nabla^E, A\big] = \operatorname{tr}\Big[\tilde{\nabla}^E, A\Big].\end{aligned}$$

since $\tilde{\nabla}^E$ was arbitrary we might as well take $\tilde{\nabla}^E = d$ since this is also a connection. Which gives the result

$$\operatorname{tr}\big[\nabla^E, A\big] = \operatorname{tr}[d, A] = d \operatorname{tr}[A]. \tag{5.10}$$

□

The final thing we define before the Chern-Weil theorem is the power series, given by $f(x) = a_0 + a_1 x + ... + a_n x^n + ....$ Then we have the following

**Theorem 5.2.3** (Chern-Weil theorem)**.** *Taking $f$ a power function, and $R^E = (\nabla^E \circ \nabla^E)$ the curvature of the connection $\nabla^E$*

1. *The form form $\operatorname{tr}\big[f(R^E)\big]$ is closed. That is,*

$$d \operatorname{tr}\big[f(R^E)\big] = 0. \tag{5.11}$$

2. *If $\tilde{\nabla}^E$ is another connection on $E$ and $\tilde{R}^E$ its curvature, then there is a differential form $\omega \in \Omega^*(M)$ such that*

$$\operatorname{tr}\big[f(R^E)\big] - \operatorname{tr}\Big[f(\tilde{R}^E)\Big] = d\omega. \tag{5.12}$$

*Proof.* Proof of 1:

$$d \operatorname{tr}\big(f(R^E)\big) = \operatorname{tr}\big[\nabla^E, f(R^E)\big] = \operatorname{tr}\big(a_1[\nabla^E, R^E] + ... + a_n[\nabla^E, (R^E)^n] + ...\big) = 0.$$

The first equality comes from lemma 5.2.2, the second is because of the definition of $f$ and the third is because

$$[\nabla^E, (R^E)^n] = [\nabla^E, ((\nabla^E)^2)^n] = 0,$$

by commutativity. Which proves part 1. Now part 2
For any $t \in [0,1]$ define

$$\nabla_t^E = (1-t)\nabla^E + t\tilde{\nabla}^E.$$

Then we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\nabla_t^E = \tilde{\nabla}^E - \nabla^E \in \Omega^1(M, End(E)).$$

Let $R_t^E$ be the curvature of $\nabla_t^E$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{tr}\left(f(R_t^E)\right) &= \operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}t}f(R_t^E)\right) = \operatorname{tr}\left(\frac{\mathrm{d}R_t^E}{\mathrm{d}t}f'(R_t^E)\right) = \operatorname{tr}\left(\frac{\mathrm{d}(\nabla_t^E)^2}{\mathrm{d}t}f'(R_t^E)\right) \\
&= \operatorname{tr}\left((\nabla_t^E\frac{\mathrm{d}}{\mathrm{d}\nabla_t^E} + \frac{\mathrm{d}\nabla_t^E}{\mathrm{d}t}\nabla_t^E)tf'(R_t^E)\right) = \operatorname{tr}\left([\nabla_t^E, \frac{\mathrm{d}}{\mathrm{d}\nabla_t^E}]f'(R_t^E)\right) \\
&= \operatorname{tr}\left([\nabla_t^E, \frac{\mathrm{d}}{\mathrm{d}\nabla_t^E}f'(R_t^E)]\right) \\
&= d\operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}\nabla_t^E}f'(R_t^E)\right).
\end{aligned}$$

Then when we integrate both sides from $t=0$ to $t=1$ we get by the fundamental theorem of calculus

$$\begin{aligned}
\operatorname{tr}\left(f(R_1^E)\right) - \operatorname{tr}\left(f(R_0^E)\right) &= \int_0^1 d\operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}\nabla_t^E}f'(R_t^E)\right)dt \\
\operatorname{tr}\left(f(R^E)\right) - \operatorname{tr}\left(f(\tilde{R}^E)\right) &= -\int_0^1 d\operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}\nabla_t^E}f'(R_t^E)\right)dt.
\end{aligned}$$

□

The right hand side term is usually called the Chern-Simons term.

## 5.3 Chern-Simons theory

We will now consider $E = TM$ the tangent bundle of a smooth oriented three dimensional manifold $M$. (Taking an appropriate connection and function $f$, we see the "usual" Chern-Simons term appear).
We take the connection $\nabla_t^{TM} = d + tA$, and take the power function to be $f(x) = -x^2$, then plugging this into the Chern-Simons term we find

$$\begin{aligned}
-\int_0^1 d\operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}\nabla_t^{TM}}f'(R_t^{TM})\right)dt &= -\int_0^1 d\operatorname{tr}\left(\frac{\mathrm{d}}{\mathrm{d}d+tA}f'((d+tA)^2)\right)dt \\
&= -\int_0^1 d\operatorname{tr}\left(A(-2)(d+tA)^2\right)dt \\
&= 2\operatorname{tr}\left[\int_0^1 A\wedge dd + A\wedge tdA + A\wedge tAd + A\wedge tA\wedge tA\right] \\
&= \operatorname{tr}\left[A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right]. \qquad (5.13)
\end{aligned}$$

This is the term that is most commonly referred to as the Chern-Simons term in the physics context.

# Chapter 6

# Inflation

## 6.1 Introduction to inflation

The lectures of Stathes Paganis were used (10) as well as the book of Weinberg (11) to write the following.
We know that the universe is homogeneous and isotropic, what does this then mean for the metric. There are three geometries that are homogeneous and isotropic. The first is the flat space geometry (or euclidean geometry), $ds^2 = d\vec{x}^2$ We can rotate $\vec{x}$ and nothing will change, we can also displace $\vec{x}$ and nothing will change. The second geometry is the 4 dimensional euclidean geometry, (Spherical space). $ds^2 = d\vec{x}^2 + dz^2$ and $z^2 + x^2 = a^2$. Naturally a sphere can be rotated without changing the way it looks, and also any place looks the same wherever you are on a sphere. The third geometry is called the 4 dimensional hyperbolic geometry. That this is also homogeneous and isotropic similar to the sphere but then the other way around. These last two can be summarized into

$$ds^2 = d\vec{x}^2 \pm dz^2 \quad \text{and} \quad z^2 \pm x^2 = a^2. \tag{6.1}$$

The coordinate system can be rescaled to $\vec{x}' = a\vec{x}$ and $z' = az$, $t' = t$. When filling in the primed values into 6.1 we get: $ds^2 = a[d\vec{x}^2 \pm dz^2]$ and $z^2 \pm \vec{x}^2 = 1$ and after differentiating $z^2 \pm \vec{x}^2 = 1$, we can summarize all three geometries in one equation:

$$ds^2 = a^2[d\vec{x}^2 + K\frac{(\vec{x} \cdot d\vec{x})^2}{1 - K\vec{x}^2}], \tag{6.2}$$

where

$$K = \begin{cases} +1 & \text{Spherical space,} \\ -1 & \text{Hyperbolic space,} \\ 0 & \text{Euclidean space.} \end{cases} \tag{6.3}$$

Here $a$ is only dependent on time because the universe is isotropic. To extend this to space time we can use the definition of proper time

$$d\tau^2 \equiv -g_{\mu\nu}(x)dx^\mu dx^\nu. \tag{6.4}$$

For $d\tau > 0$, a particle is called "time like", and $d\tau$ does not depend on whether such a particle is moving or stationary. Then we use the Robertson-Walker scale factor, which gives

$$d\tau^2 = dt^2 - ds^2 = dt^2 - a^2(t)[d\vec{x}^2 + K\frac{(\vec{x} \cdot d\vec{x})^2}{1 - K\vec{x}^2}]. \tag{6.5}$$

We have that for $ds = 0$, $d\tau = dt$, so this means that we can find all values of $g_{\mu\nu}$

$$g_{ij} = a^2(t)\left[\delta_{ij} + K\frac{x^i x^j}{1 - K\vec{x}^2}\right], \quad g_{00} = -1, \quad g_{i0} = 0, \quad g_{0i} = 0, \tag{6.6}$$

$i, j$ are running over 1,2,3 and $x^0 \equiv t$.

When rewriting 6.5 into spherical polar coordinates

$$d\tau^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1 - K\vec{x}^2} + r^2 d\Omega\right] \tag{6.7}$$

($d\vec{x}^2 = dr^2 + r^2 d\Omega$, $\Omega \equiv d\theta^2 + \sin^2\theta d\phi^2$). In this case the metric will become diagonal, with

$$g_{rr} = \frac{a^2(t)}{1 - Kr^2}, \quad g_{\theta\theta} = a^2(t)r^2, \quad g_{\phi\phi} = a^2(t)r^2\sin^2\theta, \quad g_{00} = -1. \tag{6.8}$$

For $d\tau = 0$ we can find the proper distance between two objects,

$$dt^2 = a^2(t)\left[\frac{dr^2}{1 - K\vec{x}^2} + r^2 d\Omega\right]. \tag{6.9}$$

Since we are only interested in the distance and not in direction we set the change in $\theta$ and $\phi$ to zero. After which we can take the square root on both sides

$$dt = a(t)\left[\frac{dr}{\sqrt{1 - K\vec{x}^2}}\right]. \tag{6.10}$$

If we then know the time of the emission ($t_e$) and the distance (r) at the time of emission, and we consider us to be at time zero and distance zero. We have

$$d(r,t) = \int_0^{t_e} dt = \int_0^r a(t)\left[\frac{dr}{\sqrt{1 - K\vec{x}^2}}\right]$$

$$d(r,t) = a(t) \times \begin{cases} \sin^{-1} r & K = +1 \\ \sinh^{-1} r & K = -1 \\ r & K = 0 \end{cases}$$

$\frac{dr}{\sqrt{1-K\vec{x}^2}}$ is called the co-moving distance, and importantly, this is not a function of time. When rewriting the equation by dividing both sides with $a(t)$ we can see that this makes it appear as if the speed of light changes with the motion of the universe. Even though at any place in the universe the speed of light will be constant. Finally we can observe that $\dot{d}/d = \dot{a}/a$, and therefore that $v_{galaxy} = H(t)d_{galaxy}$. Where $H(t) \equiv \dot{a}/a$ To get a physical idea of $a$, it is the size of the universe. where it's derivative is the expansion speed and the second derivative the acceleration.

### 6.1.1 Christoffel symbols

We would like to look at the (proper) acceleration, this is to understand forces in a relativistic sense. From this we can also look at geodesics. First velocities,

the four vector (proper) velocity is given by $v^\mu = \vec{v} = \frac{d\vec{s}}{du^\nu}$ where $\vec{s} = x^\mu \hat{e}_\mu$. $u$ can be any type of time, and proper velocity is when $u = \tau$ for massive (and therefore time like) particles. Now let us look at acceleration, where this is proper when $u = \tau$.

$$\frac{d\vec{v}}{du^\nu} = \frac{d}{du^\nu}(v^\mu \hat{e}_\mu) = \frac{dv^\mu}{du^\nu}\hat{e}_\mu + v^\mu \frac{d\hat{e}_\mu}{du^\nu}. \tag{6.11}$$

Now define the Christoffel symbol $\Gamma^t$ to satisfy

$$\Gamma^\lambda_{\nu\mu} \cdot \hat{e}_\lambda = \frac{\partial \hat{e}_\mu}{\partial u^\nu}. \tag{6.12}$$

We then get

$$\frac{dV^\mu}{du} = \frac{\partial V^\mu}{\partial u} + \Gamma^\mu_{\lambda\nu} V^\lambda V^\nu$$
$$\frac{dV^\mu}{du} = \frac{d^2x^\mu}{du^2} + \Gamma^\mu_{\lambda\nu}\frac{dx^\lambda}{du}\frac{dx^\nu}{du}.$$

When this is equal to zero we have a geodesic. Where this formula can be seen as the force, when multiplying both sides by $m$ mass. So a geodesic is the path when no force is applied. The $\Gamma$ term represents the curvature of spacetime and therefore the gravitational field.
These $\Gamma$'s are not unique, but when we assume that we have torsion free spacetime and parallel transport of two vectors leaves lengths and angles invariant (Levi-Civita). We get that

$$\Gamma^\mu_{\nu\kappa} = \frac{1}{2}g^{\mu\lambda}\left[\frac{\partial g_{\lambda\nu}}{\partial x^\kappa} + \frac{\partial g_{\lambda\kappa}}{\partial x^\nu} - \frac{\partial g_{\nu\kappa}}{\partial x^\lambda}\right]. \tag{6.13}$$

Here we can understand the first $g$ term to be the gravitational field, and the derivatives as gradient of the gravity. From this and the Robertson Walker values for $g_{\mu\nu}$ the gammas can be found.

Let us now find the geodesics

$$\frac{d^2x^\mu}{du^2} + \Gamma^\mu_{\nu\kappa}\frac{dx^\nu}{du}\frac{dx^\kappa}{du} = 0. \tag{6.14}$$

By first defining $\tilde{g}_{ij} \equiv \left(\delta_{ij} + K\frac{x^i x^j}{1-K\vec{x}^2}\right)$ and then using equation 6.6

$$\Gamma^0_{ij} = a\dot{a}\tilde{g}_{ij},$$
$$\Gamma^i_{0j} = \frac{\dot{a}}{a}\delta_{ij},$$
$$\Gamma^i_{jl} = K\tilde{g}_{jl}x^i.$$

We can further simplify this by assuming that x is close to the origin, then $\tilde{g}_{ij} = \delta_{ij}$,

$$\Gamma^0_{ij} = a\dot{a}\delta_{ij},$$
$$\Gamma^i_{0j} = \frac{\dot{a}}{a}\delta_{ij},$$
$$\Gamma^i_{jl} = 0.$$

Using this we can solve

$$\frac{d^2x^i}{d\tau^2} = -\Gamma^i_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -2\frac{\dot{a}}{a}\frac{dx^i}{d\tau}\frac{dt}{d\tau}. \tag{6.15}$$

Then multiplying both sides with $\frac{d\tau}{dt}$,

$$\frac{d\tau}{dt}\left(\frac{dx^i}{d\tau}\right) = -2\frac{\dot{a}}{a}\frac{dx^i}{d\tau}. \tag{6.16}$$

So we find

$$\frac{dx^i}{d\tau} \propto \frac{1}{a^2(t)}. \tag{6.17}$$

To find how the momentum changes with a we use $p = m\sqrt{g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau}}$ This means that $p \propto 1/a(t)$.

### 6.1.2 Energy momentum tensor

We would now like to know how momentum and energy influence the universe, and how the things the universe is made of change when the expansion factor $a(t)$ changes. We will do this by first defining the energy-momentum tensor as $T^{\mu\nu}$: Flux of the $p^\mu$ component through a surface of constant $x^\nu$. Where $p^0 = E$ and $p^i$ is momentum in their respective direction. Now consider what each of the elements mean, $T^{00}$ has constant time, and changing x, y, z, which is a volume in which there is energy, This gives the equation $\frac{E}{\Delta x\Delta y\Delta z} = \rho$ which is energy density. The $T^{ii}$ are a momentum going though a surface in a certain time, leading to the equation $\frac{p^1}{\Delta t\Delta y\Delta z} = P^1$ which is force over area, which is pressure in the x direction, a similar equation applies for the other 2 spacial directions. Then we have $T^{0i}$ which is the energy flux, $T^{i0}$ the momentum density and the other terms are shear stress. Finally $T^{\mu\nu}$ has to satisfy the continuity equation $\nabla T^{\mu\nu} = 0$, since whatever flows in has to flow out.

#### Perfect fluid and our Energy momentum tensor

A perfect fluid has at every point has a locally inertial frame, co-moving with the fluid that appears the same in all directions (note that this is very similar to our homogeneous and isotropic universe). From this we can conclude that the energy momentum tensor only has diagonal elements, and the pressure is the same in all directions. We get

$$T^{\mu\nu} = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{bmatrix}.$$

The difference with this and our universe is that in our universe each of the diagonal terms is a function of time. To figure out what the exact values are

for our universe we have to take a closer look at the definition of isotropy and homogeneity.
Starting with isotropy, with energy momentum tensor $t^{\mu\nu}$
At $\vec{x} = 0$. This means that in general we have that $t^{ij} \propto g^{ij} = \tilde{g}^{ij}/a^2(t)$. Note here the difference between $g_{ij} = a^2(t)\tilde{g}_{ij}$ and $g^{ij} = \tilde{g}^{ij}/a^2(t)$. Which since we are at $\vec{x} = 0$ gives $t^{ij} \propto \delta_{ij}/a^2(t)$. (we can lower the index because delta is identity). Then adding homogeneity, which means the proportionality coefficients can only depend on time. Which gives: $t^{ij} = \frac{f(t)}{a^2(t)}\delta_{ij}$. $t^{00} =$ only a function of time, $t^{\mu 0} = t^{0\mu} = 0$. Then combining the physics with the mathematics, we get that $T^{00} = \rho(t)$ energy density and $f(t) = P(t)$ pressure. Writing this in matrix form gives

$$T^{\mu\nu} = \begin{bmatrix} \rho(t) & 0 & 0 & 0 \\ 0 & \frac{P(t)}{a^2(t)} & 0 & 0 \\ 0 & 0 & \frac{P(t)}{a^2(t)} & 0 \\ 0 & 0 & 0 & \frac{P(t)}{a^2(t)} \end{bmatrix}.$$

### 6.1.3 Motion of our universe

Now we give some motivation for the Einstein equations, from which the Friedmann equations can be derived. Both of these concepts will be used to explore further into inflation.

#### Einstein equations

We would like to describe the curvature of space, since this is what gravity is. To check the curvature we can move a vector around in a loop. This vector will not be pointing in the same direction for a curved space. The difference between the initial vector and the rotated vector tells us everything we need to describe the curvature (where we assume there is no torsion). This is described by the Riemann tensor given by $R^{\rho}_{\mu\sigma\nu} = \partial_\sigma \Gamma^{\rho}_{\nu\mu} - \partial_\nu \Gamma^{\rho}_{\sigma\mu} + \Gamma^{\rho}_{\sigma\lambda}\Gamma^{\lambda}_{\nu\mu} + \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\sigma\mu}$.
We however do not need this tensor to understand gravity but can use the Ricci tensor: given by $R_{\mu\nu} = R^{\rho}_{\mu\rho\nu} = \partial_\rho \Gamma^{\rho}_{\nu\mu} - \partial_\nu \Gamma^{\rho}_{\rho\mu} + \Gamma^{\rho}_{\rho\lambda}\Gamma^{\lambda}_{\nu\mu} + \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\rho\mu}$. This Ricci tensor tells us how a volume element changes because of the curvature of space in any direction. However it does not describe the complete deformation. However for gravity this is all we need. In curved space parallel lines can meet, which means that there is a changing distance between two parallel lines, which can be seen as an area. This is something that the Ricci tensor describes.
Let us also define the Ricci scalar $R = g^{\mu\nu}R_{\mu\nu}$, which describes how much the spacetime volume of an object moving along a geodesic changes due to tidal forces. And only describes volume (so no directions). To get to the Einstein equations we define the Einstein Tensor as $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$. Which leads to the general relativity field equations: $G_{\mu\nu} = \frac{8\pi G}{c^2}T_{\mu\nu}$.

**Friedmann equation**

Now let's see what we can do with the Einstein equations, where we can rewrite it for convenience,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}.$$

Taking the trace on both sides we find that

$$\begin{aligned} \mathrm{Tr}(R_{\mu\nu}) - \mathrm{Tr}(1/2 g_{\mu\nu}R) &= 8\pi G\,\mathrm{Tr}(T_{\mu\nu}) \\ R - 2R &= 8\pi G T \qquad (Tr(g_{\mu\nu}) = 4) \\ R &= -8\pi G T. \end{aligned}$$

Then by defining $S_{\mu\nu} = T_{\mu\nu} - 1/2 g_{\mu\nu}T$ we get:

$$R_{\mu\nu} = \frac{1}{2}g_{\mu\nu}R + 8\pi G T_{\mu\nu} = 8\pi G(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T) = 8\pi G S_{\mu\nu}. \tag{6.18}$$

With the Robertson Walker metric we have $g_{\mu\nu}$. We can compute $R_{\mu\nu}$ since we know $g_{\mu\nu}$ and the Christoffel symbols and finally we have $T^{\mu\nu}$ for our universe, which we can invert. Then we get

$$\begin{aligned} R_{00} = -3\frac{\ddot{a}(t)}{a(t)}, \quad R_{ij} = (2k + 2\dot{a}^2 + a\ddot{a})\tilde{g}_{ij}, \\ T_{00} = \rho, \quad T_{0i} = T_{i0} = 0, \quad T_{ij} = a^2 P \tilde{g}_{ij}, \end{aligned}$$

giving

$$\begin{aligned} S_{00} &= T_{00} - 1/2 g_{00}T \\ S_{00} &= \rho + 1/2(3P - \rho) \end{aligned}$$

and

$$\begin{aligned} S_{ij} &= T_{ij} - 1/2 g_{\mu\nu}T \\ S_{ij} &= a^2 P \tilde{g}_{ij} - \frac{1}{2}a^2(t)\tilde{g}_{ij}(3P - \rho). \end{aligned}$$

Which simplifies the Einstein equations to only two equations.
The first one being:

$$\begin{aligned} R_{ij} &= 8\pi G S_{ij} \\ (2K + 2\dot{a}^2 + a\ddot{a})\tilde{g}_{ij} &= 8\pi G[a^2 P\tilde{g}_{ij} - \frac{1}{2}a^2\tilde{g}_{ij}(3P - \rho)] \\ 2K + 2\dot{a}^2 + a\ddot{a} &= 4\pi G a^2(\rho - P) \\ \frac{2K}{a^2} + 2\left(\frac{\dot{a}}{a}\right)^2 + \frac{\ddot{a}}{a} &= 4\pi G(\rho - P). \end{aligned}$$

This is for space, called the $(ij)$ equation. Where $K$ is defined as in 6.1.
For the second one we have

$$\begin{aligned} R_{00} &= 8\pi G S_{00} \\ -3\frac{\ddot{a}(t)}{a(t)} &= 4\pi G(\rho + 3P), \end{aligned}$$

which we call the (00) equation.
We can now subtract the two equations such that the second derivative disappears which gives the Friedmann equation:

$$\dot{a}^2(t) + K = \frac{8\pi G\rho(t)}{3}a^2(t). \tag{6.19}$$

If we know $a(t_{init}), \rho(t_{init})$ and $P = P(\rho)$ the equation of state, than with this equation we can find $a(t)$ and $\rho(t)$.

## 6.2 Space Lagrangian

For the rest of this chapter the book of Carroll is used(12), the book of Weinberg is continued to be used(11).

We will consider a "guess" for an action called the Hilbert action given by

$$S_H = \int d^4x\sqrt{-g}R. \tag{6.20}$$

Solving this action, and recalling that $R = g^{\rho\sigma}R_{\rho\sigma}$, we find

$$\begin{aligned}
\frac{\delta S_H}{\delta g^{\mu\nu}} &= \frac{\partial}{\partial g^{\mu\nu}}\sqrt{-g}R + \partial_\lambda\frac{\partial}{\partial\partial_\lambda g^{\mu\nu}}\sqrt{-g}R \\
&= \frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}}R + \sqrt{-g}\frac{\partial g^{\rho\sigma}R_{\rho\sigma}}{\partial g^{\mu\nu}} \\
&= -\frac{1}{2}\sqrt{-g}g_{\mu\nu}R + \sqrt{-g}\delta^\rho_\mu\delta^\sigma_\nu R_{\rho\sigma} \\
&= \sqrt{-g}\big[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\big] = 0,
\end{aligned}$$

where $\frac{\delta S_H}{\delta g^{\mu\nu}} = 0$ are the Euler-Lagrange equations with respect to $g^{\mu\nu}$. We have that $g = \det(g^{\mu\nu})$ so we use Jacobi's formula $\frac{d}{dA}\det(A) = \det(A)(A^{-1})^T$. Which we see is exactly the Einstein equation without a source, i.e. it only explains space.

Now we can consider adding a matter term to this, making

$$S = \frac{1}{16\pi G}S_H + S_M, \tag{6.21}$$

where the constant is there for consistency. Also solving this with respect to $g^{\mu\nu}$ we have

$$\begin{aligned}
\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}} &= \frac{1}{\sqrt{-g}}\frac{1}{16\pi G}\frac{\delta S_H}{\delta g^{\mu\nu}} + \frac{1}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}} \\
&= \frac{1}{16\pi G}\big[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\big] + \frac{1}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}} = 0.
\end{aligned}$$

We will define $T_{\mu\nu} = -2\frac{1}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}}$. This gives

$$\frac{1}{16\pi G}\big[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\big] = \frac{1}{2}T_{\mu\nu}. \tag{6.22}$$

Which is exactly the Einstein equation!

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}. \tag{6.23}$$

## 6.3 Inflaton matter field

We will now add a matter action related to the inflaton pseudo scalar field $\phi$. Which has action

$$S_M = \int d^4x\sqrt{-g}\big(-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\big) \tag{6.24}$$

where $\phi$ is only dependent on time. We do this because it represents the inflation, which we have claimed to be isotropic and homogeneous, and should therefore not depend on where it is only on when. The $g$, $g^{\mu\nu}$ terms are there to weakly couple the inflaton to gravity. Solving this action with respect to $g^{\mu\nu}$ to find the energy tensor we get

$$\frac{\delta S_M}{\delta g^{\mu\nu}} = \frac{1}{2}\sqrt{-g}g_{\mu\nu}\big(\frac{1}{2}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + V(\phi)\big) - \sqrt{-g}(\frac{1}{2}\delta^\rho_\mu\delta^\sigma_\nu\partial_\rho\phi\partial_\sigma\phi). \tag{6.25}$$

Which is basically the same calculation we did for $S_H$. From this we can find the energy-stress tensor

$$\begin{aligned} T_{\mu\nu} &= -2\frac{1}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}} \\ &= -g_{\mu\nu}\big(\frac{1}{2}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + V(\phi)\big) + \partial_\mu\phi\partial_\nu\phi. \end{aligned}$$

Now simplifying, and returning to flat FLRW-metric, we have that $\rho = T^{00} = T_{00}$, also remembering that $\phi$ is only a function of time

$$\begin{aligned} T_{00} &= -g_{00}\big(\frac{1}{2}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + V(\phi)\big) + \partial_0\phi\partial_0\phi \\ &= (\frac{1}{2}g^{00}\dot{\phi}^2 + V(\phi)) + \dot{\phi}^2 \\ &= -\frac{1}{2}\dot{\phi}^2 + V(\phi) + \dot{\phi}^2 \\ &= \frac{1}{2}\dot{\phi}^2 + V(\phi), \end{aligned}$$

so $\rho(t) = \frac{1}{2}\dot{\phi}^2 + V(\phi)$. Filling that into the Friedmann equation for flat space (K = 0) we have

$$H^2 = \frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right). \tag{6.26}$$

We will define $M_{pl}^2 = \frac{1}{8\pi G}$ making it of the form

$$H^2 = \frac{1}{3M_{pl}^2}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right). \tag{6.27}$$

## 6.4 Rewriting the inflation conditions

For the remainder of this chapter the lecture notes of Baumann (13) are also used.

We know by experiment that the universe is accelerating, meaning that $\ddot{a} > 0$, we also know that the universe is of some positive size $a > 0$. We have that the Hubble constant $H = \frac{\dot{a}}{a}$, then $\dot{H} = \frac{\ddot{a}a - \dot{a}^2}{a^2} = \frac{\ddot{a}}{a} - H^2$. Then we find the following

$$\frac{\dot{H}}{H^2} + 1 = \frac{\ddot{a}}{aH^2} > 0. \tag{6.28}$$

So then

$$\epsilon := -\frac{\dot{H}}{H^2} < 1. \tag{6.29}$$

We also want $\epsilon$ to stay small for even for large Hubble times, therefore we define:

$$|\eta| := \frac{|\dot{\epsilon}|}{H\epsilon} \ll 1. \tag{6.30}$$

Where these conditions on $\epsilon$ and on $\eta$ are now the new conditions for inflation.

## 6.5 Slow-roll inflation

Let us consider our earlier action $S = \frac{M_{pl}}{2} S_H + S_M$. Where the $S_M$ term is the action we defined for the inflaton field,

$$S = \int d^4x \sqrt{-g} \left[ \frac{M_{pl}}{2} R - \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right]. \tag{6.31}$$

Then solving the Lagrange equations with respect to $\phi$ we get

$$\frac{\delta S}{\delta \phi} = \partial_\sigma \frac{\partial \mathcal{L}}{\partial(\partial_\sigma \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{6.32}$$

Solving the terms individually we get

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\sqrt{-g} V'(\phi)$$

and

$$\begin{aligned} \frac{1}{\sqrt{-g}} \frac{\partial \mathcal{L}}{\partial \partial_\sigma \phi} &= -\frac{1}{2} g^{\mu\nu} \delta_{\mu\sigma} \partial_\nu \phi + \frac{1}{2} g^{\mu\nu} \delta_{\nu\sigma} \partial_\mu \phi \\ &= -\partial^\sigma \phi. \end{aligned}$$

In the FLRW-metric we have that $g = \det g^{\mu\nu} = -a^6(t)$, then we get the Euler-Lagange equations

$$\begin{aligned} -\partial_\sigma \sqrt{-(-a^6(t)}\partial^\sigma \phi + \sqrt{-(-a^6(t)} V'(\phi) &= 0 \\ \frac{\partial}{\partial t} + a^3(t) V'(\phi) &= 0 \\ 3\dot{a}a^2\dot{\phi} + a^3\ddot{\phi} + a^3 V'(\phi) &= 0 \\ 3\frac{\dot{a}}{a}\dot{\phi} + \ddot{\phi} + V'(\phi) &= 0 \\ \ddot{\phi} + 3H\dot{\phi} + V'(\phi) &= 0. \end{aligned}$$

Now using the Friedmann equation we are able to find a new restriction that is equivalent to the restriction of inflation in the following way

$$\frac{\mathrm{d}H^2}{\mathrm{d}t} = 2H\dot{H}.$$

Now with the Friedmann equation

$$\begin{aligned}\frac{\mathrm{d}H^2}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{3M_{pl}^2}\big[\frac{1}{2}\dot{\phi} + V(\phi)\big] \\ &= \frac{1}{3M_{pl}^2}\big[\dot{\phi}\ddot{\phi} + \dot{\phi}V'(\phi)\big].\end{aligned}$$

We then fill in the Euler-Lagrange equation for $\phi$ we found earlier.

$$\begin{aligned}\frac{\mathrm{d}H^2}{\mathrm{d}t} &= \frac{1}{3M_{pl}^2}\big[-\dot{\phi}(3H\dot{\phi} + V'(\phi)) + \dot{\phi}V'(\phi)\big] \\ &= \frac{1}{3M_{pl}^2}\big[-3H\dot{\phi}^2\big].\end{aligned}$$

Giving us

$$\dot{H} = -\frac{\dot{\phi}^2}{2M_{pl}^2}. \tag{6.33}$$

We have that

$$\epsilon = -\frac{\dot{H}}{H^2} = \frac{\dot{\phi}^2}{2M_{pl}^2}\frac{3M_{pl}^2}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} = 3\frac{\frac{1}{2}\dot{\phi}^2}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} < 1. \tag{6.34}$$

Which is equivalent to

$$2(\frac{1}{2}\dot{\phi}^2) < V(\phi). \tag{6.35}$$

So for inflation to happen we need that the kinetic energy is smaller than the potential energy.

# Chapter 7

# Chern-Simons term in inflation

## 7.1 Adding the Chern-Simons term

We are considering the following action (3):

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{pl}^2}{2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\alpha}{4f}\phi F_{\mu\nu}\tilde{F}^{\mu\nu}\right]. \quad (7.1)$$

Where $F_{\mu\nu}$ is the field strength and $\tilde{F}^{\mu\nu} = \frac{\epsilon^{\mu\nu\rho\sigma}}{2}F_{\rho\sigma}$. The $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ term is the term we found from the Yang-Mills approach with equation 2.14. The $\phi F\tilde{F}$ term is called the Chern-Simons term, the reason for it will be discussed soon, but first we verify if it is an allowed term in the Lagrangian. For the Chern-Simons term to be allowed we need it to be Lorentz invariant, which it is manifestly. We need it to be conserved under parity, and it has to be gauge invariant. Gauge invariant is easy to show, we have that $F_{\mu\nu} \mapsto F'_{\mu\nu} = S^{-1}F_{\mu\nu}S$. So for the term we have

$$\begin{aligned}\phi F_{\mu\nu}\tilde{F}_{\rho\sigma} = \phi F_{\mu\nu}\epsilon^{\rho\sigma\mu\nu}F_{\rho\sigma} \mapsto \phi F'_{\mu\nu}\epsilon^{\rho\sigma\mu\nu}F'_{\rho\sigma} &= \phi S^{-1}F_{\mu\nu}S\epsilon^{\rho\sigma\mu\nu}S^{-1}F_{\rho\sigma}S \\ &= \phi S^{-1}F_{\mu\nu}SS^{-1}\epsilon^{\rho\sigma\mu\nu}F_{\rho\sigma}S \\ &= \phi S^{-1}F_{\mu\nu}\epsilon^{\rho\sigma\mu\nu}F_{\rho\sigma}S \\ &= \phi S^{-1}F_{\mu\nu}\tilde{F}_{\rho\sigma}S,\end{aligned}$$

since $\epsilon^{\rho\sigma\mu\nu}$ is a scalar, and $\phi$ a (pseudo) scalar field. Since we are really considering this in a trace, we are actually able to cyclically permute the matrices, also cancelling $S$, $S^{-1}$ as in section 2.2.2. Therefore the term is gauge invariant. As for parity, we have that in general $F_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu + [B_\mu, B_\nu]$, and the following transformations under parity

$$\begin{aligned}\partial_0 &\mapsto \partial_0 & \qquad B_0 &\mapsto B_0 \\ \partial_i &\mapsto -\partial_i & \qquad B_i &\mapsto -B_i\end{aligned}$$

then we see that $F_{\mu\nu}$ is mapped in the following way

$$\begin{aligned}F_{00} &\mapsto F_{00} & \qquad F_{ij} &\mapsto F_{ij} \\ F_{0k} &\mapsto -F_{0k} & \qquad F_{k0} &\mapsto -F_{k0}.\end{aligned}$$

Because of the $\epsilon^{\rho\sigma\mu\nu}$, we always have a 0, 1, 2, 3 combination. So one $F_{0k}$ and an $F_{ij}$, so $F\tilde{F}$ changes sign under parity. Since $\phi$ is a pseudo scalar field, it will also change sign under parity. That means that the term $\phi F\tilde{F}$ is conserved under parity. So this so called Chern-Simons term is indeed allowed in the Lagrangian, but how is it related to Chern-Simons theory we just discussed?

### 7.1.1 How is this term related to Chern-Simons theory?

For this part the book of Nakahara (14) was used.

First we see that the term $F_{\mu\nu}\epsilon^{\rho\sigma\mu\nu}F_{\rho\sigma}$, does not have any repeating indices, and is therefore equal to $\mathrm{tr}(F \wedge F)$ where $F = F_{\mu\nu}dx_\mu \wedge dx_\nu$. Now we define

$$K = \mathrm{tr}\left(A \wedge dA + \frac{2}{3}A \wedge A \wedge A\right). \tag{7.2}$$

The trace of the Chern-Simons term. Now see what happens when we take the exterior derivative of $K$

$$\begin{aligned} dK &= \mathrm{tr}\left(d(A \wedge dA) + \frac{2}{3}d(A \wedge A \wedge A)\right) \\ &= \mathrm{tr}\left(dA \wedge dA + \frac{2}{3}(dA \wedge A \wedge A - A \wedge dA \wedge A + A \wedge A \wedge dA)\right). \end{aligned}$$

Now we will use that $F = dA + A \wedge A$ meaning that $dA = F - A \wedge A$,

$$\begin{aligned} dK &= \mathrm{tr}((F - A \wedge A) \wedge (F - A \wedge A) \\ &+ \frac{2}{3}((F - A \wedge A) \wedge A \wedge A - A \wedge (F - A \wedge A) \wedge A + A \wedge A \wedge (F - A \wedge A))) \\ &= \mathrm{tr}(F \wedge F - F \wedge A \wedge A - A \wedge A \wedge F + A \wedge A \wedge A \wedge A \\ &+ \frac{2}{3}(F \wedge A \wedge A - A \wedge A \wedge A \wedge A - A \wedge F \wedge A \\ &+ A \wedge A \wedge A \wedge A + A \wedge A \wedge F - A \wedge A \wedge A \wedge A)). \end{aligned}$$

Note that the $\mathrm{tr}(A \wedge A \wedge A \wedge A) = 0$ and that $\mathrm{tr}(A \wedge F \wedge A) = -\mathrm{tr}(A \wedge A \wedge F) = -\mathrm{tr}(F \wedge A \wedge A)$ due to the cyclicity of the trace and the anti-commutativity of the wedge product (see appendix B). This leaves us with

$$\begin{aligned} dK &= \mathrm{tr}(F \wedge F) - 2\,\mathrm{tr}(F \wedge A \wedge A) + \frac{2}{3}(3\,\mathrm{tr}(F \wedge A \wedge A)) \\ &= \mathrm{tr}(F \wedge F). \end{aligned}$$

Since in the Lagrangian we are integrating over a manifold M, which is all of space in this case, we see the following:

$$\int_M F_{\mu\nu}\tilde{F}^{\mu\nu} = \int_M dK = \int_{\partial M} K. \tag{7.3}$$

Which makes the name Chern-Simons term plausible, in the action we also have a $\phi$ in front. When doing integration by parts there should also be some integral term as given above and therefore one could still call it the Chern-Simons term. Note here that this does not mean that this interaction is simply a boundary term, this just shows the reason behind the naming of this term.

### 7.1.2 What does the term do for inflation?

Let's look at what this term changes is the Euler-Lagrange equations with respect to the inflaton $\phi$, since $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ is independent of $\phi$ this does not appear

in these equations. The Chern-Simons term is linearly dependent on $\phi$. We therefore get the following

$$\frac{\delta S}{\delta \phi} = \partial^\mu(-\sqrt{-g}\partial_\mu\phi) + \sqrt{-g}V'(\phi) + \sqrt{-g}\frac{\alpha}{4f}F_{\mu\nu}\tilde{F}^{\mu\nu} = 0.$$

This gives, since $\phi$ is only dependent on time, the following

$$\ddot{\phi} + 3H = -V'(\phi) - \frac{\alpha}{4f}F_{\mu\nu}\tilde{F}^{\mu\nu}. \tag{7.4}$$

This shows that we can see the Chern-Simons term as an additional source term for the inflaton field. This can amplify one helicity of the gauge field. Depending on the background and backreaction, this can affect slow-roll dynamics and source cosmological perturbations, including gravitational waves.

# Chapter 8

# Conclusion

In this thesis we explained Gauge theory from a physics perspective using the Yang-Mills approach and from a mathematics perspective using connections. We found that there are many kinds of connections, so we explained the relation between them. Eventually we found the link between the Yang-Mills approach and the connection approach through gauge fields and the field strength. This lead to the introduction of Chern-Simons theory using these connections. We gave a first introduction to inflation to find the inflaton field Lagrangian, with this we realised that the inflaton coupled to the Chern-Simons term is an allowed term in the Lagrangian. Solving this inflaton Lagrangian it became clear that the Chern-Simons term acts as a source term.

### Outlook

This thesis lays the foundation to be able to use Chern-Simons theory in topics like inflation and quantum gravity. For example, the helicity of a gauge field is closely related to the Chern-Simons term (15). The Chern-Simons term also appears in quantum gravity. Namely through the Chern-Simons state given by $\Psi_{CS} = e^{-\frac{6}{\Lambda}S_{CS}(A)}$, where $S_{CS}$ is the Chern-Simons action. This state is then in the physical state space if the cosmological constant $\Lambda$ is non-zero(16).

# Bibliography


[1] Q Shafi. Unified gauge theories and the gravitational cut-off. *Journal of Physics A: Mathematical and General*, 10(1):99–104, 1977.

[2] David Bleecker. *Gauge theory and variational principles*. Courier Corporation, 2005.

[3] Angelo Caravano. Simulating the inflationary universe: from single-field to the axion-u (1) model. *arXiv preprint arXiv:2209.13616*, 2022.

[4] TR Govindarajan and Pichai Ramadevi. *Geometry and Topology of Low Dimensional Systems*. Springer, 2024.

[5] Stephen Bruce Sontz. Connections on principal bundles. In *Principal Bundles: The Classical Case*, pages 121–144. Springer, 2015.

[6] Valery Rubakov. Classical theory of gauge fields. 2009.

[7] Tohru Eguchi, Peter B Gilkey, and Andrew J Hanson. Gravitation, gauge theories and differential geometry. *Physics reports*, 66(6):213–393, 1980.

[8] Loring W Tu. *Differential geometry: connections, curvature, and characteristic classes*. Springer, 2017.

[9] Weiping Zhang. *Lectures on Chern-Weil theory and Witten deformations*, volume 4. World Scientific, 2001.

[10] S. Paganis. *Particle astrophysics lectures*. NTU, 2024.

[11] Steven Weinberg. *Cosmology*. OUP Oxford, 2008.

[12] Sean M Carroll. An introduction to general relativity. *Spacetime and geometry*, 101:102, 2004.

[13] Daniel Baumann. *The Physics of Inflation: A Course for Graduate Students in Particle Physics and Cosmology*. 2011.

[14] Mikio Nakahara. *Geometry, topology and physics*. CRC press, 2018.

[15] R Jackiw and So-Young Pi. Creation and evolution of magnetic helicity. *Physical Review D*, 61(10):105015, 2000.

[16] John C Baez and Javier P Muniain. *Gauge fields, knots and gravity*, volume 4. World Scientific Publishing Company, 1994.

# Appendix A

# Intuition with Schrödinger equation

Starting with the Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x_i^2} + V\psi$$

The wave function is a function $\psi : \mathbb{R}^4 \to \mathbb{C}^n$ The physical quantity of the wave function is $\psi^*\psi$. When considering some $n$-dimensional unitary matrix $U$, applying $U$ to $\psi$ gives $\tilde{\psi} = U\psi$ then $\tilde{\psi}^*\tilde{\psi} = (U\psi)^*U\psi = \psi^*U^*U\psi = \psi^*\psi$ so the physical quantity $\psi^*\psi$ is invariant under a unitary operation. Let us now look at what happens to the Schrödinger equation when changing to $\tilde{\psi}$. Since the physical quantity does not change then also the Schrödinger equation should stay invariant. We can write any unitary matrix $U$ as $U = e^{iH}$ where $H$ is an

hermitian matrix. $\psi = e^{-iH}\tilde{\psi}$, looking at the equation part by part we have

$$\begin{aligned}
\frac{\partial \psi}{\partial t} &= \frac{\partial}{\partial t} e^{-iH}\tilde{\psi} \\
&= \frac{\partial e^{-iH}}{\partial t}\tilde{\psi} + e^{-iH}\frac{\partial \tilde{\psi}}{\partial t} \\
&= \frac{\partial e^{-iH}}{\partial H}\frac{\partial H}{\partial t}\tilde{\psi} + e^{-iH}\frac{\partial \tilde{\psi}}{\partial t} \\
&= -ie^{-iH}\frac{\partial H}{\partial t}\tilde{\psi} + e^{-iH}\frac{\partial \tilde{\psi}}{\partial t} \\
\frac{\partial \psi}{\partial x_i} &= \frac{\partial}{\partial x_i} e^{-iH}\tilde{\psi} \\
&= -ie^{-iH}\frac{\partial H}{\partial x_i}\tilde{\psi} + e^{-iH}\frac{\partial \tilde{\psi}}{\partial x_i} \\
\frac{\partial^2 \psi}{\partial x_i^2} &= \frac{\partial^2}{\partial x_i^2} e^{-iH}\tilde{\psi} \\
&= \frac{\partial}{\partial x_i}[-ie^{-iH}\frac{\partial H}{\partial x_i}\tilde{\psi} + e^{-iH}\frac{\partial \tilde{\psi}}{\partial x_i}] \\
&= -i\frac{\partial e^{-iH}}{\partial x_i}\frac{\partial H}{\partial x_i}\tilde{\psi} - ie^{-iH}\frac{\partial^2 H}{\partial x_i^2}\tilde{\psi} - ie^{-iH}\frac{\partial H}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i} + \frac{\partial e^{-iH}}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i} + e^{-iH}\frac{\partial^2 \tilde{\psi}}{\partial x_i^2} \\
&= -e^{-iH}(\frac{\partial H}{\partial x_i})^2\tilde{\psi} - ie^{-iH}\frac{\partial^2 H}{\partial x_i^2}\tilde{\psi} - ie^{-iH}\frac{\partial H}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i} - ie^{-iH}\frac{\partial H}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i} + e^{-iH}\frac{\partial^2 \tilde{\psi}}{\partial x_i^2} \\
&= e^{-iH}((\frac{\partial H}{\partial x_i})^2\tilde{\psi} - i\frac{\partial^2 H}{\partial x_i^2}\tilde{\psi} - 2i\frac{\partial H}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i} + \frac{\partial^2 \tilde{\psi}}{\partial x_i^2}) \\
&= e^{-iH}((\frac{\partial H}{\partial x_i})^2 - i\frac{\partial^2 H}{\partial x_i^2} - 2i\frac{\partial H}{\partial x_i}\frac{\partial}{\partial x_i} + \frac{\partial^2}{\partial x_i^2})\tilde{\psi}
\end{aligned}$$

This final expression we can simplify in the following way, shortly forgetting about $e^{-iH}$. Notice that this is a quadratic formula, which we could possibly decompose, rewriting gives:

$$(\frac{\partial^2}{\partial x_i^2} - i\frac{\partial^2 H}{\partial x_i^2} - 2i\frac{\partial H}{\partial x_i}\frac{\partial}{\partial x_i} + (\frac{\partial H}{\partial x_i})^2)\tilde{\psi}$$

Then lets take the decomposition step by step:

$$(\frac{\partial}{\partial x_i} + \quad)(\frac{\partial}{\partial x_i} + \quad)\tilde{\psi} = \frac{\partial^2}{\partial x_i^2}\tilde{\psi}$$

This is the first term, trying to get the second term:

$$\begin{aligned}
(\frac{\partial}{\partial x_i} + \quad)(\frac{\partial}{\partial x_i} - i\frac{\partial H}{\partial x_i})\tilde{\psi} &= \frac{\partial^2 \tilde{\psi}}{\partial x_i^2} - \frac{\partial}{\partial x_i} i\frac{\partial H}{\partial x_i}\tilde{\psi} \\
&= \frac{\partial^2 \tilde{\psi}}{\partial x_i^2} - i\frac{\partial^2 H}{\partial x_i^2}\tilde{\psi} - i\frac{\partial H}{\partial x_i}\frac{\partial \tilde{\psi}}{\partial x_i}
\end{aligned}$$

Then making the last two terms:

$$(\frac{\partial}{\partial x_i} - i\frac{\partial H}{\partial x_i})(\frac{\partial}{\partial x_i} - i\frac{\partial H}{\partial x_i})\tilde{\psi} = \frac{\partial^2\tilde{\psi}}{\partial x_i^2} - i\frac{\partial^2 H}{\partial x_i^2}\tilde{\psi} - i\frac{\partial H}{\partial x_i}\frac{\partial\tilde{\psi}}{\partial x_i} - i\frac{\partial H}{\partial x_i}\frac{\partial}{\partial x_i}\tilde{\psi} - (\frac{\partial H}{\partial x_i})^2\tilde{\psi}$$
$$= (\frac{\partial^2}{\partial x_i^2} - i\frac{\partial^2 H}{\partial x_i^2} - 2i\frac{\partial H}{\partial x_i}\frac{\partial}{\partial x_i} + (\frac{\partial H}{\partial x_i})^2)\tilde{\psi}$$

So apparently we can write the new Schrödinger equation as:

$$i\hbar e^{-iH}(\frac{\partial}{\partial t} - i\frac{\partial H}{\partial t})\tilde{\psi} = -\frac{\hbar^2}{2m}e^{-iH}(\frac{\partial}{\partial x_i} - i\frac{\partial H}{\partial x_i})^2\tilde{\psi} + Ve^{-iH}\tilde{\psi}$$

This motivates us to keep $\psi$ invariant under derivatives. Consider $\psi_U = U\psi$ and $\psi_V = V\psi$, here $\psi_U, \psi_V$ play the role of $\tilde{\psi}$. Then we have that

$$\frac{\partial\psi}{\partial x_\mu} = U^{-1}(\frac{\partial}{\partial x_\mu} - igB_U)\psi_U \quad \text{and}$$
$$\frac{\partial\psi}{\partial x_\mu} = V^{-1}(\frac{\partial}{\partial x_\mu} - igB_V)\psi_V$$

This gives us

$$U^{-1}(\frac{\partial}{\partial x_\mu} - igB_U)\psi_U = V^{-1}(\frac{\partial}{\partial x_\mu} - igB_V)\psi_V \quad \text{and}$$
$$\psi_V = V\psi = VU^{-1}\psi_U$$

Defining $\psi_U = \psi$, $\psi_V = \tilde{\psi}$, $UV^{-1} = S$, $B_U = B$, $B_V = B'$ then we get:

$$(\frac{\partial}{\partial x_\mu} - igB)\psi = S(\frac{\partial}{\partial x_\mu} - igB')\tilde{\psi} \tag{A.1}$$

Then we have an invariance between any two wave functions. We want to find a relation between $B$ and $B'$, which represent the action of the wave functions. Assume we know the wave function $\psi$ and therefore the $B$ associated with it. We can consider $S$ to be some unitary matrix

$$(\frac{\partial}{\partial x_\mu} - igB_\mu)\psi = S(\frac{\partial}{\partial x_\mu} - igB'_\mu)\tilde{\psi}$$
$$= S(\frac{\partial}{\partial x_\mu} - igB'_\mu)S^{-1}\psi$$
$$= SS^{-1}\frac{\partial\psi}{\partial x_\mu} + S\frac{\partial S^{-1}}{\partial x_\mu}\psi - igSB'_\mu S^{-1}\psi$$
$$= (\frac{\partial}{\partial x_\mu} + S\frac{\partial S^{-1}}{\partial x_\mu} - igSB'_\mu S^{-1})\psi$$

This gives

$$\begin{aligned}
\frac{\partial}{\partial x_\mu} - igB_\mu &= \frac{\partial}{\partial x_\mu} + S\frac{\partial S^{-1}}{\partial x_\mu} - igSB'_\mu S^{-1} \\
-igB_\mu &= S\frac{\partial S^{-1}}{\partial x_\mu} - igSB'_\mu S^{-1}) \\
B_\mu &= \frac{i}{g}S\frac{\partial S^{-1}}{\partial x_\mu} + SB'_\mu S^{-1}) \\
B'_\mu &= S^{-1}B_\mu S + \frac{1}{ig}\frac{\partial S^{-1}}{\partial x_\mu}S \\
&= S^{-1}B_\mu S - \frac{1}{ig}S^{-1}\frac{\partial S}{\partial x_\mu}
\end{aligned}$$

Where the last equality comes from

$$\begin{aligned}
S^{-1}S &= I \\
\frac{\partial}{\partial x_\mu}S^{-1}S &= \frac{\partial}{\partial x_\mu}I \\
\frac{\partial S^{-1}}{\partial x_\mu}S + S^{-1}\frac{\partial S}{\partial x_\mu} &= 0 \\
\frac{\partial S^{-1}}{\partial x_\mu}S &= -S^{-1}\frac{\partial S}{\partial x_\mu}
\end{aligned}$$

Then finally rewriting this gives:

$$B'_\mu = S^{-1}\big(B_\mu - \frac{1}{ig}\frac{\partial S}{\partial x_\mu}S^{-1}\big)S$$

# Appendix B

# Trace identities

$$
\begin{aligned}
\operatorname{tr}(A \wedge A \wedge A \wedge A) &= \operatorname{tr}(A_\mu A_\nu A_\rho A_\sigma dx_\mu \wedge dx_\nu \wedge dx_\rho \wedge dx_\sigma) \\
&= \operatorname{tr}(A_\sigma A_\mu A_\nu A_\rho dx_\mu \wedge dx_\nu \wedge dx_\rho \wedge dx_\sigma) \\
&= -\operatorname{tr}(A_\sigma A_\mu A_\nu A_\rho dx_\sigma \wedge dx_\mu \wedge dx_\nu \wedge dx_\rho) \\
&= -\operatorname{tr}(A \wedge A \wedge A \wedge A) \\
\implies \operatorname{tr}(A \wedge A \wedge A \wedge A) &= 0
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{tr}(A \wedge F \wedge A) &= \frac{1}{2} \operatorname{tr}(A_\mu F_{\nu\rho} A_\sigma) dx_\mu \wedge dx_\nu \wedge dx_\rho \wedge dx_\sigma \\
&= -\frac{1}{2} \operatorname{tr}(A_\sigma A_\mu F_{\nu\rho}) dx_\sigma \wedge dx_\mu \wedge dx_\nu \wedge dx_\rho \\
&= -\operatorname{tr}(A \wedge A \wedge F)
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{tr}(A \wedge F \wedge A) &= \frac{1}{2} \operatorname{tr}(A_\mu F_{\nu\rho} A_\sigma) dx_\mu \wedge dx_\nu \wedge dx_\rho \wedge dx_\sigma \\
&= -\frac{1}{2} \operatorname{tr}(F_{\nu\rho} A_\sigma A_\mu) dx_\nu \wedge dx_\rho \wedge dx_\sigma \wedge dx_\mu \\
&= -\operatorname{tr}(F \wedge A \wedge A)
\end{aligned}
$$