Utrecht University

Formalising local symmetries
An introduction to mathematical gauge theory

by

Thijs Douwes

A thesis

Submitted to the Departments of Mathematics and Physics
in partial fulfilment of the requirements
for the double degree of

Bachelor of Science

under the joint supervision of

Dr. Fabian Ziltener
Department of Mathematics

Prof. Thomas W. Grimm
Department of Physics

January 2023

Abstract

Many of our theories of nature contain objects that may be altered as a variable function of spacetime without changing the physical content of the theory—one example of this is the electromagnetic potential. Such objects are said to exhibit local symmetry under gauge transformations. Like any other type of symmetry, these local symmetries can be formalised mathematically. This is precisely the objective of mathematical gauge theory, which essentially studies the properties of spaces on which local symmetry is imposed geometrically.

In this exploratory work, a rigorous mathematical framework is laid out as a foundation in which to understand more advanced concepts of gauge theory. Beginning with the basics of Lie group theory, more abstraction is gradually added such as fibre bundles, vector and principal bundles, and connections and curvature on them. The level of abstraction culminates in a highly general formulation of the Yang–Mills functional for any compatible spacetime, metric and structure group.

The simple case of free space electromagnetism is used as a working example throughout, to serve as a model for the way in which physical theories arise out of relatively simple geometric structures. This includes a derivation of the homogeneous Maxwell equations from the second Bianchi identity, and a derivation of the free space inhomogeneous equations from the Yang–Mills functional.

Contents

1 Introduction **1**
1.1 Main results . . . 3
1.2 Overview . . . 4
1.3 Remarks on sources . . . 4
1.4 Prerequisites . . . 5
1.5 Notational conventions . . . 5

2 Lie groups **6**
2.1 Basic properties . . . 6
2.2 Lie algebras . . . 8
2.3 Adjoint representation . . . 12

3 Fibre bundles **14**
3.1 General fibre bundles . . . 14
3.2 Vector bundles . . . 16
3.2.1 Frame bundles . . . 18
3.3 Principal bundles . . . 19
3.3.1 Fundamental vector fields . . . 20
3.3.2 Gauges and their transformations . . . 21

4 Gauge fields **25**
4.1 Ehresmann connections . . . 25
4.2 Connections on G-bundles . . . 27
4.2.1 Gauge potentials . . . 30
4.3 Curvature . . . 33
4.3.1 Gauge field strength . . . 35

5 Yang–Mills theory **38**
5.1 Basic definitions . . . 39
5.2 Scalar products . . . 40
5.3 General Y–M theory . . . 42
5.4 Electromagnetism . . . 44
5.4.1 Closing remarks . . . 45

Chapter 1

Introduction

When James Clerk Maxwell finalised his theory of the electromagnetic field in 1864, he wrote down twenty different equations of just as many variables.[Max65, p486] Some 20 years later, Oliver Heaviside used the vector calculus of his own devising to simplify Maxwell's description down to just four equations in two unknowns. To date, the most familiar form of Maxwell's equations is a variant of Heaviside's formulation:[Gri17, p337]

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{1.4}$$

These equations were instrumental in the development of special relativity; indeed, Einstein's famous 1905 paper was titled *Zur Elektrodynamik bewegter Körper* ("On the electrodynamics of moving bodies"), and its second half is dedicated entirely to electromagnetism.[Ein05] In return, the relativistic formalism of four-vectors in Minkowski spacetime allows us to simplify the Maxwell equations even further:* the modern covariant formulation is[Sop08, p96]

$$\epsilon^{\mu\nu\rho\sigma} \partial_\nu F_{\rho\sigma} = 0, \tag{1.5}$$

$$\partial_\nu F^{\mu\nu} = J^\mu. \tag{1.6}$$

Here J^μ is the four-current (ρ, J_x, J_y, J_z); $\epsilon^{\mu\nu\rho\sigma}$ is the fully antisymmetric tensor for which $\epsilon^{0123} = 1$ and exchanging two indices induces a sign change; and $F^{\mu\nu}$ is the

*At this point, we move to natural units where $\mu_0 = \varepsilon_0 = c$ to ease the notational burden and avoid confusion.

electromagnetic field tensor given by

$$F^{\mu\nu} = \begin{Bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{Bmatrix}.$$

Equation (1.5) is the equivalent of the homogeneous equations (1.2) and (1.3), while equation (1.6) replaces the inhomogeneous equations (1.1) and (1.4).

It is common to introduce a four-potential $A^\mu(x)$ such that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.7}$$

In a sense, this equation is equivalent to (1.5): writing $F^{\mu\nu}$ in this form ensures that it always holds, and any $F^{\mu\nu}$ which satisfies (1.5) can be written as such. (We will demonstrate this equivalence in the course of this work.) Introducing A^μ also allows us to define the Lagrangian density

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + J_\mu A^\mu, \tag{1.8}$$

for which the Euler–Lagrange equations of motion are precisely equation (1.6).

All in all, the potential formulation is very useful and the formalism is elegant—but there is one kink in the cable. We have somewhat of a mathematical redundancy in the form of a *gauge freedom* for A^μ: if we add the gradient of a scalar field $\lambda(x)$ to A^μ such that

$$\overline{A}^\mu = A^\mu + \partial^\mu \lambda, \tag{1.9}$$

then $\overline{A}^\mu$ gives rise to the same field $F^{\mu\nu}$ as A^μ.

This gauge freedom can be thought of as a symmetry of the theory; more specifically, since changes to the four-potential are not necessarily universal across spacetime, we are dealing with a *local symmetry*—i.e. a symmetry that may manifest differently for each point in spacetime. Modern physics is rife with local symmetries, especially in the context of quantum field theory (QFT) and physics beyond the Standard Model (BSM).

While the gauge freedom described above is fairly benign and easy to ignore in a classical setting, in general local symmetries can be more complex and result in fields that are not invariant under their own gauge transformations. In these cases, it becomes much more necessary to develop a mathematical framework in which these symmetries are formalised. This formalism comes in the form of mathematical *gauge theory*, and it is one of the fundamental underpinnings of the very successful Standard Model of particle physics. Its basic premise is to associate to each point in spacetime a copy of the so-called *structure group* that describes a certain local symmetry. The objects that naturally arise on the resulting total spaces are called *gauge fields*, and their properties encode meaningful information about the corresponding physical system.

As noted before, electromagnetism (EM) is not the theory that benefits the most from being described as a gauge theory. However, its simple structure makes it a great stepping stone, and a place to draw examples from as we explore the mathematical theory in greater generality. As a result, we will keep our focus on EM throughout this work and mostly forego more advanced physics discussions in favour of greater mathematical substance.

1.1 Main results

The following are the main goals of this thesis:

- The first goal is to demonstrate that local group symmetries over a spacetime can be formalised in terms of a space called a principal bundle and the connections thereon. Specifically, we will see objects ω and Ω that look like the EM potentials and field tensor arise from the act of imposing a $\mathrm{U}(1) \cong S^1$ symmetry.

- Next, we will witness how geometric properties of this principal bundle give rise to physical information about the fields on it. Specifically, it is demonstrated that a geometric law called the second Bianchi identity

  $$\mathrm{D}\Omega = 0$$

  is the equivalent of the homogeneous Maxwell equation

  $$\epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0.$$

- Yang–Mills theory is formulated at a high level of abstraction, which results in the very general expression of the Yang–Mills functional

  $$\mathcal{S}_{YM}[\omega] = \frac{1}{2}\int_M \|\mathcal{F}\|^2 \epsilon.$$

  The critical points of this functional are shown to obey the Yang–Mills equation

  $$\mathrm{D} \star \mathcal{F} = 0.$$

- Finally, the Yang–Mills functional for $\mathrm{U}(1)$ symmetry over Minkowski spacetime is shown to be equivalent to the action for free space EM,

  $$\mathcal{S}_M = -\frac{1}{4}\int F_{\mu\nu}F^{\mu\nu}\,\mathrm{d}^4\mathbf{x}\,.$$

  As a result, the Yang–Mills equation is shown to be equivalent in this setting to the free space inhomogeneous Maxwell equation

  $$\partial_\nu F^{\mu\nu} = 0.$$

1.2 Overview

The rest of this thesis is organised as follows:

In Chapter 2, we begin the mathematical contents in earnest with a discussion of Lie groups. We start off with a quick refresher on basic properties, before moving on to discussions of the Lie algebra of a Lie group and the adjoint representation.

We then move on to a discussion of fibre bundles in Chapter 3. We begin in full generality, discuss vector bundles briefly and then move on to principal bundles, which are the main spaces that our gauge theories will take place in. We end off by formalising the notion of a gauge transformation.

In Chapter 4 we introduce the actual gauge fields themselves, which take on the form of connections and their curvature on principal bundles. We begin this chapter with a highly general description of connections on manifolds, and work our way up to connections on principal bundles and their curvature. We also discuss how to recover fields on the base spacetime from these objects, and finish by deriving the homogeneous Maxwell equations from their properties.

Finally, Chapter 5 is about the Yang–Mills theory behind defining a certain class of gauge invariant Lagrangian. To this end, we introduce some machinery such as metrics first, then focus our attention on different kinds of scalar products on spacetime before finally defining the Yang–Mills Lagrangian in great generality. We finish by deriving the inhomogeneous Maxwell equation in free space.

1.3 Remarks on sources

Mathematical gauge theory is an advanced and highly active area of research, with applications in some of our most successful physical theories. Put in plain terms, no stone has been left unturned at the undergraduate level. This is why, instead of sourcing recent research, this thesis is built almost exclusively on textbooks. It is the culmination of a learning process over the course of some months, and an attempt to convey the learnings at an undergraduate appropriate level. In some sense, the main focus of the writing is on didactic value; it strives to be the resource that its author would have wanted to have access to from the start.

The specific sources used are mainly detailed at the start of each following chapter, but the following texts have been some of the most influential:

- [Lee09] and its supplement [Lee12], for the basic underpinnings of Lie groups and fibre bundles;
- [Nab11] and [KMS93], for their treatment of the basics of connections and their curvatures;
- [Ham17] and [Fra12], for shaping a very mathematical approach to Yang–Mills theory.

1.4 Prerequisites

Mathematically, it is assumed that the reader is familiar with the very basics of differential geometry: this includes smooth manifolds, differential k-forms, the different differential operators etc. Some basic abstract algebra (especially group theory and some basic tensor algebra) is also assumed.

On the physics side, the reader is assumed to have some familiarity with the language of (classical) field theory, including Einstein summation, raising and lowering of indices, and contraction of indices. The Lagrangian formalism is also taken to be known.

1.5 Notational conventions

- Many of the definitions and theorems discussed hold for general C^k objects (and even topological spaces), but we will only concern ourselves with the smooth (C^∞) case; realistically, this is the only case of relevance to any physics we will be doing. Similarly, where possible we will only deal with spaces of finite dimension to avoid unnecessary complications.

- We use the notation $\subset$ for a subset and $\subsetneq$ for a proper subset.

- In what is perhaps a slight abuse of notation, we will always take $f|_A$ to be the restriction of a function f to the set $A \cap \mathrm{dom}(f)$, even when A is not completely contained the domain. Similarly, we will always take the domain of a composition $f \circ g$ to be the largest set such that $f \circ g$ is well defined. These conventions save superfluous notation without causing any real confusion.

- We define the wedge product of a differential k-form α and l-form β to be

  $$\alpha \wedge \beta = \frac{(k+l)!}{k!\,l!}\,\mathrm{Alt}(\alpha \otimes \beta),$$

  rather than the common alternative $\alpha \wedge \beta = \mathrm{Alt}(\alpha \otimes \beta)$. Both conventions have their advantages, but in our case we would like to have expressions such as

  $$\mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}x \otimes \mathrm{d}y - \mathrm{d}y \otimes \mathrm{d}x$$

  in order to simplify some of the results.

- Finally, on a more physical note, we will use the "mostly plus" signature $(-+++)$ for the Lorentzian metric.

Chapter 2

Lie groups

Group theory is the mathematical language of symmetries of all sorts, from $\mathbb{Z}/2\mathbb{Z}$ for basic mirror symmetry to SO(3) for rotational symmetry in 3-space. The local symmetries we deal with in physical gauge theories are continuous symmetries that manifest as a smooth function of spacetime. In these cases we want our symmetry groups to be smooth, which gives rise to the concept of a Lie group.

In this chapter, we give a quick refresher on Lie groups and introduce some features that are pertinent to the rest of this work. The subject matter comes broadly from [Lee09], but the discussion of Maurer–Cartan forms was partly inspired by [Nab11], and a few more advanced results are lifted from [Kna02].

2.1 Basic properties

Lie groups are smooth manifolds with group operations that behave smoothly. More formally:

Definition 2.1.1. A **Lie group** is a smooth manifold G that is also a group, where the group multiplication $G \times G \to G,\ (g, h) \mapsto gh$ is a smooth map.

Remark. Most definitions of a Lie group require the inversion map $g \mapsto g^{-1}$ to also be smooth, but in fact (surprisingly!) this is implied by the smoothness of the multiplication.[Nab11, Lemma 5.8.1]

Example 2.1.2. The real numbers form a Lie group $G = (\mathbb{R}, +)$ under addition: $\mathbb{R}$ is trivially a smooth manifold, and its group multiplication (addition) is clearly smooth.

In physical applications (especially in gauge theory), we are primarily interested in symmetries that may be represented by compact Lie groups. There is a fairly non-trivial result from the theory of Lie group representations that is of special interest in this case:

Theorem 2.1.3. [Kna02, Corollary 2.44]: *"Any compact Lie group G [...] is isomorphic to a closed linear group."*

Consequently, physicists prefer to work with matrix Lie groups in virtually all cases, as matrices tend to facilitate both manual and numerical calculations. Of course, from the physicists perspective this works the other way around—certain symmetries are found to be described by classes of matrices, which then turn out to be isomorphic to a compact Lie group.

We will strive not to restrict ourselves to matrix Lie groups only in this work. Still, whenever a result looks simpler (or even sufficiently different) from the matrix point of view, we will make note of this in a remark.

Example 2.1.4. The following are some matrix Lie groups of interest in physics:

- The complex **circle group**

$$S^1 = \{\, z \in \mathbb{C} : |z| = 1 \,\}$$

  is a Lie group under multiplication. This group is compact and connected. Examples of physical applications include rotations in two dimensions and the complex phase of a wave function.

- The real **orthogonal group** of degree n is the group of $n \times n$ matrices

$$\mathrm{O}(n) = \{\, Q \in \mathrm{GL}_n(\mathbb{R}) : Q^\mathsf{T} = Q^{-1} \,\}$$

  under matrix multiplication, where Q^T denotes the transpose of Q. For all $n \in \mathbb{N}^+$ this is a compact Lie group with two connected components: one with matrices of determinant 1, and the other with those of determinant -1. The former contains the identity of the group, and it constitutes a subgroup called the **special orthogonal group**:

$$\mathrm{SO}(n) = \{\, Q \in \mathrm{O}(n) : \det(Q) = 1 \,\}\,.$$

  This compact, connected Lie group may be used to represent rotations in n-dimensional Euclidean space, and as such we find that SO(2) is isomorphic to S^1 in particular.

- Similarly, the **unitary group** of degree n is the group of complex-valued $n \times n$ matrices

$$\mathrm{U}(n) = \{\, Q \in \mathrm{GL}_n(\mathbb{C}) : Q^\dagger = Q^{-1} \,\}\,,$$

  where $Q^\dagger$ denotes the conjugate transpose of Q. In contrast to $\mathrm{O}(n)$, this is a compact, connected Lie group. Note that U(1) is essentially equivalent to S^1, so that $S^1 \cong \mathrm{O}(2) \cong \mathrm{U}(1)$. As before, the matrices of unit determinant in $\mathrm{U}(n)$ form another Lie group called the **special unitary group**:

$$\mathrm{SU}(n) = \{\, Q \in \mathrm{U}(n) : \det(Q) = 1 \,\}\,.$$

An important property of Lie groups is that much of their structure is captured near the identity element $e \in G$. This is illustrated well by the following result:

Theorem 2.1.5. [Lee09, Theorem 5.6]: *"If G is a connected Lie group and U is a neighbourhood of the identity element e, then U generates the group. In other words, every element of G is a product of elements of U."*

Outline of proof. We choose an open subset V of U that is closed under the group inversion, and then define

$$V^n = \{ g_1 g_2 \cdots g_n \mid g_i \in V \} .$$

Then we can show that

$$V^\infty = \bigcup_{n\in\mathbb{N}} V^n$$

is a subgroup of G that is both open and closed, and so must be equal to G. □

2.2 Lie algebras

Theorem 2.1.5 hints strongly at the fact that Lie groups can be linearised in some way around $e \in G$ without losing too much of their structure. This is precisely what leads us to the concept of a Lie algebra: effectively, the Lie algebra of G is T_eG with some additional structure to preserve the group structure on G. Before we make this precise, we first introduce the concept of a Lie algebra in a more abstract setting.

Definition 2.2.1. A **Lie algebra** is a vector space $\mathfrak{g}$, equipped with an alternating bilinear map

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

called the **Lie bracket**, which satisfies the Jacobi identity:

$$\big[x,[y,z]\big] + \big[y,[z,x]\big] + \big[z,[x,y]\big] = 0 \qquad \forall x,y,z \in \mathfrak{g}.$$

Remark. Note that the combination of alternativity and bilinearity also implies that the Lie bracket is anti-commutative: we have

$$[x,y] + [y,x] = [x,x] + [x,y] + [y,x] + [y,y] = [x+y, x+y] = 0. \qquad \square$$

In what follows, it will be useful to know when different Lie algebras may be considered to be isomorphic:

Definition 2.2.2. Given two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ over a field $\mathbb{F}$, a **Lie algebra homomorphism** from $\mathfrak{g}$ into $\mathfrak{g}'$ is an $\mathbb{F}$-linear map $\varphi : \mathfrak{g} \to \mathfrak{g}'$ such that

$$\varphi([v,w]_{\mathfrak{g}}) = [\varphi(v), \varphi(w)]_{\mathfrak{g}'}$$

for all $v, w \in \mathfrak{g}$. If φ is also bijective, we call it a **Lie algebra isomorphism** and the Lie algebras are said to be isomorphic.

In some sense, the definition of a Lie bracket serves as a generalisation of the Lie bracket of vector fields on a manifold M, defined such that

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X.$$

Indeed, the set of vector fields $\mathfrak{X}(M)$ naturally has the structure of a vector space under addition. Since the Lie bracket of vector fields satisfies the Jacobi identity, it follows that $\mathfrak{X}(M)$ is a Lie algebra equipped with this bracket.

By extension, the space $\mathfrak{X}(G)$ of vector fields on a Lie group G is also a Lie algebra. However, $\mathfrak{X}(G)$ is infinite-dimensional and does not reflect the structure of G as a group. It turns out we can do better: we can obtain a much more interesting finite-dimensional Lie algebra by restricting our attention to vector fields that agree well with the group structure of G.

To make this more precise, consider the map of left* multiplication by a certain element $g \in G$:

$$L_g : G \times G, \quad h \mapsto gh.$$

This map is clearly a diffeomorphism from G to itself: it is smooth by the definition of a Lie group, and its inverse is $L_g^{-1} = L_{g^{-1}}$. This means L_g induces the standard push-forward operation on the vector fields on G: that is, we have the map

$$(L_g)_* : \mathfrak{X}(G) \to \mathfrak{X}(G)$$

such that

$$\mathcal{L}_{(L_g)_*(X)}(f)(gh) = \mathcal{L}_X(f \circ L_g)(h) \tag{2.1}$$

for all $h \in G$ and $f \in C^\infty(G)$. Using the general identity $\mathcal{L}_X(f)(h) = (\mathrm{d}f)_h(X_h)$ and the chain rule, this simplifies to

$$(L_g)_*(X)_{gh} = (\mathrm{d}L_g)_h(X_h) \quad \forall h \in G, \tag{2.2}$$

which is the usual formula for the push-forward of a diffeomorphism. Of special interest now are those vector fields which are invariant under this map:

Definition 2.2.3. A vector field $X \in \mathfrak{X}(G)$ on a Lie group is called **left invariant** when $(L_g)_*(X) = X$ for all $g \in G$. The set of all such vector fields is denoted $\mathfrak{X}^L(G)$.

One notable property of such vector fields is that they are fully determined by their element at the identity e of G: indeed, if we apply equation (2.2) with $h = e$ to the above definition, we find that

$$X_g = (L_g)_*(X)_g = (\mathrm{d}L_g)_e(X_e). \tag{2.3}$$

This means there is a one-to-one correspondence between vectors in T_eG and vector fields in $\mathfrak{X}^L(G)$: each $v \in T_eG$ induces a unique left invariant vector field, and two

*The choice between left and right multiplication in this discussion is one of convention, but it is the more natural one for our purposes: as we will see, the Lie algebra corresponds to the commutator of matrices in our case, but there would be a sign discrepancy if we proceeded using right multiplication.

distinct vector fields $X, Y \in \mathfrak{X}^L(G)$ must clearly have distinct vectors X_e and Y_e in T_eG. Moreover, since differentials are linear, we find for $X, Y \in \mathfrak{X}^L(G)$ and a scalar λ that

$$X_g + \lambda \cdot Y_g = (\mathrm{d}L_g)_e(X_e) + \lambda \cdot (\mathrm{d}L_g)_e(Y_e) = (\mathrm{d}L_g)_e(X_e + \lambda \cdot Y_e).$$

In other words, the correspondence between T_eG and $\mathfrak{X}^L(G)$ is properly an isomorphism of vector spaces. It follows that $\mathfrak{X}^L(G)$ has the same dimension as G, while also encoding its multiplicative behaviour. This makes it a strong candidate for our more interesting Lie algebra—all that's left to show is that $\mathfrak{X}^L(G)$ is closed under the Lie bracket operation:

Proposition 2.2.4. *For any two vector fields $X, Y \in \mathfrak{X}^L(G)$, the Lie bracket $[X, Y]$ is again left invariant.*

Proof. We make repeated use of equation (2.1) with an arbitrary test function $f \in C^\infty(G)$:

$$\begin{aligned}
\mathcal{L}_{(L_g)_*[X,Y]}(f)(gh) &= \mathcal{L}_{[X,Y]}(f \circ L_g)(h) \\
&= \mathcal{L}_X \mathcal{L}_Y (f \circ L_g)(h) - \mathcal{L}_Y \mathcal{L}_X (f \circ L_g)(h) \\
&= \mathcal{L}_X \mathcal{L}_{(L_g)_*(Y)}(f)(gh) - \mathcal{L}_Y \mathcal{L}_{(L_g)_*(X)}(f)(gh) \\
&= \mathcal{L}_X \mathcal{L}_Y (f)(gh) - \mathcal{L}_Y \mathcal{L}_X (f)(gh) \\
&= \mathcal{L}_{[X,Y]}(f)(gh).
\end{aligned}$$

Since this holds for arbitrary f and h, it follows that $(L_g)_*[X, Y] = [X, Y]$ as desired. □

Corollary 2.2.5. *The vector space $\mathfrak{X}^L(G)$ equipped with the Lie bracket of vector fields is a Lie algebra of dimension* $\dim(G)$.

Furthermore, because $\mathfrak{X}^L(G) \cong T_eG$ as vector spaces, we can also induce a Lie bracket on T_eG with the same structure. Writing $\vec{v}$ for the vector field induced by left translation on $v \in T_eG$, we have the following corollary:

Corollary 2.2.6. *The vector space T_eG with the bracket of $v, w \in T_eG$ defined by*

$$[v, w] := [\vec{v}, \vec{w}]_e$$

is a Lie algebra.

These isomorphic Lie algebras turn out to be invaluable in analysing the Lie group G, which leads us to the following definition:

Definition 2.2.7. Let G be a Lie group. Then **the Lie algebra of** G is the abstract Lie algebra that is isomorphic to those outlined in the previous corollaries. It is denoted using lowercase Fraktur letters, so that e.g. for $G = \mathrm{SU}(2)$ we notate $\mathfrak{g} = \mathfrak{su}(2)$ for the Lie algebra. In practice it is often most useful to identify $\mathfrak{g}$ with T_eG, so that we may write $v \in \mathfrak{g}$ when talking about vectors in T_eG.

The following are powerful examples to motivate the use of this definition:

Example 2.2.8. The Lie group $G = \mathrm{SO}(3)$ represents rotations in $\mathbb{R}^3$, and its Lie algebra $\mathfrak{so}(3)$ is precisely isomorphic to $\mathbb{R}^3$ with the Lie bracket given by the regular cross product of vectors. That is, $[v, w] := v \times w$ for $v, w \in \mathbb{R}^3$.

Example 2.2.9. A more general example is provided by the matrix Lie group $G = \mathrm{GL}_n(\mathbb{C})$ for some $n \in \mathbb{N}$. Since $\mathrm{GL}_n(\mathbb{C})$ is open in the set of all matrices $\mathcal{M}_n(\mathbb{C})$, its Lie algebra is $\mathfrak{gl}_n(\mathbb{C}) = T_I \, \mathrm{GL}_n(\mathbb{C}) = \mathcal{M}_n(\mathbb{C})$, and the Lie bracket turns out to be the normal commutator of matrices, i.e.

$$[A, B] = AB - BA$$

for $A, B \in \mathfrak{gl}_n(\mathbb{C})$. If G is any other Lie matrix group, we may view it as a subgroup of $\mathrm{GL}_n(\mathbb{C})$, and its Lie algebra is a subalgebra of $\mathcal{M}_n(\mathbb{C})$ with the same commutator bracket:

- For $\mathrm{U}(n)$ the Lie algebra consists of skew-Hermitian matrices:

$$\mathfrak{u}(n) = \{ A \in \mathcal{M}_n(\mathbb{C}) \mid A^\dagger = -A \} .$$

- For $\mathrm{SU}(n)$ these matrices must also be traceless:

$$\mathfrak{su}(n) = \{ A \in \mathcal{M}_n(\mathbb{C}) \mid A^\dagger = -A, \, \mathrm{tr}\, A = 0 \} .$$

- Restricting the previous groups to real entries we find that $\mathfrak{o}(n) = \mathfrak{so}(n)$ consists of the skew-symmetric matrices (which are always traceless).

Remark. The latter example has important implications for the Lie algebra of non-matrix groups as well: there is a more general result called Ado's Theorem which, while itself beyond the scope of this work, has the direct implication that any finite-dimensional Lie algebra is in fact isomorphic to a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ with the matrix commutator as a bracket.[Kna02, Appendix B3] This means we may always think of the Lie bracket as being a commutator in some sense.

In our discussion of gauge fields in later chapters, it will be useful to formalise the opposite operation to that in equation 2.3, i.e. sending a vector $X_g \in T_g G$ to it's unique corresponding vector in $T_e G \cong \mathfrak{g}$. For reasons that will become apparent later, it is useful to think of this operation as a 1-form:

Definition 2.2.10. Let G be a Lie group with Lie algebra $\mathfrak{g} = T_e G$. Then **the (left) Maurer–Cartan form of** G is defined as the map Θ such that $\Theta_g = (\mathrm{d}L_{g^{-1}})_g$. We think of Θ as a Lie algebra-valued 1-form and notate $\Theta \in \Omega^1(G; \mathfrak{g}) := \Omega^1(G) \otimes \mathfrak{g}$.

Remark. By nature, if X is a left-invariant vector field on G, then the Maurer–Cartan form of G sends X_g to the same vector $X_e \in \mathfrak{g}$ for all $g \in G$. Conversely, the unique vector field that is dual to the Maurer–Cartan form must also be left-invariant (and it is in some sense the most "canonical" vector field on G).

Example 2.2.11. The Lie algebra of $G = \mathbb{R}$ is $\mathbb{R}$ with a trivial Lie bracket. In this case, the Maurer–Cartan form Θ is simply the coordinate form $\mathrm{d}x$.

Example 2.2.12. When G is a matrix group, we can consider G as a submanifold of $\mathcal{M}_n(\mathbb{C})$; for $g \in G \subset \mathcal{M}_n(\mathbb{C})$ we may identify the vectors $v \in T_g G$ with elements of $\mathcal{M}_n(\mathbb{C})$ again. Under this identification, we can think of left multiplication by g^{-1} as a linear operation left unchanged by differentiation: we can then write

$$\Theta_g = \mathrm{d}L_{g^{-1}} = L_{g^{-1}} \circ \mathrm{d}g = g^{-1}\,\mathrm{d}g\,,$$

where $\mathrm{d}g$ is identified with the identity map on $\mathcal{M}_n(\mathbb{C})$.

In the case of $G = \mathrm{U}(1)$, we can write $g = z \in \mathbb{C} \cong \mathcal{M}_1(\mathbb{C})$ so that

$$\begin{aligned}
\Theta_g &= z^{-1}\,\mathrm{d}z \\
&= (z_1 - iz_2)(\mathrm{d}z_1 + i\,\mathrm{d}z_2) \\
&= z_1\,\mathrm{d}z_1 + z_2\,\mathrm{d}z_2 + i(z_1\,\mathrm{d}z_2 - z_2\,\mathrm{d}z_1) \\
&= i(z_1\,\mathrm{d}z_2 - z_2\,\mathrm{d}z_1),
\end{aligned} \quad \Big) \; z_1\,\mathrm{d}z_1 + z_2\,\mathrm{d}z_2 \;\textit{ on } \mathrm{U}(1)$$

which indeed takes values in the Lie algebra $\mathfrak{u}(1) = T_1 S^1 = i\mathbb{R}$. Alternatively, if we parametrize g as $\mathrm{e}^{i\theta}$, we have $\mathrm{d}g = i\mathrm{e}^{i\theta}\,\mathrm{d}\theta$, so that

$$\Theta_g = g^{-1}\,\mathrm{d}g = \mathrm{e}^{-i\theta} i \mathrm{e}^{i\theta}\,\mathrm{d}\theta = i\,\mathrm{d}\theta\,,$$

which is in fact independent of g.

2.3 Adjoint representation

Another, perhaps more natural, characterisation of the Lie bracket for a Lie group G arises when we start thinking about representations of G. Recall that a representation is a group homomorphism $\rho : G \to \mathrm{GL}(V)$ for some vector space V. The most natural vector space in the case of a Lie group G is its Lie algebra $\mathfrak{g} \cong T_e G$, and one obvious way to obtain a linear map $T_e G \to T_e G$ is to take the differential of a function $G \to G$ that fixes e. We wish to associate such a function $f_g : G \to G$ to each $g \in G$, in a way that makes

$$\rho : G \to \mathrm{GL}(\mathfrak{g}), \quad g \mapsto (\mathrm{d}f_g)_e$$

a representation. The trivial option is to use $f_g = \mathrm{id}_G$ for all $g \in G$, but this gives us the representation $g \mapsto \mathrm{id}_{\mathfrak{g}}$, which is not useful in that it does not elucidate the group structure of G. Perhaps the most natural non-trivial choice for f_g are the inner automorphisms of G, given by left conjugation by g:

$$f_g = C_g : G \to G, \quad h \mapsto ghg^{-1}.$$

These are all homomorphisms, with $C_e = \mathrm{id}_G$. They give rise to the following important representation of G:

Definition 2.3.1. The **adjoint representation of** G is the representation given by

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}), \quad g \mapsto \mathrm{Ad}_g := (\mathrm{d}C_g)_e.$$

Remark. In the case that G is a matrix Lie group, the left conjugation by g is a linear map, and so its differential still looks like left conjugation on the matrices of $\mathfrak{g}$: we have $\mathrm{Ad}_g(h) = gAg^{-1}$ for all $A \in \mathfrak{g}$.

This representation is important in and of itself, and encodes much of the group structure of G on its algebra. We will encounter it again later when exploring connections on principal bundles (or in physical terms, gauge fields). In this context, though, we can repeat the procedure above and differentiate Ad at the identity to obtain a map from $T_eG \cong \mathfrak{g}$ to $T_{\mathrm{id}_\mathfrak{g}}\mathrm{GL}(\mathfrak{g}) = \mathrm{GL}(\mathfrak{g})$ in order to obtain a so-called representation of the Lie algebra $\mathfrak{g}$—we will not expand on these representations here, but they have the expected definition. We find the following map:

Definition 2.3.2. The **adjoint representation of** $\mathfrak{g}$ is the map

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{GL}(\mathfrak{g}), \quad v \mapsto \mathrm{ad}_v := (\mathrm{d}\,\mathrm{Ad})_e(v).$$

This representation has the following very important property:

Theorem 2.3.3. $\mathrm{ad}_v(w) = [v, w]$ *for all* $v, w \in \mathfrak{g}$.

The proof of this theorem is beyond the scope of this work, but since ad arises so naturally as a representation of $\mathfrak{g}$, it may also be taken as an alternative definition of the Lie bracket of $\mathfrak{g}$:

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \quad (v, w) \mapsto \mathrm{ad}_v(w).$$

We take the more abstract stance here of insisting that $\mathfrak{g}$ and its bracket be the abstract Lie algebra isomorphic to either construction, but the interconnectedness of the theory does reinforce how fundamental the Lie algebra $\mathfrak{g}$ and its bracket are to any Lie group G.

Chapter 3

Fibre bundles

As discussed in the introduction, the basic idea of gauge theory is to associate a copy of the structure group G to each point in the spacetime M. However, we would be remiss to restrict our attention to the product space $M \times G$ unnecessarily. While our simple case of EM can be formulated on the product space just fine, non-trivial topologies can and do arise in more complex physical theories; in fact, these are of sufficient interest to have spawned a whole area of research called topological field theory.

The space we are looking for, then, is one that looks like a product space *locally*; this is precisely what motivates the concept of a fibre bundle. We introduce these in full generality first, and then add the structure of a Lie group. Finally, we explore what gauge transformations look like on these spaces.

The structure of this chapter follows [Lee09] for the most part, with some inspiration taken from [Nab11] for the section about gauge transformations.

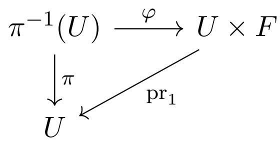
3.1 General fibre bundles

Informally, a fibre bundle is an extension of a base space M in the form of a fibre F, to create a total space E that is locally diffeomorphic to $M \times F$. More formally, we have the following definition:

Definition 3.1.1. A smooth **fibre bundle** is a quadruple (E, π, M, F) consisting of

- the **base space** M, the **(typical) fibre** F, and the **total space** E, all smooth manifolds;
- a smooth surjection $\pi : E \to M$, called the **bundle projection**,

and satisfying the following local triviality condition: for each $x \in M$, there exist an open neighbourhood U of x and a smooth diffeomorphism $\varphi : \pi^{-1}(U) \to U \times F$ (called a **local trivialisation**) such that the following diagram commutes:



In particular, the fibre bundle $(M \times F,\ \mathrm{pr}_1,\ M,\ F)$ is called the **trivial bundle**. The set $E_x := \pi^{-1}(x)$ is referred to as the **fibre over** x. Furthermore, parallel to the concept of charts and atlases on manifolds, we call a pair (U, φ) a **bundle chart**, and a family of bundle charts $\{(U_j, \varphi_j)\}_{j \in J}$ that covers the base space is called a **bundle atlas**.

We can be a bit more specific about the local trivialisations $\varphi : \pi^{-1}(U) \to U \times F$: by the diagram in the definition, we have $\mathrm{pr}_1 \circ \varphi = \pi|_{\pi^{-1}(U)}$. It is then useful to write $\varphi = (\pi, \Phi)$, where we make the restriction on π implicit for ease of notation. We call the map $\Phi : \pi^{-1}(U) \to F$ the **principal part** of φ. Since φ is a diffeomorphism, it follows that its principal part must be a diffeomorphism for each individual fibre—to be precise, $\Phi|_{E_p} : E_p \to F$ is a diffeomorphism for each $x \in U$.

An important concept in the physical application of fibre bundles is that of a section: intuitively, this constitutes a choice of a point in the fibre over each point in the base space.

Definition 3.1.2. A **(global) smooth section** of a fibre bundle (E, π, M, F) is a smooth map $\sigma : M \to E$ such that $\pi \circ \sigma = \mathrm{id}_M$. Similarly, a **local smooth section** over an open $U \subset M$ is a smooth map $\sigma : U \to E$ such that $\pi \circ \sigma = \mathrm{id}_M$. We will denote the set of smooth sections of a fibre bundle by $\Gamma(E)$.

Remark. The word "gauge" may also be used to refer to a local trivialisation; as we will see later, there is no real distinction in the case of physical gauge theories.

Depending on the topology of the total space, a fibre bundle may not admit any global sections.

There are several different concepts of morphisms that may be applied to fibre bundles, but especially relevant to us will be that of automorphisms:

Definition 3.1.3. Let (E, π, M, F) be a smooth fibre bundle. Then a **bundle automorphism** on E is a diffeomorphism $f : E \to E$ that preserves the fibres of E, i.e. such that $\pi \circ f = \pi$.

As we will see later, these automorphisms formalise the notion of gauge transformations.

In what follows we will discuss two important subclasses of fibre bundles: vector bundles and principal bundles. The principal bundle is ultimately where gauge theory takes place, but we will treat vector bundles first: partly because they illuminate how a fibre bundle can take on additional structure, and partly because they will provide a working example of a principal bundle, in the form of the frame bundle.

3.2 Vector bundles

Two important examples of fibre bundles are the tangent bundle TM and the cotangent bundle T^*M on a manifold M. In both these cases, each fibre takes on the structure of a vector space. This leads us to the concept of a vector bundle, which is essentially a fibre bundle where the typical fibre is a vector space, and each fibre of the total space takes on its structure.

In the following definition we take the field $\mathbb{F}$ to be either $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.2.1. A rank k smooth **$\mathbb{F}$-vector bundle** is a fibre bundle (E, π, M, V) where the typical fibre V is an $\mathbb{F}$-vector space of finite dimension k. Additionally, the bundle must admit a bundle atlas $\{(U_j, \varphi_j)\}_{j\in J}$ such that for each $j \in J$ and $x \in U_j$:

(i) the fibre E_x is an $\mathbb{F}$-vector space;

(ii) the map $\Phi_j|_{E_x} : E_x \to V$ is a vector space isomorphism where $\varphi_j = (\pi, \Phi_j)$.

The atlas $\{(U_j, \varphi_j)\}_{j\in J}$ is called (predictably) a **vector bundle atlas** and its elements are called **vector bundle charts**.

Example 3.2.2. The tangent and cotangent bundles of a smooth n-manifold M are rank n vector bundles with typical fibre $\mathbb{R}^n$ (or $(\mathbb{R}^n)^*$, the two are isomorphic as vector spaces). The smooth sections of the tangent bundle TM are precisely the vector fields on M (i.e. $\Gamma(TM) = \mathfrak{X}(M)$), and the smooth sections of the cotangent bundle are precisely the differential one-forms on M (i.e. $\Gamma(TM) = \Omega^1(M)$).

The vector space structure on the fibres of a vector bundle allows us to extend familiar linear operations to the context of bundles, such as that of choosing a subspace:

Definition 3.2.3. Let (E, π, M, V) be a smooth rank k vector bundle and let V' be an l-dimensional subspace of V. Suppose E' is a submanifold of E with the property that there exists an atlas of vector bundle charts (U, φ) such that $\operatorname{Im} \varphi|_{E'} = U \times V'$; then $(E', \pi|_{E'}, M, V')$ is called a smooth rank l **vector subbundle** of (E, π, M, V). The aforementioned bundle charts are said to be adapted to E'.

Subbundles of tangent bundles are important enough that we have a distinct name for them:

Definition 3.2.4. A smooth rank k **distribution** on a manifold M is the total space of a smooth rank k vector subbundle of the tangent bundle TM.

Distributions are the subject of much study in geometry, and we will see them make a return in our discussion of connections on fibre bundles in Chapter 4.

The additional structure on a vector bundle also allows us to choose sections of the bundle more readily. For instance, each vector bundle admits at least one global section (as opposed to general bundles), namely the **zero section** that assigns to each point x in the base space M the zero vector $0_x \in E_x$. Note that we do not need to choose a basis to do this, since the zero vector is a fixed element of each fibre. If we do fix a basis, either on the fibres E_x or on the typical fibre V, a richer structure emerges.

We refer to a basis of the fibre E_x over $x \in M$ as a **frame** at x. We may associate frames to neighbouring points on the base space in a smooth manner using sections of the bundle:

Definition 3.2.5. Let (E, π, M, V) be a rank k vector bundle, and let $U \subset M$ be an open set. A (local) **frame field** over U is a k-tuple $\sigma = (\sigma_1, \ldots, \sigma_k)$ of sections over U such that for all $x \in U$, $\big(\sigma_1(x), \ldots, \sigma_k(x)\big)$ is a frame at x.

We can go one step further: if we choose a fixed basis $\{\mathbf{e}_i\}_{1\leq i\leq k}$ for the typical fibre V, we can identify a frame at $x \in M$ with the unique linear isomorphism $V \to E_x$ that sends each $\mathbf{e}_i$ to the respective basis vector in the frame. Conversely, each such isomorphism represents a unique frame at x. But note that precisely such an isomorphism is contained within the definition of a vector bundle: if (U, φ) is a vector bundle chart with $\varphi = (\pi, \Phi)$, then $\Phi|_{E_x}^{-1} : V \to E_x$ is a linear isomorphism for each $x \in U$, and as such it gives rise to a frame at x. Moreover, since Φ is a smooth function, the collection of frames it induces over U is precisely equivalent to (the image of) a frame field. This means that in parallel to the correspondence between frames and linear isomorphisms, there is also a one-to-one correspondence between frame fields and local trivialisations. A similar concept will arise later in the context of principal bundles, which will be important to our understanding of gauge transformations.

One notable consequence of this correspondence between frame fields is that it gives rise to an equivalent definition of distributions:

Proposition 3.2.6. *Let M be a smooth n-manifold, and let $\Delta \subset TM$. Then the following statements are equivalent:*

(i) *Δ is a smooth rank k distribution on M.*

(ii) *There exists an open cover of M with the following property: for each open U in the cover, there exist k linearly independent vector fields $(X_1, \ldots, X_k) \in \mathfrak{X}(U)$ such that for all $x \in U$, the vectors $X_i(x)$ span the subspace $\Delta_x := \Delta \cup T_xM$.*

Proof. We prove the implication $(i) \implies (ii)$; the other direction is similar.

Let the typical fibre of TM be $V \cong \mathbb{R}^n$ and let Δ be a rank k distribution on M. Then there is a rank k subbundle of TM with a k-dimensional subspace $V' \subset V$ as the typical fibre. We choose an ordered basis for V such that the first k basis vectors form a basis of V'; in other words, if we identify V with $\mathbb{R}^n$, we identify V' with $\mathbb{R}^k \times \{0\} \subset \mathbb{R}^n$ via a change of basis.

By Definition 3.2.3, there is an atlas of bundle charts (U, φ) adapted to Δ, whose opens form a cover of M. Using our preferred basis, we can identify the chart (U, φ) with a frame field over U, the first k sections in which are also sections of Δ by inclusion. But sections of TM are precisely smooth vector fields, and the fact that they form a frame at each point implies the linear independence of these vector fields. Since $\dim \Delta_x = k$ for all $x \in U$, it follows that the k vector fields that lie on Δ precisely span the spaces Δ_x. □

Remark. Some authors use a weaker, generalized definition of a distribution, in which the vector fields that span the subspaces are not taken to be linearly independent. We have no need for such generality in the present work.

3.2.1 Frame bundles

We can further leverage the equivalence between frames and linear isomorphisms to construct another important class of bundle from any vector bundle (E, π, M, V). Consider first the space of all frames at $x \in M$. Per our discussion above, this space is isomorphic to the space $\mathrm{GL}(V, E_x)$ of all linear isomorphisms $V \to E_x$. We can use this to define a space related to E:

$$F(E) := \bigsqcup_{x \in M} \mathrm{GL}(V, E_x),$$

where the square cup indicates a disjoint union, meaning each element of $F(E)$ is of the form (x, u) where u is (identified with) a frame at $x \in M$. This gives rise to a natural projection on the first factor:

$$\widetilde{\pi} : F(E) \to M, \quad (x, u) \mapsto x.$$

We now come to our central claim, namely that $F(E)$ defines another fibre bundle over M:

Proposition 3.2.7. *Any rank k $\mathbb{F}$-vector bundle (E, π, M, V) gives rise to a fibre bundle $\big(F(E),\, \widetilde{\pi},\, M,\, \mathrm{GL}(V)\big)$.*

Proof. To show this, we need to show that $F(E)$ is a smooth manifold, and that there exists a bundle atlas with local trivialisations of the form $\widetilde{\varphi} : \widetilde{\pi}^{-1}(U) \to U \times \mathrm{GL}(V)$. Conveniently, the latter implies the former: the principal part of such a trivialisation is a smooth diffeomorphism between the open $\widetilde{\pi}^{-1}(U) \subset F(E)$ and an open in $\mathrm{GL}(V)$, which we can canonically identify with $\mathbb{F}^{k \times k}$. This means any bundle atlas we find automatically induces a smooth structure on $F(E)$.

We can adapt a bundle atlas from the existing vector bundle atlas on E in the following way: if (U, φ) is a vector bundle chart on E with $\varphi = (\pi, \Phi)$, then we can simply define

$$\widetilde{\varphi} : \widetilde{\pi}^{-1}(U) \to U \times \mathrm{GL}(V), \quad (x, u) \mapsto \big(x,\, \Phi|_{E_x} \circ u\big). \tag{3.1}$$

Note that this is well defined since u is an isomorphism from V to E_x and vice versa for $\Phi|_{E_x}$. ∎

Definition 3.2.8. The fibre bundle $\big(F(E),\, \widetilde{\pi},\, M,\, \mathrm{GL}(V)\big)$ constructed above is called the **frame bundle** of E, and we usually denote it simply by $F(E)$. For a general smooth manifold M, we may also notate $F(M)$ for the frame bundle of the tangent bundle TM.

As we will see in the next section, the frame bundle of a vector bundle is just one motivating example for the more general concept of a principal bundle.

3.3 Principal bundles

We now arrive at the main object of interest for gauge theories in physics: namely that of a principal bundle. Informally, these are bundles with a Lie group as a typical fibre, which acts on the total space in a certain coherent way. In physics, these formalise the invariance of certain fields under gauge transformations geometrically: the base space then represents space-time, and the typical fibre is the symmetry group of the theory.

We will begin by studying the example of a frame bundle more closely. As we have seen, the typical fibre of the frame bundle $F(E)$ of a vector bundle (E, π, M, V) is that its typical fibre is $\mathrm{GL}(V)$, arising from the identification of frames at $x \in M$ with elements of $\mathrm{GL}(V, E_x)$. Whereas the spaces $\mathrm{GL}(V, E_x)$ have no natural group structure due to the lack of a canonical transformation and the inability to compose the elements, $\mathrm{GL}(V)$ itself *is* a Lie group with identity element id_V.

Crucially, if we define the local trivialisations $\widetilde{\varphi}$ as in equation (3.1) above, the structure of the frame bundle is determined not only by the structure of $\mathrm{GL}(V)$ as a manifold, but also by its group structure. Indeed, the inverse of $\widetilde{\varphi}$ is

$$\widetilde{\varphi}^{-1} : U \times \mathrm{GL}(V) \to \widetilde{\pi}^{-1}(U), \quad (x, g) \mapsto \big(x,\ \Phi|_{E_x}^{-1} \circ g\big).$$

Recall that $\Phi|_{E_x}^{-1}$ is itself a frame, so that we are clearly making use of an existing right action of g on the frame bundle, defined by

$$r : F(E) \times \mathrm{GL}(V) \to F(E), \quad \big((x, u), g\big) \mapsto (x, u) \cdot g = (x, u \circ g).$$

This action is clearly free (i.e. only $g = \mathrm{id}$ keeps a frame fixed) and its orbits are precisely the individual fibres of $F(E)$ (in other words, the action is transitive on the fibres of $F(E)$).

All of this motivates us to define a class of fibre bundle with a Lie group as a typical fibre, equipped with a similar action. We arrive at the following definition:

Definition 3.3.1. Let G be a Lie group. A **principal G-bundle** is a smooth fibre bundle (P, π, M, G) equipped with a smooth free right action of G on P such that

(i) The action preserves fibres, i.e. $\pi(p \cdot g) = \pi(p)$ for all $p \in P$ and $g \in G$.

(ii) There exists a bundle atlas $\{(U_j, \varphi_j)\}_{j \in J}$ such that the principal part of φ_j is G-equivariant for all $j \in J$. That is, if $\varphi_j = (\pi, \Phi_j)$, then

$$\Phi_j(p \cdot g) = \Phi_j(p)g$$

for all $p \in \pi^{-1}(U)$ and $g \in G$.

Note that the first requirement is needed to make the second well defined, since it ensures that $pg \in \pi^{-1}(U)$ for all $p \in \pi^{-1}(U)$. As expected, we refer to the atlas $\{(U_j, \varphi_j)\}_{j \in J}$ as a **principal bundle atlas** and to its components as **principal bundle charts**. The typical fibre G is referred to as the **structure group** of the bundle.

Remark. Some definitions require the action of G on P to act transitively on fibres, but this is implicit in our definition: indeed, if p_1 and p_2 are in the same fibre, then choosing $g := \Phi^{-1}(p_2)\Phi(p_1)$ gives $p_1 = p_2 \cdot g$. Note that it follows that the action is regular on fibres: i.e. for each $p_1, p_2 \in P_x$ for some $x \in M$ there is a unique $g \in G$ such that $p_1 = p_2 \cdot g$.

Example 3.3.2. The frame bundle $F(E)$ of a vector bundle (E, π, M, V), equipped with the action discussed above, is indeed a principal bundle, and the atlas we constructed on it is a principal bundle atlas: if $\widetilde{\varphi} = (\widetilde{\pi}, \widetilde{\Phi})$ as defined in equation (3.1), then we have

$$\widetilde{\Phi}\big((x,u)g\big) = \widetilde{\Phi}\big((x, u \circ g)\big) = \Phi|_{E_x} \circ u \circ g = \widetilde{\Phi}\big((x,u)\big) \circ g$$

as required.

As far as the triviality of principal G-bundles goes, it will be useful to take note of the following more advanced result:[Nab11, p222]

Theorem 3.3.3. *Any principal G-bundle over a contractible base space (i.e. a manifold with the homotopy type of a point) is trivial. In particular, any principal G-bundle over $\mathbb{R}^n$ is trivial.*

3.3.1 Fundamental vector fields

Much like we could extend operations from linear algebra to vector bundles, the structure present on a Lie group G extends to principal G-bundles. One prominent example is that of left invariant vector fields: in section 2.2, we found that each vector $v \in T_eG \cong \mathfrak{g}$ corresponds to a unique left invariant vector field $\vec{v}$, given for $g \in G$ by

$$\vec{v}_g = (\mathrm{d}L_g)_e(v),$$

where L_g represents left translation by g. In the case of a principal bundle (P, π, M, G), the right action of G on P allows us to define a similar left multiplication on G by elements of P:

$$\ell_p : G \to P, \quad g \mapsto p \cdot g.$$

The differential of this map at e can be used to associate elements $v \in \mathfrak{g}$ to vector fields again, this time on all of the total space:

Definition 3.3.4. Let (P, π, M, G) be a principal bundle and let ℓ_p be defined as above. Then the **fundamental vector field** $v^\# \in \mathfrak{X}(P)$ corresponding to $v \in \mathfrak{g}$ is given by

$$v^\#_p = (\mathrm{d}\ell_p)_e(v)$$

for $p \in P$. The space of all fundamental vector fields on P is denoted $\mathfrak{g}^\# \subset \mathfrak{X}(P)$.

It transpires that $\mathfrak{g}^\#$ has the same vector space structure as $\mathfrak{g}$, since $(\mathrm{d}\ell_p)_e$ is an injective linear map for all $p \in P$ (indeed, ℓ_p is an immersion of G into P). In fact, the Lie algebra bracket on $\mathfrak{g}$ also carries over to the usual Lie bracket for vector fields on $\mathfrak{g}^\#$:

Proposition 3.3.5. *Let (P, π, M, G) be a principal bundle and let $v, w \in \mathfrak{g}$. Then*

$$[v^\#, w^\#] = [v, w]^\#.$$

Proof. We proceed similarly to [LM87, App. 5, Prop. 3.8]: first, we note that for all $v \in \mathfrak{g}$ and $p \in P$, we have

$$v^\#_{pg} = (\mathrm{d}\ell_{pg})_e(v) = \big(\mathrm{d}(\ell_p \circ L_g)\big)_e(v) = \big((\mathrm{d}\ell_p)_g \circ (\mathrm{d}L_g)_e\big)(v) = (\mathrm{d}\ell_p)_g(\vec{v}_g);$$

in other words, $v^\#$ is ℓ_p-related to $\vec{v}$ for all $v \in \mathfrak{g}$ and $p \in P$. But recall that for any two pairs of f-related vector fields, their Lie brackets are also f-related; it follows that

$$[v^\#, w^\#]_{pg} = (\mathrm{d}\ell_p)_g\left([\vec{v}, \vec{w}]_g\right) = (\mathrm{d}\ell_p)_g\left(\overrightarrow{[v, w]}_g\right) = [v, w]^\#_{pg},$$

so that in particular for $g = e$ we find $[v^\#, w^\#]_p = [v, w]^\#_p$ for any $p \in P$, as desired. □

Corollary 3.3.6. *The space of all fundamental vector fields $\mathfrak{g}^\#$ is a Lie algebra isomorphic to $\mathfrak{g}$.*

This equivalence will become especially powerful in the context of connections in chapter 4. For now we turn our attention to another useful consequence of the group structure on a principal bundle.

3.3.2 Gauges and their transformations

Similar to our discussion of frame fields on vector bundles, we can establish a link between principal bundle charts and sections of a principal bundle, which will give rise to the concept of a gauge. Starting with a local trivialisation, we have a similar concept to the zero section on a vector bundle:

Definition 3.3.7. Let (P, π, M, G) be a principal bundle and let (U, φ) be a principal bundle chart. Then the local smooth section

$$\sigma_\varphi : U \to P, \quad x \mapsto \varphi^{-1}(x, e)$$

is called the **canonical section** associated to φ.

Conversely, if we have a local smooth section, we can obtain a local trivialisation as follows:

Proposition 3.3.8. *Let (P, π, M, G) be a principal bundle and let $\sigma : U \to P$ be a local smooth section. Then the map*

$$f_\sigma : U \times G \to \pi^{-1}(U), \quad (x, g) \mapsto \sigma(x) \cdot g$$

is a diffeomorphism and its inverse forms a principal bundle chart (U, f_σ^{-1}).

Proof. The smoothness of f_σ is obvious. Its bijectivity follows from the fact that the action of G on P is regular: for each $x \in U$ this implies that the map

$$G \to P_x, \quad g \mapsto \sigma(x) \cdot g$$

is a bijection. This means f_σ has an inverse; this inverse must be of the form $f_\sigma^{-1} = (\pi, \Phi)$ for some $\Phi : \pi^{-1}(U) \to G$ such that $\Phi\big(\sigma(x) \cdot g\big) = g$. We need to show that Φ is smooth and G-equivariant.

For G-equivariance, note that we may write any $p \in \pi^{-1}(U)$ in the form $p = \sigma(x) \cdot g$, so that for any $h \in G$ we have

$$\Phi(p \cdot h) = \Phi\big(\sigma(x) \cdot gh\big) = gh = \Phi\big(\sigma(x) \cdot g\big)h = \Phi(p)h.$$

For the smoothness of Φ, let $p \in \pi^{-1}(U)$ arbitrarily and let (U', φ') be a principal bundle chart with $p \in U'$ and $\varphi' = (\pi, \Phi')$. Then we can write $p = \sigma(x) \cdot g = (\sigma \circ \pi)(p) \cdot g$ to obtain

$$\Phi'(p) = \Phi'\big((\sigma \circ \pi)(p) \cdot g\big) = (\Phi' \circ \sigma \circ \pi)(p)g,$$

so that

$$\Phi(p) = g = (\Phi' \circ \sigma \circ \pi)(p)^{-1}\Phi'(p).$$

It follows that Φ is smooth in p. Since p was chosen to be arbitrary, smoothness follows on all of $\pi^{-1}(U)$. We conclude that (U, f_σ^{-1}) is indeed a principal bundle chart. □

Note that in the previous proposition, σ is precisely the canonical section associated to f_σ^{-1}; conversely, if (U, φ) is a principal bundle chart with canonical section σ_φ, then for all $x \in U$ and $g \in G$ we have

$$f_{\sigma_\varphi}(x, g) = \sigma_\varphi(x) \cdot g = \varphi^{-1}(x, e) \cdot g = \Phi_x^{-1}(e) \cdot g = \Phi_x^{-1}(g) = \varphi^{-1}(x, g)$$

so that $f_{\sigma_\varphi}^{-1}$ as defined above is precisely identical to the original trivialisation φ; that is, the canonical section associated to φ is unique to φ. This means that in the case of principal bundles, there is a precise one-to-one correspondence between principal bundle charts and local sections of the bundle. This is precisely why in the physics literature, either of the two may be referred to interchangeably as a gauge:

Definition 3.3.9 (Gauge)**.** Let (P, π, M, G) be a principal bundle, and let $U \subset M$ be an open set. Then a **(local) gauge** on U refers to either one of the following equivalent concepts:

(i) A principal bundle chart (U, φ);

(ii) A local section σ over U.

Specifically, we think of (U, φ) and σ as being the same gauge when σ is the canonical section associated to φ.

In physics, it is often most useful to think of gauges as local sections, and so we will most often denote a gauge as σ and identify it with a section. The physical intuition behind this is as follows:

Say we have a physical field theory, defined on some spacetime Σ (e.g. the Minkowski space $\Sigma \cong \mathbb{R}^4$), which is invariant with respect to gauge transformations following some structure group G (for example, the invariance of electromagnetism under adding the gradient of a scalar field to the 4-potential). We decide to incorporate this symmetry by formulating the theory on a principal G-bundle which has Σ as a base space. The fields of interest (which we will make precise in Chapter 4) are differential forms defined on the entirety of the total space P of this bundle. Should we wish to recover the fields on the original spacetime, we can pull these forms back along some section σ to obtain fields on Σ. A gauge transformation in our original theory then corresponds precisely to choosing a different section σ' and pulling back along it instead. This leads us to the following preliminary definition:

Definition 3.3.10. Let (P, π, M, G) be a principal bundle with a given local gauge on $U \subset M$. Then a **(local) gauge transformation** over U refers simply to a different choice of local gauge on U.

We can also define gauge transformations more globally using the concept of a bundle automorphisms given in Definition 3.1.3:

Definition 3.3.11. Let (P, π, M, G) be a principal bundle. A **global gauge transformation** or **principal bundle automorphism** over M is a bundle automorphism f on P that is G-equivariant, i.e. $f(pg) = f(p)g$ for all $p \in P$ and $g \in G$. The global gauge transformations form a group under composition which is called the **gauge group** of the bundle and denoted $\mathcal{G}(P)$.

The two concepts (local and global) of a gauge transformation are related as follows:

Lemma 3.3.12. *Let (P, π, M, G) be a principal bundle, let f be a global gauge transformation and let $U \subset M$ be the domain of a local gauge. Then the restriction $f|_{\pi^{-1}(U)}$ induces a local gauge transformation over U.*

Proof. To begin with, we recall that $\pi \circ f = \pi$, so that $f\big(\pi^{-1}(U)\big) = \pi^{-1}(U)$. This means if (U, φ) is a principal bundle chart, the map $\varphi' := \varphi \circ f$ is well defined. We claim that (U, φ') is again a principal bundle chart; for this we need to show that φ' is of the form (π, Φ') where Φ' is G-equivariant. We do this by writing $\varphi = (\pi, \Phi)$ to obtain

$$\varphi' = (\pi, \Phi) \circ f = (\pi \circ f, \Phi \circ f) = (\pi, \Phi \circ f);$$

the G-equivariance of $\Phi' = \Phi \circ f$ then follows from that of f and Φ. □

In fact, this also provides us with an alternate characterisation of a local gauge transformation over U: we may think of it as arising from a unique principal bundle automorphism over U, or more precisely, a global gauge transformation of the restricted principal bundle $\big(\pi^{-1}(U),\, \pi|_{\pi^{-1}(U)},\, U, G\big)$.

The simplest way to think about gauge transformations arises if we think of gauges primarily as sections (which we often will): in this case, the transformation between two local sections is given by a G-valued "transition function" for each point where the domains of the sections overlap. This leads us to a more usable definition for local gauge transformations:

Definition 3.3.13. Let (P, π, M, G) be a principal bundle, and let $\sigma_1 : U_1 \to P$ and $\sigma_2 : U_2 \to P$ be local sections over non-disjoint $U_1, U_2 \subset M$. Then **the (local) gauge transformation** from σ_1 to σ_2 is the smooth transition function $g_{12} : U_1 \cup U_2 \to G$ such that $\sigma_2 = \sigma_1 \cdot g_{12}$.

This notion of a gauge transformation is the most practically workable, and consequently the one that will see most use in later chapters. That said, it is easily seen to correspond closely to the other transformations introduced above.

We will explore the effects of gauge transformations on the objects of physical interest (the so-called "gauge fields") as we study them in greater detail in the next chapter.

Chapter 4

Gauge fields: Connections and curvature

If we attempt to differentiate arbitrary objects on a principal bundle with the usual derivatives, we quickly start running into problems: the result may not play nice with gauge transformations on the bundle. That is, differentiating before or after a gauge transformation may yield different results. To prevent this, we must restrict our attention to objects that play nice with the G-structure present on the bundle, and define differentiation in a way that respects the structure of the Lie algebra (i.e. the linearisation of G). Making this precise gives rise to the concept of connections, which (along with their derivatives) are precisely the objects of interest in gauge theory.

Most influential to this chapter is the treatment of gauge fields in [Nab11], combined with the didactic approach in [KMS93]: we will introduce connections from a distinctly geometric viewpoint initially, beginning on general fibre bundles and adding more layers of abstraction along the way. Cues were also taken from [Nak17] and [Lee12, Chapter 12].

4.1 Ehresmann connections

We begin with a notion of connections that works on any fibre bundle. This will help understand the most basic features of the concept first, before moving on to the additional structure on a principal bundle. It must be emphasised that some of the terminology introduced in this section will take on additional meaning later; the material here must be treated as a stepping stone towards the "true" concept of a connection on a principal bundle.

We start by introducing the notion of verticality: vectors are said to be vertical if they are tangent to the fibres of the bundle. More precisely:

Definition 4.1.1. Let (E, π, M, F) be a smooth fibre bundle, and let $p \in E$. Then a vector $v \in T_pE$ is said to be **vertical** if $\mathrm{d}\pi(v) = 0 \in T_{\pi(p)}M$. The collection of all vertical vectors $VE := \ker(\mathrm{d}\pi)$ is called the **vertical bundle**, and the space of vertical vectors tangent to p is denoted V_pE.

Lemma 4.1.2. *The vertical bundle VE is a distribution on E.*

Proof. We will use the characterisation of distributions using vector fields from Proposition 3.2.6. Let $\dim(F) = k$; then clearly, the subspaces V_pE must have dimension k for all $p \in E$. This means we are looking for k linearly independent vector fields on an open cover of E, with values in VE. The strategy is to find such vector fields on F and push them forward to E via $M \times F$. We proceed as follows:

Let (U, φ) be an arbitrary bundle atlas on (E, π, M, F) and let (U', χ) be an arbitrary smooth chart on F. Then all sets of the form $\varphi^{-1}(U \times U') \subset \pi^{-1}(U)$ clearly constitute an open cover of E.

For each $y \in U \times U'$, $(\mathrm{d}\varphi)_{\varphi^{-1}(y)}$ is a vector space isomorphism between $T_{\varphi^{-1}(y)}E$ and $T_y(U \times U')$ that takes vertical vectors to

$$V_y := \ker(\mathrm{d}\,\mathrm{pr}_1)_y \subset T_y(U \times U');$$

subsequently, $(\mathrm{d}\,\mathrm{pr}_2)_y|_{V_y}$ is a vector space isomorphism between V_y and $T_{\mathrm{pr}_2(y)}U'$. Now, all we need to do is note that the chart (U', χ) induces k linearly independent vector fields on U' in the obvious manner. Then, we push these vector fields forwards to E via

$$\left((\mathrm{d}\varphi)^{-1}\right)_y \circ \left((\mathrm{d}\,\mathrm{pr}_2)_y|_{V_y}\right)^{-1}$$

for each $y \in U \times U'$, and obtain k linearly independent vector fields on $\varphi^{-1}(U \times U')$ that take values in VE. □

Because a fibre bundle is only equipped with a projection in the direction of the fibres, there is no corresponding canonical definition for horizontal vectors. Instead, we refer to any distribution that is complementary to VE as horizontal:

Definition 4.1.3. Let (E, π, M, F) be a smooth fibre bundle. Then an **Ehresmann connection** is a choice of a distribution H on E such that $T_pE = H_p \oplus V_pE$ for all $p \in E$. Given an Ehresmann connection, we refer to the distribution H as the **horizontal bundle**, and to its elements as **horizontal vectors**.

Defining horizontal vectors as above allows us to write any vector $v \in T_pE$ as a sum of a vertical vector $v^V \in V_pE$ and a horizontal vector $v^H \in H_p$. This means we can now meaningfully define the missing projection map in the horizontal direction, in terms of tangent vectors:

$$\mathrm{pr}_H : TE \to VE, \quad v := v^H + v^V \mapsto v^H.$$

This map is linear on T_pE for all $p \in E$, and it is smooth by virtue of the smoothness of H as a distribution. This means we may properly interpret it as a smooth 1-form on E that takes values in VE; that is, $\mathrm{pr}_H \in \Omega^1(E; VE)$. This gives rise to the following general notion:

Definition 4.1.4. Let (E, π, M, F) be a smooth fibre bundle. Then a **(general) connection form** on the bundle is a smooth VE-valued 1-form $\theta \in \Omega^1(E; VE)$ that is also a projection, that is:

- $\theta_p : T_pE \to VE_p$ is surjective for all $p \in E$;
- θ works trivially on vertical vectors, i.e. $\theta \circ \theta = \theta$.

The key point is now this:

Proposition 4.1.5. *On a fibre bundle (E, π, M, F), there is a one-to-one correspondence between Ehresmann connections and general connection forms.*

Outline of proof. We have already seen that each horizontal bundle induces a connection form on the bundle (E, π, M, F). Conversely, given a connection form θ, we claim that $H^\theta := \ker(\theta)$ is a horizontal bundle:

By the surjectivity of θ_p, the subspaces H^θ_p have the same rank $\dim(E) - \dim(F)$ for all $p \in E$, and since $\mathrm{Im}(\theta) = VE$, H^θ must be complementary to VE. The smoothness of H^θ as a distribution then follows from the smoothness of θ as a form.

Finally, the uniqueness of connection forms corresponding to distinct horizontal bundles and vice versa can be easily established pointwise: at each point $p \in E$, each choice of a projection onto V_pE corresponds uniquely to a choice of a complementary subspace of T_pE. □

The equivalence between Ehresmann connections and fixing a general connection form means we have two ways of looking at connections. Both of these notions can be extended naturally to vector and principal bundles by imposing extra restrictions that ensure they respect the additional structure on the bundle. The main objects of interest for our description of gauge theories live on the principal bundle, so we will not concern ourselves with the case of vector bundles here.

4.2 Connections on principal bundles

Moving to principal bundles, the first thing to note is that the concept of the vertical bundle is already compatible with the group action. Indeed, pushing a vertical vector forward via the right action

$$r_g : P \to P, \quad p \mapsto p \cdot g$$

of some $g \in G$ will always return another vertical vector, simply because $\pi \circ r_g = \pi$ by definition.

This same property of G-equivariance does not hold automatically for horizontal bundles as introduced in the last section, but it makes sense to impose it in this context: pushing a horizontal vector forward via r_g should give another horizontal vector.

Definition 4.2.1. Let (P, π, M, G) be a principal bundle. Then a **(principal) Ehresmann connection** is a choice of a distribution H on P such that

(i) $T_pP = H_p \oplus V_pP$ for all $p \in P$;

(ii) $(\mathrm{d}r_g)_p(H_p) = H_{p\cdot g}$ for all $p \in P$ and $g \in G$.

We will still refer to H as the horizontal bundle in this context.

This additional restriction on Ehresmann connections translates naturally to a similar property of the general connection forms they induce:

Lemma 4.2.2. *Let (P, π, M, G) be a principal bundle. Then a principal Ehresmann connection H on P is equivalent to a general connection form $\theta \in \Omega^1(P; VP)$ such that θ is G-equivariant, in the sense that $\mathrm{d}r_g \circ \theta = \theta \circ \mathrm{d}r_g$ for all $g \in G$. More explicitly, we have*

$$(\mathrm{d}r_g)_p \circ \theta_p = \theta_{p\cdot g} \circ (\mathrm{d}r_g)_p$$

for all $p \in P$ and $g \in G$.

Remark. We leave the proof of this lemma for what it is, since—as will become apparent—general connection forms are just a stepping stone and ultimately unimportant to gauge theories.

It turns out we can do a bit better still; the structure of the principal bundle allows for a more powerful definition of connection forms in terms of the Lie algebra $\mathfrak{g}$. To set this up, we return to our discussion of fundamental vector fields from section 3.3.1. There, we found that the space of fundamental vector fields $\mathfrak{g}^\#$ is a Lie algebra isomorphic to $\mathfrak{g}$. It turns out we can also make a pointwise identification on P in the following way:

Lemma 4.2.3. *Let (P, π, M, G) be a principal bundle. Then the Lie algebra isomorphism $\mathfrak{g} \to \mathfrak{g}^\#$ induces a linear isomorphism $\mathfrak{g} \to V_pP$ at each point $p \in P$.*

Proof. Recall that for $v \in \mathfrak{g}$, the fundamental vector field $v^\#$ is given at $p \in P$ by $v^\#_p = (\mathrm{d}\ell_p)_e(v)$, where ℓ_p is the map $g \mapsto pg$. This means linearity is immediate, and we find

$$(\mathrm{d}\pi)_p(v^\#_p) = \big((\mathrm{d}\pi)_p \circ (\mathrm{d}\ell_p)_e\big)(v) = \big(\,\mathrm{d}(\pi \circ \ell_p)\,\big)_e(v)$$

for all $v \in \mathfrak{g}$; but $(\pi \circ \ell_p)(g) = \pi(p \cdot g) = \pi(p)$ is constant with respect to g, so that $(\mathrm{d}\pi)_p(v^\#_p) = 0$. In other words, $v^\#_p \in V_pP$ for all $v \in \mathfrak{g}$ and $p \in P$.

Finally, since ℓ_p is an immersion, $(\mathrm{d}\ell_p)_e$ is injective; combined with the fact that $\dim(V_pP) = \dim(G)$, we conclude that $(\mathrm{d}\ell_p)_e$ is indeed an isomorphism between $\mathfrak{g}$ and V_pP.* □

It is through this isomorphism that we finally arrive at the main object of interest: if we compose the general connection form $\theta \in \Omega^1(P; VP)$ with the inverse isomorphism $V_pP \to \mathfrak{g}$ at each point, we obtain a 1-form ω that takes values in the Lie algebra everywhere, i.e. $\omega \in \Omega^1(P; \mathfrak{g})$. We will look at a precise definition first, and then demonstrate equivalence with the principal Ehresmann connection H.

*Note that this proof relies implicitly on the finite dimension of G—we are not concerned with the infinite dimensional case here.

Definition 4.2.4 (Gauge field)**.** Let (P, π, M, G) be a principal bundle. Then a **(principal) connection form** or **gauge field** on P is a Lie algebra-valued 1-form $\omega \in \Omega^1(P; \mathfrak{g})$ such that

(i) $\omega_p(v_p^{\#}) = v$ for all $p \in P$ and $v \in \mathfrak{g}$;

(ii) $r_g^*\omega := \omega \circ \mathrm{d}r_g = \mathrm{Ad}_{g^{-1}} \circ \omega$ for all $g \in G$, where Ad is the adjoint representation from section 2.3.

Remark. This interpretation of a connection on a principal bundle is so prevalent and so useful that we will often refer to it simply as a **connection**. Unless otherwise stated, whenever we talk about a connection, we will be referring to a gauge field.

Proposition 4.2.5. *Let $\omega \in \Omega^1(P; \mathfrak{g})$ be a gauge field on the principal bundle (P, π, M, G). Then $H := \ker \omega$ is a principal Ehresmann connection on P.*

Proof. We proceed as [Lee12, Theorem 12.11]:

Let $X_p \in T_pP$ for some $p \in P$. Then $X_p - \omega_p(X_p)_p^{\#} \in H_p$, since

$$\omega_p\big(X_p - \omega_p(X_p)_p^{\#}\big) = \omega_p(X_p) - \omega_p(X_p) = 0.$$

But $\omega_p(X_p)_p^{\#} \in V_pP$, so that we can write

$$X_p = \big(X_p - \omega_p(X_p)_p^{\#}\big) + \omega_p(X_p)_p^{\#}$$

So that $T_pP = H_p + V_pP$. But if X_p is both vertical and horizontal (in H), then we have $X_p = \omega_p(X_p)_p^{\#} = 0_p^{\#} = 0$, so that $T_pP = H_p \oplus V_pP$. Now let $X_p \in H_p$. Then

$$(r_g^*\omega)_{p\cdot g}(X_p) = \mathrm{Ad}_{g^{-1}}\, \omega_p(X_p) = 0,$$

so that $(\mathrm{d}r_g)_p(X_p) \in H_{p\cdot g}$. Finally, the desired equality $(\mathrm{d}r_g)_p(H_p) = H_{p\cdot g}$ follows from $\dim(H_p) = \dim(H_{p\cdot g})$. □

We are now ready to take our first cursory dive into the fields for electromagnetism:

Example 4.2.6. The simplest form of classical free space electromagnetism can be formulated on the trivial bundle $(\mathbb{R}^4 \times \mathbb{R}, \mathrm{pr}_1, \mathbb{R}^4, \mathbb{R})$. The Lie algebra of $G = \mathbb{R}$ is again $\mathbb{R}$, so that the connections on this bundle are usual real-valued forms. If we denote an element of $P = \mathbb{R}^4 \times \mathbb{R}$ as (x, y), then a general smooth 1-form on P can be written as

$$\omega = A_\mu(x, y)\, \mathrm{d}x^\mu + f(x, y)\, \mathrm{d}y\,,$$

where $A_\mu, f \in C^\infty(P)$ and summation over μ is implied. To make ω into a gauge field, we impose the restrictions from 4.2.4: firstly, property (i) simply implies that $f(x, y) = 1$, since vertical vectors are unaffected by $\mathrm{d}x^\mu$. Secondly, since $\mathbb{R}$ is abelian, $\mathrm{Ad}_g = \mathrm{id}_{\mathbb{R}}$ for each $g \in \mathbb{R}$. This means property (ii) reduces to $r_g^*\omega = \omega$; in other words, ω must be independent of y. All things considered, a gauge field on P must be of the following form:

$$\omega = A_\mu(x)\,\mathrm{d}x^\mu + \mathrm{d}y\,.$$

The first term is essentially the obvious extension of any smooth 1-form $A_\mu(x)\,\mathrm{d}x^\mu \in \Omega^1(\mathbb{R}^4)$ to P; as we will see in the next subsection, this form on the base space corresponds to one possible EM four-potential A_μ. Importantly, the second term $\mathrm{d}y$ can also be viewed as an extension of the Maurer–Cartan form $\Theta = \mathrm{d}y \in \Omega^1(\mathbb{R})$ as introduced in Definition 2.2.10.

Example 4.2.7. More commonly, electromagnetism is formulated on a principal U(1)-bundle. The connections on this bundle take values in $\mathfrak{u}(1) \cong i\mathbb{R}$, meaning the purely gauge theoretic formulation of EM takes on an extra factor i compared to the classical formulation. Let us take a look at the trivial bundle with $P = \mathbb{R}^4 \times \mathrm{U}(1)$ in particular. If we parametrize P as $p = (x, \mathrm{e}^{i\theta})$ for $p \in P$, we can use a similar argument to the previous example to arrive at the general gauge field

$$\omega = iA_\mu(x)\,\mathrm{d}x^\mu + i\,\mathrm{d}\theta\,,$$

where the term $i\,\mathrm{d}\theta$ once again takes on the shape of the Maurer–Cartan form Θ, this time from Example 2.2.12. This is effectively equivalent to the last example up to a factor i, and it would seem that the choice between $G = \mathbb{R}$ and $G = \mathrm{U}(1)$ is a rather arbitrary one. This is indeed largely the case from a classical point of view, but we will see that having a compact group is rather useful in defining the Lagrangian in the next chapter. There are also more physical motivations to go for U(1), mostly from quantum mechanics, but we will not be delving into them—suffice it to say we will be working with U(1) wherever possible.

These examples are suggestive of the importance of the Maurer–Cartan form Θ of G in the construction of gauge fields. This should come as no surprise: if we extend $\Theta \in \Omega^1(M;\mathfrak{g})$ to $\widetilde{\Theta} \in \Omega^1(M \times G;\mathfrak{g})$ in the obvious way on a trivial bundle, then for $p \in P$, then $\widetilde{\Theta}_p$ is precisely the linear isomorphism $V_pP \to \mathfrak{g}$ that we used to arrive at the definition of a gauge field. It follows that $\widetilde{\Theta}$ is always a connection on any trivial bundle. This is significant, in part because principal bundles over $\mathbb{R}^n$ are always trivial by Theorem 3.3.3. Moreover, as we will see in the next subsection, the picture is not necessarily much more complicated on a more general non-trivial bundle.

4.2.1 Gauge potentials

In physical gauge theories, the base space of the principal bundle most often represents the spacetime in which the physics take place. It is therefore often useful to fix a (local) gauge (in the sense of section 3.3.2), in order to obtain a physical object on the base space:

Definition 4.2.8. Let (P, π, M, G) be a principal bundle, let ω be a gauge field on P and let σ be a local section over $U \subset M$. Then the pull-back $\mathcal{A}_\sigma = \sigma^*\omega \in \Omega^1(M;\mathfrak{g})$ is called a **(local) gauge potential** on M.

Example 4.2.9. Continuing from Example 4.2.7, let $\omega = iA_\mu(x)\,\mathrm{d}x^\mu + i\,\mathrm{d}\theta$ be a gauge field on $P = \mathbb{R}^4 \times \mathrm{U}(1)$. Then we can pull ω back along the global canonical section $\sigma : x \mapsto (x, 1)$ in order to obtain the gauge potential $\mathcal{A} = iA_\mu(x)\,\mathrm{d}x^\mu \in \Omega^1\big(\mathbb{R}^4, \mathfrak{u}(1)\big)$. Note that the EM potential has an extra factor i in the gauge description: we have $\mathcal{A}_\mu = iA_\mu$.

Now, we will take a closer look at how general gauge potentials transform under gauge transformations, as given in Definition 3.3.13.

Proposition 4.2.10. *Let (P, π, M, G) be a principal bundle, let ω be a gauge field on P and let $\sigma_1 : U_1 \to P$ and $\sigma_2 : U_2 \to P$ be local sections over non-disjoint subsets $U_1, U_2 \subset M$, related by the local gauge transformation g_{12} on $U_1 \cup U_2$. Then the local gauge potentials $\mathcal{A}_1 = \sigma_1^*\omega$ and $\mathcal{A}_2 = \sigma_2^*\omega$ are related by*

$$\mathcal{A}_2 = \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{A}_1 + \Theta_{12},$$

where $\Theta_{12} = g_{12}^*\Theta$ and Θ is the Maurer–Cartan form for G.

Remark. We defer to [Lee12, Theorem 12.13] for the proof.

When G is a matrix Lie group, we have

$$\mathrm{Ad}_{g_{12}^{-1}} \mathcal{A}_1 = g_{12}^{-1}\mathcal{A}_1 g_{12}$$

as per the remark under Definition 2.3.1, and

$$\Theta_{12} = g_{12}^*\Theta = \Theta_{g_{12}} \circ \mathrm{d}g_{12} = g_{12}^{-1}\,\mathrm{d}g_{12}$$

as per Example 2.2.12, so that the transformation law reads

$$\mathcal{A}_2 = g_{12}^{-1}\mathcal{A}_1 g_{12} + g_{12}^{-1}\,\mathrm{d}g_{12}\,.$$

Before we interpret this result physically, we return briefly to the question of constructing explicit gauge fields on non-trivial G-bundles. This is done using the reverse process of taking gauge potentials; we begin with a set of local 1-forms covering the base space that "look like" gauge potentials in some sense, and reverse-engineer a gauge field that is compatible with them. For this, we use the following fairly intuitive result (presented here without proof, see e.g. [Nab11, Theorem 6.1.1]):

Theorem 4.2.11. *Let (P, π, M, G) be any principal bundle, and let $\{(U_j, \varphi_j)\}_{j\in J}$ be a principal bundle atlas on it. Suppose that we have a 1-form $\mathcal{A}_j \in \Omega^1(U_j; \mathfrak{g})$ for all $j \in J$, such that the different $\mathcal{A}_j$ satisfy the transformation law in Proposition 4.2.10 whenever their domains overlap. Then there is a unique gauge field ω on P such that each $\mathcal{A}_j$ is precisely the gauge potential $\sigma_j^*\omega$, where σ_j is the canonical section associated to φ_j.*

The real usefulness of this theorem presents itself when we write the canonical section $\sigma_j : M \to P$ as $\sigma_j = \varphi_j^{-1} \circ \sigma_0$, where $\sigma_0 : M \to M \times G$ is simply the global canonical section $x \mapsto (x, e)$ of the trivial bundle. But we have already seen from previous examples how to construct a connection on the product bundle: we simply take any smooth 1-form $\mathcal{A} \in \Omega^1(U_j; \mathfrak{g})$ and add the Maurer–Cartan form Θ, both trivially extended to the product space. The above theorem then ensures that there is a compatible connection ω on P, such that $\omega|_{\pi^{-1}(U_j)} = \varphi_j^*(\mathcal{A} + \Theta)$. Locally, then, we are free to choose any 1-form on the base space to work with, as long as its domain is covered by a single principal bundle chart for P.

Example 4.2.12. Returning to our pet example of $P = \mathbb{R}^4 \times \mathrm{U}(1)$, global sections σ of the bundle can be identified with smooth functions $f \in C^\infty(\mathbb{R}^4)$ up to addition of 2π, via $\sigma(x) = (x, \mathrm{e}^{if(x)})$. This means that when we pick two global sections σ_1 and σ_2, we find that the transition function satisfies

$$\left(x, \mathrm{e}^{if_2(x)}\right) = \sigma_2(x) = \sigma_1(x) \cdot g_{12}(x) = \left(x, \mathrm{e}^{if_1(x)} g_{12}(x)\right),$$

so that $g_{12} = \mathrm{e}^{i(f_2 - f_1)}$. Writing $f_2 = f_1 + \lambda$, it becomes clear that g_{12} may be identified with any smooth function $\lambda \in C^\infty(\mathbb{R}^4)$, again up to addition of 2π. Furthermore, because $\mathbb{R}$ is abelian the adjoint representation Ad acts trivially. This means the transformation law reduces in this case to

$$\mathcal{A}_2 = \mathcal{A}_1 + \mathrm{e}^{-i\lambda}\,\mathrm{d}(\mathrm{e}^{i\lambda}) = \mathcal{A}_1 + i\,\mathrm{d}\lambda\,,$$

which—because the addition of 2π to λ leaves $\mathrm{d}\lambda$ unchanged—is precisely equivalent to the familiar gauge freedom $\overline{A}^\mu = A^\mu + \partial^\mu \lambda$ from classical electromagnetism (equation (1.9)).

From this example, it becomes clear that we can use just one gauge field on a principal bundle to represent all different electromagnetic four-potentials that are equivalent to a certain potential A^μ. The converse is not true, however: seemingly different connections on the bundle may yield the same set of gauge potentials when pulled back along different sections.

Example 4.2.13. In the case of $P = \mathbb{R}^4 \times \mathbb{R}$, the gauge potential $\mathcal{A}_\sigma = \mathrm{d}f$ for $f \in C^\infty(\mathbb{R}^4)$ may be obtained both by pulling $\omega = \mathrm{d}f + \mathrm{d}g$ back along the zero section $\sigma_0 : x \mapsto (x, 0)$ (where $\mathrm{d}f$ is extended to the total space in the obvious way), and by pulling $\omega' = \mathrm{d}g$ back along the section $\sigma : x \mapsto (x, f(x))$.

In general this is the case precisely when two different connections ω and ω' are related by a global gauge transformation $f \in \mathcal{G}(P)$ (as in Definition 3.3.11), i.e. $\omega' = f^*\omega$. In this case ω and ω' are said to be **gauge equivalent**. Since equivalent gauge fields yield the same gauge potentials, we may consider them “isomorphic” in a way: if we take a connection to be a fixed part of the structure of the bundle, moving to an equivalent gauge is a natural result of transforming the whole bundle by a bundle automorphism.

More precisely, gauge equivalence defines an equivalence relation on the set $\mathcal{A}(P)$ of all connections on P, with the equivalence classes being elements of the set $\mathcal{A}(P)/\mathcal{G}(P)$ (called the **moduli space of connections**). These equivalence classes can be thought of as the real fundamental objects in a physical theory, since they uniquely represent different configurations of the fields.

4.3 Curvature

Now that we have established the concept of a connection on a principal bundle, it becomes possible to define a way of differentiating forms on the bundle in a way that respects the group structure that is present. This is important in order to maintain G-equivariance of forms, and in the general context, it will be crucial to formulating Lagrangian densities that are invariant under gauge transformations. For this derivative, we will make use of the fact that a connection allows us to split a tangent vector up into horizontal and vertical components:

Definition 4.3.1. Let (P, π, M, G) be a principal bundle equipped with a connection ω, which induces a horizontal distribution $H \subset TP$. Then the **covariant exterior derivative** $\mathrm{D}\eta$ of a k-form η on P is defined for all $p \in P$ as

$$(\mathrm{D}\eta)_p(v_1, \ldots, v_k) := (\mathrm{d}\eta)_p(v_1^H, \ldots, v_2^H),$$

where v^H denotes the horizontal component of $v = v^H + v^V \in T_pP = H_p \oplus V_pP$.

We may think of this derivative as enforcing the structure of the gauge field onto higher order objects. Perhaps the most obvious such object to consider is the covariant derivative of the gauge field itself:

Definition 4.3.2. Let (P, π, M, G) be a principal bundle equipped with a connection ω. Then the **curvature** of ω is the $\mathfrak{g}$-valued 2-form $\Omega \in \Omega^2(P; \mathfrak{g})$ defined by

$$\Omega_p(v, w) = (\mathrm{D}\omega)_p(v, w) = (\mathrm{d}\omega)_p(v^H, w^H)$$

for $p \in P$. A connection is said to be **flat** if its curvature is identically zero.

Remark. Since Ω is defined in terms of horizontal components of vectors, we always have $\Omega_p(v, w) = \Omega_p(v^H, w^H)$. Note also the analogy between *all* connections sending horizontal vectors to zero (i.e. $\omega_p(v^H) = 0$) and *flat* connections having a differential that sends horizontal vectors to zero (i.e. $(\mathrm{d}\omega)_p(v^H, w^H) = 0$).

Example 4.3.3. If P is the trivial bundle $M \times G$, then the extension $\widetilde{\Theta}$ of the Maurer–Cartan form to P is flat: it induces horizontal vectors that are parallel to M in the total space, while being constant with respect to $x \in M$. It follows that any non-zero curvature in a trivial bundle results from adding the extension of a 1-form $\mathcal{A}$ on the base space to $\widetilde{\Theta}$.

An important result for the explicit computation of curvature forms comes in the form of Cartan's structure equation. In order to formulate this equation, we need to introduce some notation: we would like to define a wedge operation on $\mathfrak{g}$-valued forms on P. Since the Lie algebra does not necessarily have a product defined on it, we make use of the bracket on $\mathfrak{g}$; essentially, for two such forms η and θ we write $[\eta \wedge \theta]$ to mean the regular wedge product, but involving the Lie bracket on $\mathfrak{g}$ instead of multiplication in $\mathbb{R}$. For example, if η and θ are 1-forms, we want

$$[\eta \wedge \theta]_p(v, w) = [\eta_p(v), \theta_p(w)] - [\eta_p(w), \theta_p(v)].$$

One very natural way to formalise this is to make use of $\Omega(P;\mathfrak{g}) := \Omega(P) \otimes \mathfrak{g}$, and write $\eta = \alpha \otimes X$ and $\theta = \beta \otimes Y$ for $\alpha, \beta \in \Omega(P)$ and $X, Y \in \mathfrak{g}$. We can then define $[\eta \wedge \theta] := \alpha \wedge \beta \otimes [X, Y]$. This also makes it fairly easy to see that this operation satisfies the same graded Leibniz rule as the regular wedge product: if η is a k-form, then

$$\mathrm{d}[\eta \wedge \theta] = [\mathrm{d}\eta \wedge \theta] + (-1)^k[\eta \wedge \mathrm{d}\theta].$$

This brings us to the following result:

Theorem 4.3.4 (Cartan's structure equation)**.** *Let (P, π, M, G) be a principal bundle equipped with a connection ω, and let Ω be the curvature of ω. Then the curvature can be written as*

$$\Omega = \mathrm{d}\omega + \frac{1}{2}[\omega \wedge \omega].$$

Remark. This equation is so useful and ubiquitous that it is often used as an equivalent definition of the curvature Ω. The proof of this identity is fairly simple but laborious, so we will not include it here.

Example 4.3.5. Let $P = \mathbb{R}^4 \times \mathrm{U}(1)$ again, equipped with the most general gauge field $\omega = iA_\mu(x)\,\mathrm{d}x^\mu + i\,\mathrm{d}\theta$ from Example 4.2.7. In this case, since $\mathrm{U}(1)$ is abelian, the Lie bracket on $\mathfrak{u}(1)$ is trivial, so that Cartan's structure equation reduces to $\Omega = \mathrm{d}\omega$. Writing this out coordinate-wise we find

$$\Omega = i\,\partial_\nu A_\mu\,\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu,$$

using Einstein summation as usual; note that the $\mathrm{d}\theta$ term in ω is constant and vanishes when we differentiate.

Cartan's structure equation allows us to more easily prove basic properties of the curvature such as the following:

Lemma 4.3.6. *Let ω be a gauge field on a principal bundle (P, π, M, G) with curvature Ω. Then Ω is G-equivariant, i.e. $r_g^*\Omega = \mathrm{Ad}_{g^{-1}} \circ \Omega$ for all $g \in G$.*

Outline of proof. The idea is to use Cartan's structure equation, then show that the pull-back by r_g commutes in everywhere so that

$$r_g^*\Omega = \mathrm{d}(r_g^*\omega) + \frac{1}{2}\big[r_g^*\omega \wedge r_g^*\omega\big] = \mathrm{d}\big(\mathrm{Ad}_{g^{-1}} \circ \omega\big) + \frac{1}{2}\big[\mathrm{Ad}_{g^{-1}} \circ \omega \wedge \mathrm{Ad}_{g^{-1}} \circ \omega\big],$$

then finally show that the composition with $\mathrm{Ad}_{g^{-1}}$ commutes out everywhere. □

We provide and prove one more extremely important result here:

Theorem 4.3.7 (Second Bianchi identity)**.** *Let (P, π, M, G) be a principal bundle equipped with a connection ω, and let Ω be the curvature of ω. Then*

$$\mathrm{D}\Omega = 0.$$

Proof. We make use of Cartan's structure equation: for $p \in P$ and $u, v, w \in T_pP$ we have

$$\begin{aligned}
(\mathrm{D}\Omega)_p(u,v,w) &= \Big(\mathrm{D}\big(\mathrm{d}\omega + \frac{1}{2}[\omega\wedge\omega]\big)\Big)_p(u,v,w) \\
&= \big(\mathrm{d}^2\omega + \frac{1}{2}\,\mathrm{d}[\omega\wedge\omega]\,\big)_p(u^H,v^H,w^H) \\
&= \frac{1}{2}\big([\mathrm{d}\omega\wedge\omega] - [\omega\wedge\mathrm{d}\omega]\big)_p(u^H,v^H,w^H) \qquad \mathrm{d}^2 = 0,\ \textit{Leibniz} \\
&= [\mathrm{d}\omega\wedge\omega]_p(u^H,v^H,w^H) \\
&= 0, \qquad \omega(X^H) = 0 \textit{ for all } X \in TP
\end{aligned}$$

as needed. □

Remark. In the case of an abelian structure group G, Cartan's structure equation reduces to $\Omega = \mathrm{d}\omega$, and the second Bianchi identity is just a result of Ω being exact (and therefore closed). In this case we may just write down the stronger, easier to compute $\mathrm{d}\Omega = 0$; as we will see in the next subsection, this version second Bianchi identity provides us with the homogeneous Maxwell equations in the context of U(1) gauge theory.

4.3.1 Gauge field strength

Just like the gauge field itself, the curvature may be pulled back along a section of the principal bundle in order to obtain a physical object on the base space:

Definition 4.3.8. Let ω be a gauge field on a principal bundle (P, π, M, G) with curvature Ω, and let $\sigma : U \to P$ be a local section of the bundle. Then the pull-back $\mathcal{F}_\sigma = \sigma^*\Omega \in \Omega^2(M; \mathfrak{g})$ is called a **(local) field strength** on M.

We can commute the pull-back by a section σ through Cartan's structure equation (similar to the pull-back by r_g in Lemma 4.3.6) to obtain

$$\sigma^*\Omega = \mathrm{d}(\sigma^*\omega) + \frac{1}{2}[\sigma^*\omega \wedge \sigma^*\omega]$$

So that in terms of local gauge potentials $\mathcal{A}$ and field strength $\mathcal{F}$ we have

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \frac{1}{2}[\mathcal{A}\wedge\mathcal{A}]. \tag{4.1}$$

Similar to gauge potentials, we can write down a transformation law describing how the field strength changes under gauge transformations:

Proposition 4.3.9. *Let (P, π, M, G) be a principal bundle, let ω be a gauge field with curvature Ω and let $\sigma_1 : U_1 \to P$ and $\sigma_2 : U_2 \to P$ be local sections over non-disjoint subsets $U_1, U_2 \subset M$, related by the local gauge transformation g_{12} on $U_1 \cup U_2$. Then the local field strengths $\mathcal{F}_1 = \sigma_1^*\omega$ and $\mathcal{F}_2 = \sigma_2^*\omega$ are related by*

$$\mathcal{F}_2 = \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_1.$$

Remark. See [Nab11, Theorem 6.2.3] for a proof.

When G is a matrix Lie group, the transformation law reads $\mathcal{F}_2 = g_{12}^{-1}\mathcal{F}_1 g_{12}$. When G is any abelian group, the relation reduces to $\mathcal{F}_2 = \mathcal{F}_1$, so that the local field strength is invariant to gauge transformations; in this case, we often just denote $\mathcal{F}$ for any pull-back of Ω. Perhaps more notably, gauge theories with a non-abelian structure group generally do not have gauge invariant field strengths on the base space. In these cases, the Lagrangian density will turn out to be the main gauge invariant object of interest.

Example 4.3.10. Returning to $P = \mathbb{R}^4 \times \mathrm{U}(1)$, we are now equipped for a full geometric description of the EM field. We saw in Example 4.3.5 that the curvature on a U(1) bundle can be written $\Omega = i\,\partial_\nu A_\mu\,\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu \in \Omega^2(P;\mathfrak{g})$. We may pull this back along the global canonical section $x \mapsto (x,1)$ to obtain simply

$$\mathcal{F} = i\,\partial_\nu A_\mu\,\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu \in \Omega^2(\mathbb{R}^4;\mathfrak{g}),$$

and since U(1) is abelian, this field strength is in fact invariant to gauge transformations. Alternatively, we may first write down the gauge potential $\mathcal{A} = iA_\mu\,\mathrm{d}x^\mu$ and then use equation (4.1) to write down $\mathcal{F} = \mathrm{d}\mathcal{A}$, which leads to the same result.

Now, writing $\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu = \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu - \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$, we may consider $\mathcal{F}$ to be an anti-symmetric 2-tensor with components

$$\mathcal{F}_{\mu\nu} = i\,(\partial_\mu A_\nu - \partial_\nu A_\mu).$$

If A_μ is taken to be the EM four-potential, then by equation (1.7), $\mathcal{F}_{\mu\nu}$ corresponds to the EM field tensor $F_{\mu\nu}$ up to a factor i. To summarize, we have a gauge potential $\mathcal{A}$ with a corresponding field strength $\mathcal{F} = \mathrm{d}\mathcal{A}$, satisfying

$$\mathcal{A}_\mu = iA_\mu, \quad \mathcal{F}_{\mu\nu} = iF_{\mu\nu};$$

all arising naturally from the U(1) structure we've added to the base space.

With this background in mind, we are now ready to prove the following result:

Theorem 4.3.11. *Let $P = \mathbb{R}^4 \times \mathrm{U}(1)$ be the total space of the trivial* U(1)*-bundle over $\mathbb{R}^4$. Let ω be a gauge field on P with curvature Ω. Then the second Bianchi identity $\mathrm{D}\Omega = \mathrm{d}\Omega = 0$ is equivalent to the homogeneous Maxwell equation $\epsilon^{\mu\nu\sigma\rho}\partial_\nu F_{\sigma\rho}$ on the base space (equation (1.5)).*

Proof. First, $\mathrm{d}\Omega = 0$ directly implies that $\mathrm{d}\mathcal{F} = 0$; and indeed, this also follows naturally from $\mathcal{F} = \mathrm{d}\mathcal{A}$. Writing out

$$\mathcal{F} = i\,\partial_\nu A_\mu\,\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu,$$

we find

$$\mathrm{d}\mathcal{F} = i\partial_\rho\partial_\nu A_\mu\,\mathrm{d}x^\rho \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\mu\,.$$

Making the summation explicit and setting $\rho < \nu < \mu$ to consolidate like terms, we find

$$\begin{aligned}\mathrm{d}\mathcal{F} &= i \sum_{\rho<\nu<\mu} \big(\partial_\rho(\partial_\nu A_\mu - \partial_\mu A_\nu) + \partial_\nu(\partial_\mu A_\rho - \partial_\rho A_\mu) + \partial_\mu(\partial_\rho A_\nu - \partial_\nu A_\rho)\big)\,\mathrm{d}x^\rho \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\mu \\ &= i \sum_{\rho<\nu<\mu} \big(\partial_\rho F_{\nu\mu} + \partial_\nu F_{\mu\rho} + \partial_\mu F_{\rho\nu}\big)\,\mathrm{d}x^\rho \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\mu = 0.\end{aligned}$$

Since this must be identically zero, we find that for any $\rho < \nu < \mu$,

$$\partial_\rho F_{\nu\mu} + \partial_\nu F_{\mu\rho} + \partial_\mu F_{\rho\nu} = 0.$$

Multiplying on the left by $\epsilon^{\sigma\rho\nu\mu}$ gives

$$\epsilon^{\sigma\rho\nu\mu}(\partial_\rho F_{\nu\mu} + \partial_\nu F_{\mu\rho} + \partial_\mu F_{\rho\nu}) = \epsilon^{\sigma\rho\nu\mu}\partial_\rho F_{\nu\mu} + \epsilon^{\sigma\rho\nu\mu}\partial_\nu F_{\mu\rho} + \epsilon^{\sigma\rho\nu\mu}\partial_\mu F_{\rho\nu} = 0;$$

but then we are free to relabel the three contracted indices in each term to write

$$(\epsilon^{\sigma\rho\nu\mu} + \epsilon^{\sigma\mu\rho\nu} + \epsilon^{\sigma\nu\mu\rho})\partial_\rho F_{\nu\mu} = 0.$$

But all three of these terms differ by an even permutation of indices, meaning they are even: finally, then, we may write

$$\epsilon^{\sigma\rho\nu\mu}\partial_\rho F_{\nu\mu} = 0,$$

Which is just a relabelled form of the desired Maxwell equation. □

Chapter 5

The Lagrangian: Yang–Mills theory

Up to this point our discussion of gauge fields has been purely geometric in nature. All properties that we have derived for them thus far, such as the homogeneous Maxwell equation, are a direct result of the structure present on the principal bundle that they live on. If we want to describe the dynamics and interactions of those fields, we must impose physical laws on them; in practice we do this by defining a Lagrangian on the base space in terms of the gauge field. There are many ways to do this, but in this chapter we discuss just one physically sensible option: the gauge-invariant Yang–Mills Lagrangian, which generalises the Lagrangian for EM in free space ($J^\mu = 0$).

We might ask why we shouldn't formulate this Lagrangian on the total space: philosophically, since the Lagrangian is gauge invariant, the group structure on the bundle becomes redundant, and we can think of any resulting physics as taking place on the base space—the purpose of the principal bundle is to find gauge invariant objects in spacetime, and when we find them, the local symmetry is "spent". More practically, we want to be able to define multiple gauge fields with different structure groups over the same spacetime, and capture their interactions in one Lagrangian. This is only possible if we define the invariants for each field on the base space.

There are many different approaches to constructing the Lagrangian, ranging from the pragmatic to the highly abstract. On the pragmatic side are texts such as [Ton18] and (less so) [Nak17], which work in local coordinates whenever possible and only bother mathematically with whatever gets physical results. On the other end of the scale are highly mathematical texts such as [Ham17] and [Fra12], which resort to constructions such as the adjoint bundle in order to avoid ever being forced to go to a local description at all. This chapter tries to strike a happy medium between the two approaches: it leans more heavily on the latter two sources in attempting to achieve full generality, but our description of the Lagrangian itself is ultimately local.

An honourable mention goes out to [Ble81], who takes a uniquely different approach to formalising Lagrangian mechanics in general.

5.1 Basic definitions

We begin by making precise what we mean by a Lagrangian in this context:

Definition 5.1.1. Let (P, π, M, G) be a principal bundle. Then a **Lagrangian** (or technically, **Lagrangian density**) is a map

$$\mathscr{L} : \mathcal{A}(P) \to C^\infty(M),$$

where $\mathcal{A}(P)$ is the collection of gauge fields on P. When the gauge field ω is given we notate $\mathscr{L}[\omega] : M \to \mathbb{R}$.

We will want to integrate the Lagrangian for given gauge fields over M. This means we need M to be an oriented manifold, and since the Lagrangian for a gauge field is a real-valued function, we will also require a volume form on M. This volume form will be induced by a metric on M, which we formalise as follows:

Definition 5.1.2. Let M be a manifold. Then a **metric** g on M assigns to each point $x \in M$ a symmetric bilinear map $g_x : T_xM \times T_xM \to \mathbb{R}$ such that

(i) g_x is non-degenerate, i.e. for each non-zero $X_x \in T_xP$ the map $Y_x \mapsto g_x(X_x, Y_x)$ is not identically zero;

(ii) g varies smoothly with x, i.e. if $X, Y \in \mathfrak{X}(M)$ then $x \mapsto g_x(X_x, Y_x)$ is smooth.

The pair (M, g) is called a **pseudo-Riemannian manifold**.

Example 5.1.3. In classical field theory, the familiar metric tensor $g_{\mu\nu}$ corresponds to the Lorentzian metric on $\mathbb{R}^4$ given by $g(v, w) = v^\mathsf{T} \cdot \mathrm{diag}(-1, 1, 1, 1) \cdot w$.

Remark. This example also works the other way around: in any local coordinate description of a pseudo-Riemannian manifold (M, g), the metric may be used to raise and lower indices on tensors on M, exactly the same way as in classical field theory.

From now on, we will assume that the base space of our principal bundle is a pseudo-Riemannian manifold (M, g). As noted before, this allows a natural choice for the volume form:

Definition 5.1.4. If (M, g) is an oriented pseudo-Riemannian m-manifold, then the **canonical volume form** is the m-form ϵ on M such that for all $x \in M$, if $(e_1, \ldots, e_m)$ is an oriented basis of T_pM that is orthonormal with respect to g, then $\epsilon_x(e_1, \ldots, e_m) = 1$.

Example 5.1.5. On Minkowski spacetime $\mathbb{R}^4$, the canonical volume form is just the unit coordinate 4-form $\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$. In the tensor notation of classical field theory, this corresponds precisely to the fully antisymmetric Levi–Civita tensor $\epsilon_{\mu\nu\rho\sigma}$; this is what motivates the use of the notation ϵ.

For completeness, we can now define the action (in the physics sense):

Definition 5.1.6. Let (P, π, M, G) be a principal bundle with an oriented pseudo-Riemannian manifold (M, g) as a base space, and suppose we have a Lagrangian $\mathscr{L}$. Then the **action** corresponding to $\mathscr{L}$ is the functional

$$\mathcal{S} : \mathcal{A}(P) \to \mathbb{R}, \quad \mathcal{S}[\omega] := \int_M \mathscr{L}[\omega]\, \epsilon.$$

Remark. Note that, in order for the action to be well defined, we either require the base space M to be compact, or the Lagrangian $\mathscr{L}[\omega]$ to be compactly supported. In cases where M is not compact (such as $M = \mathbb{R}^4$ for electromagnetism) we will implicitly enforce this by restricting our attention to compactly supported gauge fields ω.

As usual, the equations of motion for the field arise as critical points of this action. We make this precise as follows:

Definition 5.1.7. We call a connection ω on P a **critical point** of the action $\mathcal{S}$ if

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathcal{S}[\omega + t\alpha] = 0,$$

where α is the difference $\eta - \theta$ between any two possible connections $\theta, \eta \in \mathcal{A}(P)$.

5.2 Scalar products

Given that the Lagrangian we are looking for is a real-valued functional of the Lie algebra-valued gauge field, we need some way to extract real numbers from elements of a Lie algebra. The approach we take in Yang–Mills theory is to recognise that the Lie algebra is a vector space, which means we can equip it with a scalar product $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$. In order to ensure gauge invariance of the resulting Lagrangian, it is important that this scalar product be Ad-invariant, i.e. $\langle \mathrm{Ad}_g(v), \mathrm{Ad}_g(w) \rangle_{\mathfrak{g}} = \langle v, w \rangle_{\mathfrak{g}}$. The following result is useful in this regard:[Ham17, Theorem 2.2.3]

Proposition 5.2.1. *Let G be a Lie group. If G is compact, then there exists a positive definite,* Ad*-invariant scalar product $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$ on $\mathfrak{g}$.*

Rather than prove this now, we will assume now that G is compact and demonstrate such a scalar product later. Before we do, recall from Theorem 2.1.3 that any compact Lie group is isomorphic to a matrix Lie group, and add to that the following result:[Fra12, p. 20.4c]

Proposition 5.2.2. *Any compact matrix Lie group $G \subset \mathrm{GL}_n(\mathbb{C})$ is isomorphic to a subgroup of $\mathrm{U}(n)$ up to a change of basis.*

Remark. We leave the proof for this proposition up to [Fra12].

It follows that we may restrict ourselves to any subgroup of $\mathrm{U}(n)$ without loss of generality, in the sense that the Yang–Mills action will not lead to any different dynamics for any other compact Lie group. We are now ready to define the scalar product on the Lie algebra of G:

Definition 5.2.3. Let $G \subset \mathrm{U}(n)$ be a Lie group with Lie algebra $\mathfrak{g} \subset \mathfrak{u}(n)$. Then we define the scalar product $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$ on $\mathfrak{g}$ as

$$\langle A, B \rangle_{\mathfrak{g}} := -\operatorname{tr} AB.$$

Proposition 5.2.4. *The scalar product $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$ is real-valued, symmetric, positive definite and* Ad*-invariant.*

Proof. From Example 2.2.9 we know that $\mathfrak{u}(n)$ consists of the skew-Hermitian matrices for which $A^\dagger = -A$. This means we have

$$\langle A, B \rangle_{\mathfrak{g}} = -\operatorname{tr} AB. = \operatorname{tr}\big((-A)B\big) = \operatorname{tr} A^\dagger B,$$

which is just the classic Frobenius inner product from linear algebra; we know this inner product to be real-valued, symmetric and positive definite.

We can check Ad-invariance directly:

$$\langle \mathrm{Ad}_g(A), \mathrm{Ad}_g(B) \rangle_{\mathfrak{g}} = \langle gAg^{-1}, gBg^{-1} \rangle_{\mathfrak{g}} = -\operatorname{tr} gAg^{-1}gBg^{-1} = -\operatorname{tr} gABg^{-1} = \langle A, B \rangle_{\mathfrak{g}},$$

where we use that the trace is invariant under basis changes. □

Now that we have a well-defined scalar product on the Lie algebra, we can adapt it to a scalar product of Lie algebra-valued forms on M in a way that respects the metric on M:

Definition 5.2.5. Let (M, g) be a pseudo-Riemannian manifold, and let $G \subset \mathrm{U}(n)$ be a Lie group with Lie algebra $\mathfrak{g}$. Then the scalar product $\langle \cdot, \cdot \rangle$ of $\mathfrak{g}$-valued k-forms $\alpha, \beta \in \Omega^k(M; \mathfrak{g})$ is defined pointwise as

$$\langle \alpha, \beta \rangle := \sum_{\mu_1 < \cdots < \mu_k} \langle \alpha_{\mu_1 \cdots \mu_k}, \beta^{\mu_1 \cdots \mu_k} \rangle_{\mathfrak{g}} = \frac{1}{k!} \langle \alpha_{\mu_1 \cdots \mu_k}, \beta^{\mu_1 \cdots \mu_k} \rangle_{\mathfrak{g}}.$$

Here the indices of β have been raised using the metric g on M.

Remark. Although this scalar product is defined in terms of local coordinates, it is well defined in the sense that it is independent of our choice of those coordinates.

If we establish a basis of generators $\{T_a\}$ for $\mathfrak{g}$, we can obtain an even more explicit formulation: we can break $\alpha \in \Omega^k(M; \mathfrak{g})$ down into components as

$$\alpha^a_{\mu_1 \cdots \mu_k} T_a := \alpha_{\mu_1 \cdots \mu_k},$$

where we sum over a as well as μ and ν; we then obtain

$$\langle \alpha, \beta \rangle = \frac{1}{k!} \langle \alpha^a_{\mu_1 \cdots \mu_k} T_a, \beta^{b \mu_1 \cdots \mu_k} T_b \rangle_{\mathfrak{g}} = \frac{1}{k!} \alpha^a_{\mu_1 \cdots \mu_k} \beta^{b \mu_1 \cdots \mu_k} \langle T_a, T_b \rangle_{\mathfrak{g}}.$$

This is independent of our choice of $\{T_a\}$, so we may as well fix a basis that is orthonormal with respect to $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$; the expression then reduces to

$$\langle \alpha, \beta \rangle = \sum_{a,b} \frac{1}{k!} \alpha^a_{\mu_1 \cdots \mu_k} \beta^{b \mu_1 \cdots \mu_k}. \tag{5.1}$$

Note that the right hand side of this equation need not be positive; it follows that this scalar product is not positive definite.

5.3 General Yang–Mills theory

We are now fully equipped to formulate the general form of the Yang–Mills Lagrangian and its action:

Definition 5.3.1. Let (P, π, M, G) be a principal bundle with $G \subset \mathrm{U}(n)$ and (M, g) an oriented pseudo-Riemannian manifold. Then the **Yang–Mills Lagrangian** is given by

$$\mathscr{L}_{YM}[\omega] := \frac{1}{2}\|\mathcal{F}\|^2$$

where $\mathcal{F}$ is the local field strength of ω in any gauge and $\|\mathcal{F}\|^2$ is a shorthand for $\langle \mathcal{F}, \mathcal{F} \rangle$. Integrating over M gives rise to an action named the **Yang–Mills Functional**:

$$\mathcal{S}_{YM}[\omega] := \frac{1}{2}\int_M \|\mathcal{F}\|^2 \, \epsilon$$

where ϵ is the canonical volume form on M.

Remark. While the notation $\|\mathcal{F}\|^2 = \langle \mathcal{F}, \mathcal{F} \rangle$ is very suggestive of a squared norm, we must keep in mind that the scalar product $\langle \cdot, \cdot \rangle$ may allow negative values depending on the metric. This means we can only properly define a norm $\|\cdot\|$ for metrics with a positive *signature*; the use is otherwise purely notational.

Theorem 5.3.2. *The Yang–Mills Lagrangian is well-defined, i.e. gauge invariant. More precisely, $\mathscr{L}_{YM}[\omega]$ is independent of our choice of section σ to obtain $\mathcal{F} = \sigma^*\Omega$.*

Proof. Proposition 4.3.9 tells us that the local field strength transforms as

$$\mathcal{F}_2 = \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_1.$$

Plugging both sides into the Yang–Mills Lagrangian gives

$$\begin{aligned}
\frac{1}{2}\langle \mathcal{F}_2, \mathcal{F}_2 \rangle &= \frac{1}{2}\left\langle \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_1, \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_1 \right\rangle \\
&= \frac{1}{4}\left\langle \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_{1\mu\nu}, \mathrm{Ad}_{g_{12}^{-1}} \circ \mathcal{F}_1^{\mu\nu} \right\rangle_{\mathfrak{g}} \quad \Big\rangle\, \langle \cdot, \cdot \rangle_{\mathfrak{g}} \text{ is } \mathrm{Ad}\text{-invariant} \\
&= \frac{1}{4}\langle \mathcal{F}_{1\mu\nu}, \mathcal{F}_1^{\mu\nu} \rangle_{\mathfrak{g}} \\
&= \frac{1}{2}\langle \mathcal{F}_1, \mathcal{F}_1 \rangle,
\end{aligned}$$

which is exactly what we need. □

As stated before, the critical points of the Yang–Mills functional give rise to equations of motion on the base space. Before we can formulate these in full generality, we need to introduce some final notation.

Definition 5.3.3. Let (M, g) be a pseudo-Riemannian m-manifold, and let $G \subset \mathrm{U}(n)$ be a Lie group with Lie algebra $\mathfrak{g}$. Then the **Hodge star operator**

$$\star : \Omega^k(M; \mathfrak{g}) \to \Omega^{m-k}(M; \mathfrak{g})$$

is the linear map uniquely determined by

$$\langle \theta, \eta \rangle \, \epsilon = \theta \wedge \star \eta$$

for all $\theta, \eta \in \Omega^k(M, \mathfrak{g})$.

Example 5.3.4. Perhaps the most striking example is provided by the canonical volume form on M: we have

$$\epsilon = \star(1).$$

The equations of motion are now defined in terms of this Hodge star:

Theorem 5.3.5. *Let (P, π, M, G) be a principal bundle with $G \subset \mathrm{U}(n)$ and (M, g) an oriented pseudo-Riemannian manifold. Then a connection ω on P is a critical point of the Yang–Mills functional $\mathcal{S}_Y M$ if and only if*

$$\mathrm{D} \star \mathcal{F}_\sigma = \mathrm{d} \star \mathcal{F}_\sigma + [\mathcal{A}_\sigma, \star \mathcal{F}_\sigma] = 0$$

for any local field strength $\mathcal{F}$ corresponding to ω. We call this requirement the ***Yang–Mills equation****.*

Outline of proof. We adapt the terminology from Definition 5.1.7 here.

We want to calculate the change in $\mathcal{F}$ when we add $t\alpha$ to the connection ω. We work locally and write $\mathcal{A} + t\alpha' = \sigma^*(\omega + t\alpha)$. We also drop the lower index for σ on $\mathcal{F}$ and use an upper index to indicate which connection we pull back from instead. Using Cartan's structure equation, we find

$$\begin{aligned}
\mathcal{F}^{\omega+t\alpha} &= \mathrm{d}\big(\mathcal{A} + t\alpha'\big) + \frac{1}{2}[\mathcal{A} + t\alpha', \mathcal{A} + t\alpha'] \\
&= \mathcal{F}^\omega + t\big(\mathrm{d}\alpha' + [\mathcal{A}, \alpha']\big) + \frac{1}{2}t^2[\alpha', \alpha'] \\
&= \mathcal{F}^\omega + t\mathrm{D}\alpha' + \frac{1}{2}t^2[\alpha', \alpha']
\end{aligned}$$

We can calculate $||\mathcal{F}^{\omega+t\alpha}||^2$ to first order:

$$\begin{aligned}
\left\|\mathcal{F}^{\omega+t\alpha}\right\|^2 &= \langle \mathcal{F}^\omega + t\mathrm{D}\alpha', \mathcal{F}^\omega + t\mathrm{D}\alpha' \rangle + \mathcal{O}(t^2) \\
&= \|\mathcal{F}^\omega\|^2 + 2t\langle \mathrm{D}\alpha', \mathcal{F}^\omega \rangle + \mathcal{O}(t^2)
\end{aligned}$$

It follows that the critical points are solutions to

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \int_M \left\|\mathcal{F}^{\omega+t\alpha}\right\|^2 \epsilon = 2 \int_M \langle \mathrm{D}\alpha', \mathcal{F}^\omega \rangle \, \epsilon = 0.$$

The idea for the last step is now to write

$$\int_M \langle \mathrm{D}\alpha', \mathcal{F}^\omega \rangle \, \epsilon := \int_M \langle \alpha', \mathrm{D}^* \mathcal{F}^\omega \rangle \, \epsilon$$

and prove that $\star \mathrm{D}^* \mathcal{F}^\omega = \mathrm{d} \star \mathcal{F}^\omega$; for this we defer to [Ham17, Theorem 7.2.12]. □

5.4 Electromagnetism

Between the Yang–Mills equation and the second Bianchi identity $\mathrm{D}\Omega = 0$, we now have a full description for the free space dynamics of a single gauge field with a compact structure group. In this final section, we will return to our familiar case of a local U(1) symmetry over a Minkowski spacetime, and demonstrate that the formalism reduces exactly to free space EM.

We have already shown that the second Bianchi identity gives rise to the homogeneous Maxwell equation 1.5. We will now derive the free space inhomogeneous equation from the U(1) Yang–Mills theory. The simplest way to do this is to show that the Yang–Mills action is equivalent to the free space action from classical field theory; since their critical points correspond to the Yang–Mills equation and the inhomogeneous Maxwell equation respectively, this will immediately imply their equivalence as well.

Theorem 5.4.1. *Let $P = \mathbb{R}^4 \times \mathrm{U}(1)$ be the trivial principal $\mathrm{U}(1)$-bundle over spacetime $\mathbb{R}^4$ equipped with the Lorentzian metric. Then the Yang–Mills Lagrangian*

$$\mathcal{S}_{YM}[\omega] := \frac{1}{2} \int_{\mathbb{R}^4} \|\mathcal{F}\|^2 \, \epsilon$$

is exactly equivalent to the free space Maxwell action

$$\mathcal{S}_M = -\frac{1}{4} \int F_{\mu\nu} F^{\mu\nu} \, \mathrm{d}^4\mathbf{x} \, .$$

Proof. We begin by fixing a potential ω; by Theorem 5.3.2, we may pick any random gauge σ to identify $\mathcal{F}$ with $\mathcal{F}_\sigma = \sigma^*\omega$. We can then freely write

$$\mathcal{S}_{YM}[\omega] = \frac{1}{2} \int_{\mathbb{R}^4} \|\mathcal{F}\|^2 \, \epsilon = \frac{1}{2} \int_{\mathbb{R}^4} \langle \mathcal{F}, \mathcal{F} \rangle \, \epsilon = -\frac{1}{4} \int_{\mathbb{R}^4} \mathrm{tr}(\mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu}) \, \epsilon.$$

On $\mathfrak{u}(1) = i\mathbb{R} \subset \mathbb{C} \cong \mathcal{M}_1(\mathbb{C})$ we have $\mathrm{tr} = \mathrm{id}_{\mathbb{C}}$, so that

$$\mathcal{S}_{YM}[\omega] = -\frac{1}{4} \int_{\mathbb{R}^4} \mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu} \, \epsilon;$$

the matrix multiplication on $\mathfrak{u}(1)$ is just regular multiplication of complex numbers, and we know from Example 4.3.10 that $\mathcal{F}_{\mu\nu} = iF_{\mu\nu}$, meaning

$$\mathcal{S}_{YM}[\omega] = \frac{1}{4} \int_{\mathbb{R}^4} F_{\mu\nu} F^{\mu\nu} \, \epsilon.$$

From Example 5.1.5, we know that $\epsilon = \mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$ in this case, so that $\int_{\mathbb{R}^4} \cdot \, \epsilon$ is just regular integration $\int \cdot \, \mathrm{d}^4\mathbf{x}$. Finally, then, we're left with

$$\mathcal{S}_{YM}[\omega] = \frac{1}{4} \int F_{\mu\nu} F^{\mu\nu} \, \mathrm{d}^4\mathbf{x} = -\mathcal{S}_M;$$

the two actions wind up being identical up to a sign change, but since sign changes do not affect critical points, both actions produce the same equations of motion. □

As stated before, the following now follows immediately from Theorem 5.3.5 and the standard result using the Euler–Lagrange equations from classical field theory:

Corollary 5.4.2. *The Yang–Mills equation for a principal* $\mathrm{U}(1)$*-bundle over Minkowski spacetime*

$$\mathrm{d} \star \mathcal{F} = 0$$

is precisely equivalent to the free space inhomogeneous Maxwell equation

$$\partial_\nu F^{\mu\nu} = 0.$$

Remark. For the Yang–Mills equation we have made use of the fact that the Lie bracket on $\mathfrak{u}(1)$ is trivial.

5.4.1 Closing remarks

All in all, we have seen that the crudest form of electromagnetism can be captured perfectly in terms of a highly geometric formalism with one very natural restriction in the form of a Lagrangian. The natural question to ask is whether we can account for non-zero charge density within this formalism; the exciting answer is that we can, and the underlying idea is the rich theory of interactions between different gauge fields. The result ends up looking more quantum mechanical than classical, and it is a great mathematical motivation behind the physical ideas of quantum mechanics: properties like quantisation may arise naturally from the topology of and interactions on the fields. This in turn provides great motivation for the formalism itself: many aspects of quantum field theory, for example, would be difficult to conceptualise without a rigorous way to encode local symmetries geometrically. It should come as no surprise, then, that gauge theory continues to be an active area of research in physics and mathematics alike.

For those interested in further reading, the author heartily recommends [Ham17]'s treatment of the subject: Hamilton goes on to describe the entirety of the Standard Model Lagrangian in great mathematical detail. For those interested in the topological side of the theory, the latter chapters in [Nak17] offer a fairly in-depth look.

Bibliography

[Ble81] David Bleecker. *Gauge theory and variational principles.* Vol. 1. Global analysis, pure and applied. Reading, MA: Addison–Wesley Publishing Company, Inc., 1981. ISBN: 0-201-10096-7.

[Ein05] Albert Einstein. “Zur Elektrodynamik bewegter Körper”. In: *Annalen der Physik* 17 (1905), p. 891. DOI: 10.1002/andp.19053221004.

[Fra12] Theodore Frankel. *The geometry of physics. An introduction.* 3rd ed. Cambridge: Cambridge University Press, 2012. ISBN: 978-1-107-60260-1.

[Gri17] David J. Griffiths. *Introduction to electrodynamics.* 4th ed. Cambridge: Cambridge University Press, 2017. ISBN: 978-1-108-42041-9.

[Ham17] Mark J. D. Hamilton. *Mathematical gauge theory. With applications to the Standard Model of particle physics.* Electronic. Universitext. Cham: Springer, 2017. DOI: 10.1007/978-3-319-68439-0.

[KMS93] Ivan Kolář, Peter W. Michor, and Jan Slovák. *Natural operations in differential geometry.* Electronic. Berlin etc.: Springer–Verlag, 1993. ISBN: 3-540-56235-4. URL: https://www.emis.de/monographs/KSM (visited on 10/31/2022).

[Kna02] Anthony W. Knapp. *Lie groups beyond an introduction.* 2nd ed. Vol. 140. Progress in mathematics. Boston, MA: Birkhäuser, 2002. ISBN: 0-8176-4259-5.

[Lee09] Jeffrey M. Lee. *Manifolds and differential geometry.* Vol. 107. Graduate studies in mathematics. Providence, RI: American Mathematical Society, 2009. ISBN: 978-1-4704-6982-5. URL: https://bookstore.ams.org/gsm-107 (visited on 10/28/2022).

[Lee12] Jeffrey M. Lee. *Online supplement for “Manifolds and differential geometry”.* 2012. URL: https://www.ams.org/bookpages/gsm-107/Supp.pdf (visited on 01/23/2023).

[LM87] Paulette Libermann and Charles Marle. *Symplectic geometry and analytical mechanics.* Mathematics and its applications. Dordrecht: D. Reidel Publishing Company, 1987. ISBN: 90-277-2438-5. URL: https://archive.org/details/symplecticgeomet0000libe.

[Max65] James C. Maxwell. “VIII. A dynamical theory of the electromagnetic field”. In: *Phil. Trans. R. Soc.* 155 (1865), pp. 459–512. DOI: 10.1098/rstl.1865.0008.

[Nab11] Gregory L. Naber. *Topology, geometry and gauge fields. Foundations.* 2nd ed. Vol. 25. Texts in Applied Mathematics. New York, NY: Springer, 2011. DOI: 10.1007/978-1-4419-7254-5. URL: http://link.springer.com/10.1007/978-1-4419-7254-5 (visited on 10/28/2022).

[Nak17] Mikio Nakahara. *Geometry, topology and physics.* 2nd ed. Boca Raton: CRC Press, 2017. DOI: 10.1201/9781315275826. URL: https://www.taylorfrancis.com/books/9781420056945 (visited on 10/28/2022).

[Sop08] Davison E. Soper. *Classical field theory.* Dover. Dover, 2008. ISBN: 978-0-486-46260-8.

[Ton18] David Tong. *Gauge theory.* 2018. URL: http://www.damtp.cam.ac.uk/user/tong/gaugetheory.html.