Instantons and the ADHM construction

AJ Lindenhovius

Advisor: Prof. Dr. R.H. Dijkgraaf

Final Version

Contents

Introduction

Overview of gauge theories and instantons

For centuries the marriage between physics and mathematics has been extremely fruitful. Newton, for instance, had to develop the calculus in order to be able to formulate his mechanics. In the beginning of previous century Einstein was able to formulate his theory of general relativity since its mathematical framework of differential geometry had already been developed by Riemann. Another great success of the interplay between physics and mathematics is the functional analysis, whose roots lie in Von Neumann's successful attempts of uniting Schrödinger's wave mechanics formulation of quantum mechanics with Heisenberg's matrix mechanics formulation.

In the second half of previous century another love child, *gauge theories*, was born. The importance for physics is immediately clear since three of nature's four fundamental forces, are described by a gauge theory. What a gauge theory is can be best explained by the example of electromagnetism, the most simple gauge theory. In electromagnetism the fundamental physical quantities are the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, since the electromagnetic force can be expressed only in terms of these two fields. The dynamics of the electric and magnetic fields in vacuum are described by the Maxwell equations.

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 0, \qquad & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \qquad & \nabla \times \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t}. \end{aligned}$$

Since the components of $\mathbf{E}$ are not completely independent, we can introduce a potential V such that under certain circumstances we have $\mathbf{E} = \nabla V$. A more general formula exists, but for the sake of argument we assume this one. Conversely, we can express V in terms of $\mathbf{E}$ by the following line integral

$$V(\mathbf{x}) = -\int_{\mathbf{x}_0}^{\mathbf{x}} \mathbf{E}(\mathbf{x}) \cdot \mathrm{d}\mathbf{l},$$

where the point $\mathbf{x}_0$ is a reference point. One can check that V given by this integral satisfies $\nabla V = \mathbf{E}$. However, we have a freedom in the choice of the reference point $\mathbf{x}_0$. If we denote by V' the potential obtained by the same integral formula but with $\mathbf{x}_0$ replace by $\mathbf{x}_1$, it turns out that $\nabla V' = \nabla V$. More generally, the electric field $\mathbf{E}$ does not change if we add a function f to V as long as $\nabla f = 0$. So we have a certain freedom in the choice of the potential, which is called *gauge freedom*. A more direct illustration of this gauge freedom is given by the convention that the voltage of the earth is zero. Given this convention, the maximal output of an electrical outlet is 220 V. However, we have the freedom to say that the voltage of the earth is 80.000 V and the maximum output of the outlet 80.220 V, which does not alter the physics, since only the differences in voltage are relevant. Similar to the electric potential, we could introduce a magnetic potential $\mathbf{A}$ such that $\mathbf{B} = \nabla \times \mathbf{A}$. Also in $\mathbf{A}$ we have a gauge freedom, we could always add a function $\mathbf{f}$ to $\mathbf{A}$ as long as $\nabla \times \mathbf{f} = 0$.

Our formulas become more compact if we use covariant notation. We collect the potentials $\mathbf{A}$ and V in one 4-vector $\mathcal{A}_\mu$, the *gauge potential* with entries $(\mathbf{A}, V)$. Furthermore, we introduce the field tensor $\mathcal{F}_{\mu\nu}$ given by

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

Then $\mathcal{F}_{\mu\nu}$ can be expressed in terms of $\mathcal{A}_\mu$ by

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu,$$

and the Maxwell equations are given by a single equation

$$\partial_\mu (*\mathcal{F})_{\mu\nu} = 0,$$

where $(*\mathcal{F})_{\mu\nu}$ is the dual tensor, which can be obtained from $\mathcal{F}_{\mu\nu}$ by the substitutions $\mathbf{E} \to \mathbf{B}$ and $\mathbf{B} \to -\mathbf{E}$. For mathematicians it is convenient to redefine $\mathcal{F}$ and $\mathcal{A}$ as $\mathcal{F} \to i\mathcal{F}$ and $\mathcal{A} \to i\mathcal{A}$. Then $\mathcal{F}_{\mu\nu}$ is clearly invariant if we add $i\partial_\mu \Lambda(x)$ to $\mathcal{A}_\mu$, where Λ is a function with values in $\mathbb{R}$. This transformation $\mathcal{A}_\mu \mapsto \mathcal{A}_\mu + i\partial_\mu \Lambda(x)$ is called a *gauge transformation* and all gauge freedom lies in the fact that $\mathcal{F}_{\mu\nu}$ is invariant under this transformation. In order to generalize to gauge theories in general, it is convenient to introduce the action

$$S = \int_{\mathbb{R}^4} \mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu} \mathrm{d}^4 x,$$

which is clearly invariant under gauge transformations. Then $\mathcal{F}_{\mu\nu}$ satisfies the Maxwell equations if and only if $\mathcal{F}_{\mu\nu}$ is an minimum of the action. If we examine the gauge transformation, we see that we can rewrite it in the following form: $\mathcal{A}_\mu \mapsto \mathcal{A}_\mu + e^{-i\Lambda(x)} \partial_\mu e^{i\Lambda(x)}$. Since $e^{i\Lambda(x)}$ is an element of the (abelian) group U(1), and conversely all elements $g(x) \in \mathrm{U}(1)$ are of the form $e^{i\Lambda(x)}$, we can rewrite the gauge transformation as follows.

$$\mathcal{A}_\mu \mapsto g(x) \mathcal{A}_\mu g(x)^{-1} + g(x)^{-1} \partial_\mu g(x)$$

with $g(x) \in \mathrm{U}(1)$. So we see that the gauge transformations of electromagnetism are in deep relation with the group U(1), which is is called the *structure group* of the gauge theory. Now we have arrived at the point that we can generalize to other gauge theories, whose structure is almost the same except for that they have different structure groups, which do not have to be abelian like U(1). On contrary, almost all gauge theories, except electromagnetism, have non-abelian structure groups. For all gauge theories the gauge transformations have the same form as our last version, but this has its consequences for $\mathcal{A}$ and $\mathcal{F}$. Firstly, we see that the term $g(x)^{-1}\partial_\mu g(x)$ is an element of the Lie algebra $\mathfrak{g}$ of the structure group G. We see if we want to be consequent that $\mathcal{A}_\mu$ must also be $\mathfrak{g}$-valued. Since $\mathcal{F}_{\mu\nu}$ is related to $\mathcal{A}_\mu$, we find that also $\mathcal{F}_{\mu\nu}$ must be $\mathfrak{g}$-valued. The relation between $\mathcal{F}_{\mu\nu}$ and $\mathcal{A}_\mu$ also alters slightly and becomes

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

The action does not change, except for that we implicitly assume that an inner product on $\mathfrak{g}$. The equations of motion for general gauge theories can be derived from the action by finding its minima. We obtain then the *Yang-Mills* equation

$$\mathcal{D}_\mu (*\mathcal{F})_{\mu\nu} = 0,$$

where $\mathcal{D}_\mu$ is the covariant derivative given by $\mathcal{D}_\mu = \partial_\mu + \mathcal{A}_\mu$. For electromagnetism the Yang-Mills equation turns out to be equivalent with the Maxwell equations.

The Maxwell equations are relatively easy to solve, but this is definitely not the case Yang-Mills equation for more complicated gauge theories, since in these cases the Yang-Mills equations are second-order differential equations which are non-linear due to the commutator of $[\mathcal{A}_\mu, \mathcal{A}_\nu]$ in $\mathcal{F}_{\mu\nu}$. However, in Euclidean space some solutions are easier to find. These are the absolute minima of the action, which satisfy a much simpler first-order partial differential equation

$$\mathcal{F}_{\mu\nu} = \pm (*\mathcal{F})_{\mu\nu},$$

called the *anti-self-duality equation* if we have a minus sign and the *self-duality* equation otherwise. Since solutions $\mathcal{A}_\mu$ of these equation, which we call *instantons*, are absolute minima of the action, they are automatically solutions of the Yang-Mills equation. Instantons are interesting for physicists, since they provide information about all solutions of the Yang-Mills equation, but also describe the phenomena of tunneling between different vacuum states. For mathematicians instantons are interesting since they provide information about the topology of the space in which they exist. For example, if we denote the group of all gauge transformation by $\mathcal{G}$ and the space of all self-dual instantons, that is solutions of the self-duality equation, by $\mathscr{A}^+$, then we can introduce the *moduli space* $\mathscr{M} = \mathscr{A}^+/\mathcal{G}$ of all self-dual instantons modulo gauge transformations. This space behaves neatly under certain conditions, and turns out to be a very important object for mathematicians, since Donaldson used it in order to prove his famous result of the existence of exotic differential structures on $\mathbb{R}^4$.

Since the (anti-)self-duality equation is also non-linear in $\mathcal{A}_\mu$, it is still hard to solve. It were Atiyah, Drinfeld, Hitchin and Manin, however, who found a trick for gauge theories with structure group $\mathrm{SU}(n)$. They discovered a set of quadratic algebraic equations, called the *ADHM equations*, whose solutions are in 1-1 correspondence with the instanton solutions. The ADHM construction is not the endpoint in the research of instantons. For physicists instantons are still important, since they also occur in other theories like general relativity and string theory. Moreover, instantons play an important role in the path integral formalism of quantum field theories, a formalism which is still not perfectly understood. Mathematicians have found some generalizations of the ADHM construction for other spaces than $\mathbb{R}^4$, for instance for ALE spaces. So we can conclude that gauge theories and instantons are still important objects of research for both physicists and mathematicians.

Outline of thesis

In this thesis we shall concentrate us on the mathematical aspects of gauge theories and instantons. We want to define gauge theories in curved spacetime, for which we have to introduce principal bundles in the first section, since they offer the mathematical framework for gauge theories on curved spaces. In the second section we shall introduce connections and show that these are equivalent with gauge potentials. Furthermore, we introduce the gauge field-strength, the object describing the physical configuration within gauge theories, and the gauge group. In the third section we shall discuss the topological properties of gauge theories using Chern classes, continued by introducing instantons and showing that there is a connection between former and latter. Finally, we calculate some instanton solutions with instanton number -1, introduce the moduli space, and derive from its dimension that we found all solutions with instanton number -1. In the fourth section we discuss the ADHM construction, which is a linear algebraic construction providing all possible instanton solutions for gauge theories with structure group $\mathrm{SU}(n)$. We shall show that the ADHM construction indeed yields $\mathrm{SU}(n)$-instantons by the use of holomorphic vector bundles. Finally, we introduce quivers, which are directed graphs and which can be used to give an alternative description of the ADHM construction.

Acknowledgements

First of all I would like to thank my advisor, R.H. Dijkgraaf, for his guidance, his willingness to spend time on advising me despite all his other activities, his enthusiasm, which was very inspiring, and for offering the opportunity to do my master research on this very interesting subject of instantons.

Secondly, I would like to express my gratitude to my friend and fellow student S. Ringers. You were of great support for me during this project, I appreciate all your help, insights and corrections and I sincerely hope that we will be able to collaborate together once again.

1 Principal bundles

1.1 The definitions of fibre and principal bundles

Definition 1.1.1. A differentiable *coordinate bundle* (E, π, M, F, G) consists of the following elements

1. A differentiable manifold E called the total space.
2. A differentiable manifold M called the base space.
3. A differentiable manifold F called the fibre.
4. A surjection $\pi : E \to M$ called the projection, such that $\pi^{-1}(x) \equiv F_x \cong F$. Sometimes we rather use the notation E_x instead of F_x.
5. A Lie group G called the structure group, which acts *freely* on the fibre on the left; if $gf = f$ for all $f \in F$, then $g = e$. In this thesis we shall only consider Lie subgroups of $GL(\mathbb{C}, n)$.
6. An open covering $\{U_i\}_{i\in I}$ of M, where I is an index set, and a set of diffeomorphisms $\psi_i : U_i \times F \to \pi^{-1}(U_i)$ such that $\pi \circ \psi_i(x, f) = x$. This means that ψ_i is the diffeomorphism such that

$$\begin{tikzcd} U_i \times F \arrow[rr, "\psi_i"] \arrow[dr, "p_1"'] & & \pi^{-1}(U_i) \arrow[dl, "\pi"] \\ & U_i & \end{tikzcd}$$

   commutes, where p_1 is the projection on the first coordinate. The map ψ_i is called a *local trivialization* and the set U_i a *coordinate neighborhood*. In case we can find a diffeomorphism $\psi : M \times F \to E$, we say ψ is a *global* trivialization and the bundle is called *trivial*.
7. The maps $\psi_{i,x} : F \to F_x$ defined by $\psi_{i,x}(f) = \psi_i(x, f)$ satisfy for all $i, j \in I$, and $x \in U_i \cap U_j$, the condition that $\psi_{j,x}^{-1} \circ \psi_{i,x} : F \to F$ is a diffeomorphism which coincides with the operation of an element of G (which is unique, since G acts freely on F). We shall abbreviate $U_i \cap U_j$ by U_{ij}.

Remarks.

1. Since $\psi_{i,x}$ provides the diffeomorphism between F and F_x, we shall from now on identify the latter with the first.
2. For simplicity we shall sometimes use the notation $E \xrightarrow{\pi} M$ or just E instead of (E, π, M, F, G) and refer to it as 'E is a bundle over M.'

Fibre bundles were introduced as generalization of the product of two spaces. We continue by defining two functions, which will be very helpful in the future.

Definition 1.1.2. We define the functions $f_i : E \to F$ and $g_{ij} : U_{ij} \to G$ by

$$f_i(p) = \psi_{i,x}^{-1}(p);$$
$$g_{ij}(x) = \psi_{i,x}^{-1} \circ \psi_{j,x}.$$

Clearly f_i and g_{ij} are smooth. In some occasions we will drop the argument of $f_i(p)$ and write just f_i. If $F = G$, we shall write g_i instead of f_i.

Proposition 1.1.3.

1. For all $i, j \in I$, $x \in U_{ij}$ and $p \in \pi^{-1}(U_{ij})$, the maps g_{ij} and f_i satisfy
$$g_{ij}(x)f_j(p) = f_i(p). \tag{1.1.1}$$
2. If the overlap U_{ij} is non-empty, then ψ_i and ψ_j are related to each other on U_{ij} by
$$\psi_j(x, f) = \psi_i(x, g_{ij}(x)f). \tag{1.1.2}$$

Proof.

1. By definition of both functions, we find
$$\begin{aligned} g_{ij}(x)f_j(p) &= \psi_{i,x}^{-1} \circ \psi_{j,x}\big(f_j(p)\big) \\ &= \psi_{i,x}^{-1}(p) \\ &= f_i(p). \end{aligned}$$
2. Let $p = \psi_j(x, f)$. Then $f_j(p) = f$, so we find
$$\begin{aligned} \psi_j(x, f) &= p \\ &= \psi_i\big(x, f_i(p)\big) \\ &= \psi_i\big(x, g_{ij}(x)f_j(p)\big) \\ &= \psi_i(x, g_{ij}(x)f), \end{aligned}$$
which is exactly the statement.

□

The following proposition is easily proved:

Proposition 1.1.4. We have the following identities:
$$\begin{aligned} g_{ii}(x) &= \mathrm{Id}(x) & [x \in U_i] \quad (1.1.3) \\ g_{ij}(x) &= g_{ji}^{-1}(x) & [x \in U_i \cap U_j] \quad (1.1.4) \\ g_{ij}(x)g_{jk}(x) &= g_{ik}(x) & [x \in U_i \cap U_j \cap U_k]. \quad (1.1.5) \end{aligned}$$

Definition 1.1.5. Two coordinate bundles $\mathcal{B}$ and $\mathcal{B}'$ are equivalent if they have the same total space, base space, projection, fibre and group, and their coordinate functions $\{\psi_i\}$, $\{\psi'_j\}$ satisfy the conditions that $\bar{g}_{ij}(x) = \psi'^{-1}_{i,x} \circ \psi_{x,j}$ coincides with the operation of an element of G, and the map $\bar{g}_{ij} : U'_{ij} \to G$ is smooth. The equivalence class of a coordinate bundle is called a *fibre bundle*. A fibre bundle $\pi : P \to M$ with fibre identical to a vector space is called a *vector bundle*. If the fibre F is identical to the structure group G, and the left action on the fibre F corresponds with left multiplication on G, the bundle is called a *principal bundle*, denoted with $P(M, G)$. Note that in this case the action of G on the fibre is automatically freely.

Definition 1.1.6. Let $R_a : G \to G$ be the right multiplication with an element $a \in G$. Let $P(M, G)$ be a principal bundle, and let $p = \psi_i\big(x, g_i(p)\big)$ for some trivialization $\psi_i : U_i \times G \to \pi^{-1}(U_i)$. Then the right multiplication $R_a : G \to G$ with an element $a \in G$ induces a right action $R_a : \pi^{-1}(U_i) \to \pi^{-1}(U_i)$ by $R_a(p) = \psi\big(x, R_a g_i(p)\big)$, or equivalently
$$pa = \psi_i(x, g_i(p)a). \tag{1.1.6}$$

Proposition 1.1.7. The right action defined above satisfies the following properties:

1. The action is globally defined;
2. The action is is free: We have $pa = p \implies a = e$ for all $p \in P$;
3. The action is transitive on the fibres: for all $p_1, p_2 \in \pi^{-1}(x)$ there is an $a \in G$ such that $p_1 = p_2 a$;
4. π is G-invariant: we have $\pi \circ R_a = \pi$, or equivalently, $\pi(pa) = \pi(p)$ for all $a \in G$;
5. g_i is G-equivariant: we have $g_i \circ R_a = R_a \circ g_i$, or equivalently, $g_i(pa) = g_i(p)a$ for all $a \in G$;
6. Every fibre $\pi^{-1}(x)$ coincides with $pG = \{pg : g \in G\}$ for any $p \in \pi^{-1}(x)$.

Proof.

1. Let $x \in U_{ij}$, then by (2) of proposition 1.1.3 we find
$$\begin{aligned} pa &= \psi_i(x, g_i a) \\ &= \psi_j(x, g_{ji}(x) g_i a) \\ &= \psi_j(x, g_j a), \end{aligned}$$
hence pa is globally defined.
2. For $p \in P$ let ψ_i a trivialization with $p \in \pi^{-1}(U_i)$. Then $p = \psi_{i,x}\Big(g_i(p)\Big)$ for some $x \in U_i$. Hence $pa = p$ is equivalent with $\psi_{i,x}(g_i(p)a) = \psi_{i,x}\Big(g_i(p)\Big)$. Now since $\psi_{i,x}$ is a diffeomorphism, we find $g_i(p)a = g_i(p)$, implying $a = e$.
3. Let $x \in M$ and $p_1, p_2 \in \pi^{-1}(x)$. Choose a trivialization ψ_i, such that $x \in U_i$. Now, if we choose $a = g_i(p_1)^{-1} g_i(p_2)$, we find
$$\begin{aligned} p_1 a &= \psi_i(x, g_i(p_1)a) \\ &= \psi_i\Big(x, g_i(p_2)\Big) \\ &= p_2, \end{aligned}$$
hence the actions is transitive on the fibres.
4. For $p \in P$ let $x \in M$ and $g \in G$ such that $p = \psi(x, g)$ for a trivialization ψ. Then we have
$$\begin{aligned} \pi(pa) &= \pi(\psi(x, g)a) \\ &= \pi\Big(\psi(x, ga)\Big) \\ &= x \\ &= \pi(p), \end{aligned}$$
where the third equality follows from the definition of a trivialization.
5. By previous statement in this proposition, we have
$$\begin{aligned} \psi_i\Big(\pi(p), g_i(pg)\Big) &= \psi_i\Big(\pi(pg), g_i(pg)\Big) \\ &= pg \\ &= \psi_i\Big(\pi(p), g_i(p)\Big)g \\ &= \psi_i\Big(\pi(p), g_i(p)g\Big), \end{aligned}$$
then the statement follows, since ψ_i is a diffeomorphism.

6. Follows directly from the transitiveness of the action and the G-invariance of π.

□

The proof of following theorem is quite long and technical, and can be found in chapter 9 of [27].

Theorem 1.1.8. Let G be a compact Lie group acting smoothly and freely on a smooth manifold P. Then the orbit space P/G is a topological manifold of dimension $\dim P - \dim G$, and has a unique smooth structure with the property that the quotient map $\pi : P \mapsto P/G$ is a smooth submersion.

Corollary 1.1.9. Let $P(M, G)$ be a principal bundle with G compact. Then M is diffeomorphic with P/G and $\dim M = \dim P - \dim G$.

1.2 Constructing fibre bundles

In this section we construct a fibre bundle E given manifolds M and F and a Lie group G. Furthermore, we give some constructions of new bundles from existing bundles.

Theorem 1.2.1. Let M and F be manifolds, G a Lie group, a left action $G \times F \to F$, an open cover $\{U_i\}$ of M, and $g_{ij} : U_{ij} \to G$ smooth functions. Then we can construct a fibre bundle (E, π, M, F, G) by defining

$$E = X/\sim \tag{1.2.1}$$

where

$$X = \bigcup_i U_i \times F \tag{1.2.2}$$

The equivalence relation $\sim$ between $(x, f) \in U_i \times F$ and $(x', f') \in U_j \times F$ is defined by

$$(x, f) \sim (x', f') \iff x = x' \text{ and } f = g_{ij}(x)f'. \tag{1.2.3}$$

If we denote elements of E by $[(x, f)]$ with $x \in M$ and $f \in F$, the projection is given by

$$\pi([(x, f)]) = x, \tag{1.2.4}$$

while the local trivialization $\psi_i : U_i \times F \to \pi^{-1}(U_i)$ is given by

$$\psi_i(x, f) = [(x, f)]. \tag{1.2.5}$$

Proof. We prove this theorem by checking that π and ψ_i satisfy all axioms of a fibre bundle. Firstly, we have

$$\pi \circ \psi_i(x, f) = \pi([(x, f)]) = x,$$

Secondly, let $x \in U_{ij}$ and $f \in F$. Then $(x, g_{ij}(x)f) \in U_i \times F$ is equivalent with $(x, f) \in U_j \times F$, so we find

$$\psi_{i,x}^{-1} \circ \psi_{j,x}(f) = \psi_{i,x}^{-1}\Big([(x, f)]\Big) = \psi_{i,x}^{-1}\Big([(x, g_{ij}(x)f)]\Big) = g_{ij}(x)f,$$

thus $\psi_{i,x}^{-1} \circ \psi_{j,x} = g_{ij}$. □

Corollary 1.2.2. Let (E, π, M, F, G) be a fibre bundle with overlap functions $g_{ij} : U_{ij} \to G$. Then, using previous theorem, we can construct a principal bundle $P(M, G)$ with $P = X/\sim$ with $X = \bigcup_i U_i \times G$.

Since many interesting geometrical spaces turn out to be fibre bundles, it is natural to ask which fibre bundles are equivalent to each other. In order to define the equivalence between fibre bundles we need the concept of bundle maps.

Definition 1.2.3. Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ fibre bundles. Then a smooth map $\bar{f} : E' \to E$ is called a *bundle map* if it maps each fibre $F'_{x'}$ of E onto F_x of F, where $x \in M$ and $x' \in M'$.

Remark. A bundle map $\bar{f}$ mapping a fibre $F'_{x'} \subset E'$ onto $F_x \subset E$ naturally induces a smooth map $f : M' \to M$ such that $f(x') = x$. In other words, the diagram

$$\begin{tikzcd} E' \arrow[r, "\bar{f}"] \arrow[d, "\pi'"'] & E \arrow[d, "\pi"] \\ M' \arrow[r, "f"] & M \end{tikzcd}$$

commutes.

Definition 1.2.4. Two bundles $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ are equivalent if there is a bundle map $\bar{f} : E' \to E$ such that both $\bar{f}$ and the induced map $f : M' \to M$ are diffeomorphisms.

Since vector bundles and principal bundles are fibre bundles with a specific fibre, it is clear the notion of equivalence of fibre bundles can directly be translated to the case of vector bundles and principal bundles.

From existing bundles we can construct new bundles.

Theorem 1.2.5. Let $\pi : E \to M$ be a fibre bundle with fibre F. Then given a map $f : N \to M$, the space

$$f^*E = \{(y, p) \in N \times E : f(y) = \pi(p)\},$$

can be made into a fibre bundle over N with fibre F, called the *pullback bundle*, as follows. We define the projection with projection $\pi_f : f^*E \to N$ by $(y, p) \mapsto y$. Given local trivializations (U_i, ψ_i) for E, we define local trivializations $\left(f^{-1}(U), \psi_i^f\right)$ for f^*E by

$$\psi_i^f(y, f_i) = \left(y, \psi_{i,f(y)}(f_i)\right), \tag{1.2.6}$$

for $y \in f^{-1}(U_i)$, $f_i \in F$. The inverse of ψ_i^f is given by

$$(\psi_i^f)^{-1}(y, p) = \left(y, \psi_{i,f(y)}^{-1}(p)\right). \tag{1.2.7}$$

The transition functions are given by

$$g_{ij}^f(y) = g_{ij}\left(f(y)\right), \tag{1.2.8}$$

for $y \in f^{-1}(U_{ij})$.

Proof. It is clear that ψ_i^f is a map $f^{-1}(U_i) \times F \to \pi_f^{-1}(U_i)$ satisfying $\pi_f \circ \psi_i^f(y, f_i) = y$ and that its inverse is given by (1.2.7). So we only have to show that (1.2.8) are indeed the transition functions for f^*E. By definition of the transition functions we find

$$g_{ij}^f(y)f_j = (\psi_{i,y}^f)^{-1} \circ \psi_{j,y}^f(f_j) = (\psi_{i,y}^f)^{-1}\left(y, \psi_{j,f(y)}(f_j)\right) = \psi_{i,f(y)}^{-1} \circ \psi_{j,f(y)}(f_j) = g_{ij}\left(f(y)\right)f_j,$$

last equality by definition of the transition functions for E. ∎

The proof of next theorem can be found in [31].

Theorem 1.2.6. Let $E \xrightarrow{\pi} M$ be a fibre bundle with fibre F and let f and g be homotopic maps from N to M. Then f^*E and g^*E are equivalent bundles over N.

Corollary 1.2.7. Let $E \xrightarrow{\pi} M$ be a fibre bundle with M contractible. Then E is the trivial bundle over M.

Another useful construction is the product bundle of two vector bundles.

Definition 1.2.8. Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be two vector bundles with fibres F and F' respectively. Then the *product bundle* $E\times E' \xrightarrow{\pi\times\pi'} M\times M'$ is a fibre bundle with fibre $F\oplus F'$. If $M = M'$, we can define the *Whitney sum bundle* $E\oplus E'$ as the pullback bundle $E\times E'$ over M by $f : M \to M\times M$ defined by $f(x) = (x, x)$. The structure group of $E\oplus E'$ is given by

$$G^{E\oplus E'} = \left\{ \begin{pmatrix} g^E & 0 \\ 0 & g^{E'} \end{pmatrix} : g^E \in G^E, g^{E'} \in G^{E'} \right\}. \tag{1.2.9}$$

The transition functions $g_{ij}^{E\oplus E'}$ of $E\oplus E'$ are given by

$$g_{ij}^{E\oplus E'}(x) = \begin{pmatrix} g_{ij}^E(x) & 0 \\ 0 & g_{ij}^{E'}(x) \end{pmatrix}.$$

Remark. We have $E\oplus E' = \{(x, p, p') \in M\times E\times E' : f(x) = \pi\times\pi'(p, p;)\}$. Since $f(x) = \pi\times\pi'(p, p')$ is equivalent with $\pi(p) = \pi'(p') = x$, we see that we can find another bundle $\bar{E}$ over M defined by

$$\bar{E} = \{(p, p') \in E\times E' : \pi(p) = \pi'(p)\}. \tag{1.2.10}$$

Since $\bar{f} : E\oplus E' \to \bar{E}$ given by $\bar{f}(x, p, p') = (p, p')$ is clearly a diffeomorphic bundle map with diffeomorphic induced map $M \to M$ (which is actually the identity), we see that we could give an equivalent definition of $E\oplus E'$ by (1.2.10). Similarly, we can define the Whitney sum of vector bundles $E_1, \ldots, E_n$ over M by

$$E_1 \oplus \ldots \oplus E_n = \{(p_1, \ldots, p_n) \in E_1\times\ldots\times E_n : \pi_1(p_1) = \ldots = \pi_n(p_n)\}.$$

1.3 Sections and smooth lifts

Definition 1.3.1.

1. Let $\pi : E \to M$ be a fibre bundle. A *local section* is a map $s : U \subset M \to E$ be a differential map such that $\pi\circ s = \mathrm{Id}$. If $U = M$, then s is called a *global section*.

2. The space of smooth sections $s : M \to E$ is denoted with $\Gamma^\infty(M, E)$. Let F be a manifold, then an F-valued section $s : M \to E\otimes F$ is defined as a map which can me written as $s(x) = s_E(x)\otimes s_F(x) \subset E\otimes F$ satisfying $\pi\circ s_E = \mathrm{Id}$. We denote that space of F-valued sections with $\Gamma^\infty(M, E\otimes F)$.

3. Let X be a manifold and F either a vector space or a manifold. Then we define the F-valued forms on an open subset U of X by $\Omega^r(U, F) = \Gamma^\infty\big(U, \Lambda^r(T^*X)\otimes F\big)$, or equivalently, by the set of alternating r-linear maps α such that at any point $x\in U$ we have $\alpha|_x : T_xX\times T_xX\times\ldots\times T_xX \to F$. If $F = \mathbb{R}$, we write $\Omega^r(U, F) = \Omega^r(U)$.

Remark. If $U = X$, it is clear that $\Lambda^r(T^*X)$ is a bundle over X. On the other hand, if U is a proper subset of X, it is not directly clear that $\Lambda^r(T^*X)$ is a bundle over U, and so neither what a section from U to $\Lambda^r(T^*X)$ is. However, Proposition 3.6 of [27] provides an isomorphism between T_xU with T_xX, which allows us to identify $\Lambda^r(T^*U)$ with the subset $\pi^{-1}(U)$ of $\Lambda^r(T^*X)$, where π is the projection of $\Lambda^r(T^*X)$ onto M. So a section from U into $\Lambda^r(T^*U)$ can naturally be identified with a section into $\Lambda^r(T^*X)$.

Proposition 1.3.2. Let $P(M,G)$ be a principal bundle. Then every local trivialization $\psi : U \times G \to \pi^{-1}(U)$ defines a local section $s : U \to P$ by

$$s(x) = \psi(x, e). \tag{1.3.1}$$

Conversely, a local section $s : U \to P$ defines a local trivialization $\psi : U \times G \to \pi^{-1}(U)$ by

$$\psi(x, g) = s(x)g. \tag{1.3.2}$$

Proof. Let $\psi : U \times G \to \pi^{-1}(U)$ be a trivialization. Then $s : U \to P$ defined by (1.3.1) is indeed a section, since $\pi \circ \psi(x, g) = x$ for all $x \in M$ and $g \in G$. Conversely, given a local section $s : U \to P$, then $\psi : U \times G \to \pi^{-1}(U)$ defined by (1.3.2) is a local trivialization, for we have

$$\begin{aligned} \pi\Big(\psi(x,g)\Big) &= \pi(\psi(x,e)g) \\ &= \pi(pg) \\ &= \pi(p) \\ &= x, \end{aligned}$$

where we used the equivariance of π. Notice that a global section implies the existence of a global trivialization, and thus trivialness of the fibre. ☐

Proposition 1.3.3. Let $s_i : U_i \to P$ be local sections defined as corresponding to local trivializations ψ_i by (1.3.1). Then we have for all $x \in U_i$ and $p \in \pi^{-1}(U_i)$

$$s_i(x)g_i(p) = p \tag{1.3.3}$$

$$g_i(s_i(x)g) = g \tag{1.3.4}$$

$$g_i \circ s_i(x) = e \tag{1.3.5}$$

$$g_i \circ s_j(x) = g_{ij}(x) \qquad [x \in U_{ij}] \tag{1.3.6}$$

$$s_i \circ \pi(p) = pg_i(p)^{-1} \tag{1.3.7}$$

$$s_i(x)g_{ij}(x) = s_j(x) \qquad [x \in U_{ij}]. \tag{1.3.8}$$

Proof. We have $s_i(x)g_i(p) = \psi_i(x,e)g_i(p) = \psi_i\Big(x, g_i(p)\Big) = p$, so the first identity holds. The second identity follows from (1.3.2) and by definition of g_i. The third is a special case of the second identity. Notice that the third is also a special case of the fourth identity, which holds since

$$\begin{aligned} g_i \circ s_j(x) &= g_i \circ \psi_j(x, e) \\ &= g_i \circ \psi_i\Big(x, g_{ij}(x)\Big) && \text{[By Proposition 1.1.3]} \\ &= g_{ij}(x). \end{aligned}$$

The fifth identity holds, since

$$\begin{aligned} s_i \circ \pi(p) &= \phi_i(\pi(p), e) \\ &= \phi_i\Big(\pi(p), g_i(p)\Big)g_i(p)^{-1} \\ &= pg_i(p)^{-1}, \end{aligned}$$

Finally,

$$\begin{aligned} s_j(x) &= \psi_j(x,e) \\ &= \psi_i(x, g_{ij}(x)e) && \text{[by (1.1.2)]} \\ &= \psi_i(x, eg_{ij}(x)) \\ &= \psi_i(x,e)g_{ij}(x) && \text{[by (1.1.6)]} \\ &= s_i(x)g_{ij}(x), \end{aligned}$$

which proves last identity. □

We can use sections to prove existence of lifts. First we need the following proposition:

Proposition 1.3.4. Let $P(M,G)$ be a principal bundle and $F : M \to X$ a map to a smooth manifold X. Then F is smooth if and only if $F \circ \pi$ is smooth:

$$\begin{tikzcd} P \arrow[d, "\pi"'] \arrow[dr, "F\circ\pi"] & \\ M \arrow[r, "F"'] & X \end{tikzcd}$$

Proof. If F is smooth, then $F \circ \pi$ is smooth, since it is a composition of smooth maps. Conversely, if $F \circ \pi$ is smooth, let $x \in M$ and take a coordinate neighborhood U_i of x and a section $s_i : U_i \to P$. Then

$$F|_{U_i} = F \circ \mathrm{Id}|_{U_i} = F \circ (\pi \circ s_i) = (F \circ \pi) \circ s_i,$$

so $F|_{U_i}$ is a composition of smooth maps. Therefore, F is smooth in a neighborhood of each point x, so it is smooth. □

Proposition 1.3.5. Let $P(M,G)$ be a principal bundle and $\widetilde{F} : P \to X$ a smooth map which satisfies $\widetilde{F}(pg) = \widetilde{F}(p)$. Then there is a unique smooth map $F : M \to X$ such that $F \circ \pi = \widetilde{F}$. In other words, F is the smooth map making following diagram commute:

$$\begin{tikzcd} P \arrow[d, "\pi"'] \arrow[dr, "\widetilde{F}"] & \\ M \arrow[r, "F"'] & X \end{tikzcd}$$

F is given by $F(x) = \widetilde{F}(p)$, where $p \in \pi^{-1}(x)$ may arbitrarily be chosen.

Proof. F defined as above is well-defined; take $q \in \pi^{-1}(x)$, then by the transitiveness of the right action of G on the fiber, there is a $g \in G$ such that $q = pg$, and we see that $\widetilde{F}(q) = \widetilde{F}(pg) = \widetilde{F}(p)$. By construction, we have $F \circ \pi = \widetilde{F}$ and by the previous proposition we find that F is smooth. □

1.4 Lie theory, representations and the Maurer-Cartan form

Groups are used in mathematics in order to describe symmetries of geometrical objects. In physics most systems are described by (partial) differential equations, where the space of solutions has in most cases certain symmetries described by a Lie group, which we will define now.

Definition 1.4.1. A smooth manifold G is called a *Lie group* if it is also a group with a group structure such the multiplication map $(g,h) \mapsto gh$ and the inversion map $g \mapsto g^{-1}$ are smooth. A Lie subgroup H of G is a subgroup in algebraic sense which is closed in G.

In this thesis we shall restrict us to matrix Lie groups, which are Lie subgroups of $\mathrm{GL}(n, \mathbb{C})$. Furthermore, all Lie groups used are connected and compact, unless otherwise stated.

The next notion we will introduce is that of a Lie algebra, which is close connected to Lie groups.

Definition 1.4.2. A *Lie algebra* is a vector space L endowed with a bilinear product $[\cdot, \cdot] : L \times L \to L$, called the *Lie product* such that

1. $[X, X] = 0$ for all $X \in L$
2. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in L$.

A *Lie subalgebra* of L is a subspace of L which is closed under the Lie product. Let K be another Lie algebra. An *ideal* I of a Lie algebra L is a subspace of L such that $[X, Y] \in I$ if $X \in I$ and $Y \in L$. A linear map $F : L \to K$ is called a *Lie algebra homomorphism* if $F([X, Y]) = [F(X), F(Y)]$. The image $F(L)$ is in that case a Lie subalgebra of K. A Lie algebra L is called *abelian* if $[X, Y] = 0$ for all $X, Y \in L$. A Lie algebra L is called *simple* if it is non-abelian, and its only ideals are 0 and L. A Lie algebra L is called *semisimple* if it is the direct sum of simple Lie algebras.

Example 1.4.3. Let $M_n(\mathbb{R})$ be the space of $n \times n$ matrices with values in $\mathbb{R}$. Then $M_n(\mathbb{R})$ is a Lie algebra with the matrix commutator as Lie product.

Example 1.4.4. Let $\mathscr{X}(M)$ be the space of smooth vector fields on a manifold M. Then $\mathscr{X}(M)$ is a Lie algebra with the Lie bracket as Lie product.

The next lemma is proved in [27] as Corollary 4.17.

Lemma 1.4.5. Let $F : M \to N$ be a diffeomorphism. Then $F_* : \mathscr{X}(M) \to \mathscr{X}(N)$ is a Lie algebra homomorphism.

We shall see that we can assign a unique Lie algebra to each Lie group, which is not the space of all vector fields, for that is in some way too large to be interesting. However, we shall find a subspace which is more suitable. We start by defining the left- en right-translation of an element h in a Lie group G by $g \in G$,

$$\begin{aligned} L_g(h) &= gh; \\ R_g(h) &= hg. \end{aligned}$$

These maps induce maps in the tangent space

$$L_{g*} : \quad T_h G \to T_{gh} G; \tag{1.4.1}$$
$$R_{g*} : \quad T_h G \to T_{hg} G. \tag{1.4.2}$$

A vector field X is called *left-invariant* if it satisfies

$$L_{g*}(X|_h) = X|_{gh}. \tag{1.4.3}$$

Since L_g is a diffeomorphism, we can write this as $L_{g*} X = X$. Using Lemma 1.4.5, we see that for left-invariant X and Y we have $L_{a*}[X, Y] = [L_{a*}X, L_{a*}Y] = [X, Y]$. We see that the space of left-invariant vector fields is closed under the Lie bracket, so these vector fields forms a Lie algebra, which is small enough.

Definition 1.4.6. Let G be a Lie group. Then the Lie algebra consisting of the left-invariant vector fields on G is called the *Lie algebra of* G, denoted with $\mathrm{Lie}(G)$.

We shall give an alternative definition of Lie(G), which provides other useful information. For this, we need the coordinate function $x = (x^{ij}) : G \to M^n(\mathbb{C})$, latter space is viewed as a real space. We define $x^{ij}(g)$ is the matrix entry g^{ij} of $g \in G$, so we could see the function x as the inclusion of G into $M^n(\mathbb{C})$. Given this notation, we can define the pushforward $x_* : T_gG \to T_{x(g)}M_n(\mathbb{C}) \simeq M_n(\mathbb{C})$, which is often denoted with dx in the physical literature and so we will do. Notice that dx provides the embedding of T_gG into $M_n(\mathbb{C})$ by

$$X|_g = X^{ij}(g) \left.\frac{\partial}{\partial x^{ij}}\right|_g \mapsto X(g),$$

where $X(g)$ is the matrix with matrix entries $X^{ij}(g)$. We see that every vector field $X : G \to TG$ is determined by a map $X(\cdot) : G \to M_n(\mathbb{C})$. When no confusion possible, we write X instead of $X(\cdot)$. The corollary of following proposition gives information about the behaviour of $X(\cdot)$ if X is a left-invariant vector field.

Proposition 1.4.7. Let G be a Lie group with coordinate functions x^{ij}, and X, Y left-invariant vector fields. Writing $X|_g = X^{ij}(g) \left.\frac{\partial}{\partial x^{ij}}\right|_g \in T_gG$, we have

$$X|_g = L_{g*}X|_e = x^{ij}(g)X(e)^{jk} \left.\frac{\partial}{\partial x^{ik}}\right|_g,$$
$$[X,Y]|_g = x(g)^{ij}[X(e),Y(e)]^{jk} \left.\frac{\partial}{\partial x^{ik}}\right|_g,$$

where $[X(e), Y(e)]$ is the commutator of matrices.

Proof. See equations (5.113) and (5.114) of [25]. □

Since $x^{ij}(g)$ is nothing more than the coordinates of g, the previous propositions reads in matrix form:

Corollary 1.4.8. Let G be a Lie group and X, Y be left-invariant vector fields. Then the map $X : G \to M_n(\mathbb{R})$ satisfies

$$X(g) = L_gX(e) = gX(e)$$
$$[X,Y](g) = g[X(e),Y(e)].$$

These identities motivate us to the following proposition:

Proposition 1.4.9. Let G be a Lie group and let $\mathfrak{g}$ be the space of all matrices $X(e)$ with X a left-invariant vector field. Then $\mathfrak{g}$ is a Lie algebra with matrix commutation as Lie product and is isomorphic to Lie(G).

Proof. From (1.4.3) we find that a vector $A \in T_eG$ uniquely defines a left-invariant vector field X_A on G by setting $X_A|_a = L_{a*}A$, while a left-invariant vector field X defines a unique vector $A = X|_e \in T_eG$. If we take $g = e$ in last corollary, we see that $[X,Y](e) = [X(e),Y(e)]$, the map $X \mapsto X(e)$ is a Lie algebra homomorphism, thus an isomorphism. □

Remark. Since Lie(G) is isomorphic to $\mathfrak{g}$, we see that the map $L_{g*} : \mathfrak{g} \to \mathfrak{g}$ corresponding to the map L_{g*} : Lie(G) $\to$ Lie(G) is equal to L_g.

By the Lie algebra isomorphism between Lie(G) and $\mathfrak{g}$, we see that $\mathfrak{g}$ has the same dimension as the manifold G, since latter has the same dimension as T_eG. We have seen that given a vector $X|_e \in T_eG$, we can construct a left-invariant vector field by the map L_{g*}. The converse is also possible:

Definition 1.4.10. The map $\Theta : \mathscr{X}(G) \to \mathrm{Lie}(G)$ is defined as $\Theta|_g = L_{g^{-1}*}$ is called the *Maurer-Cartan form.* Since $\mathfrak{g}$ is isomorphic with $\mathrm{Lie}(G)$, the corresponding map $\theta(\cdot) : \mathscr{X}(G) \to \mathfrak{g}$ given by $\theta(g) = L_{g^{-1}*} = L_{g^{-1}}$ is also called the Maurer-Cartan form. In order to avoid the use of parentheses, we shall write $\theta|_g$ instead of $\theta(g)$. If no confusion is possible, we shall write θ instead of $\theta(\cdot)$, and $g^*\theta|_{g(p)}$ or $(g^*\theta)|_p$ instead of $g^*\theta(p)$, where p an element from some smooth manifold X and $g : X \to G$ a smooth map.

For the following proposition we shall use the alternative notation $\mathrm{d}f$ for the pushforward f_* of some map f with values in $GL(n, \mathbb{R})$. We will see in the second remark below the proposition that we can interpret $\mathrm{d}f$ as the exterior derivative of f, which justifies this notation.

Proposition 1.4.11. We have $\theta(a) = x(a)^{-1}\mathrm{d}x|_a$, where $x : G \hookrightarrow GL(n, \mathbb{C})$ is the inclusion. If $g : X \to G$ is a smooth map, then $g^*\theta(\cdot) = g^{-1}\mathrm{d}g$.

Proof. The Maurer-Cartan form is exactly the form which sends a vector field X to $X|_e$. So $\theta(\cdot)$ has to send X to $X(e)$, which is indeed the case:

$$\begin{aligned}\theta|_g(X|_g) &= x(g)^{-1}\mathrm{d}x|_g(X|_g)\\ &= x(g)^{-1}X(g)\\ &= g^{-1}X(g)\\ &= X(e).\end{aligned}$$

If $g : X \to G$ is smooth, we find

$$\begin{aligned}g^*\theta|_{g(p)}(X|_p) &= \theta|_{g(p)}(g_*X|_p)\\ &= x\Big(g(p)\Big)^{-1}\mathrm{d}x|_{g(p)}(g_*X|_p)\\ &= x\Big(g(p)\Big)^{-1}\mathrm{d}x|_{g(p)}(\mathrm{d}g|_pX|_p) && [\mathrm{d}g \text{ is an alternative notation for } g_*]\\ &= x\Big(g(p)\Big)^{-1}\mathrm{d}(x\circ g)|_p(X|_p)\\ &= g(p)^{-1}\mathrm{d}g|_p(X|_p) && [x \text{ is the identity on } G].\end{aligned}$$

□

Remarks.

1. Since $g^*\theta|_{g(p)}$ sends vector fields on a manifold X to elements in $\mathfrak{g}$, we see that $g^*\theta \in \Omega^1(X, \mathfrak{g})$.

2. The notation $g^{-1}\mathrm{d}g$ suggests that we deal with an exterior derivative of some map $g : X \to G$. It is indeed possible to give an alternative definition of $\mathrm{d}g$, which underpins this idea. First we remark that G is a subgroup of $GL(n, \mathbb{C})$, so we have component functions $g^{ij} : X \to \mathbb{R}$ on which the exterior derivative d_X of X can act. Then we define $\mathrm{d}_X g$ to be the matrix with entries $\mathrm{d}_X g^{ij}$, which allows us to define $g(x)^{-1}\mathrm{d}_X|_p g$ by the usual matrix product. If $x : G \to M^n(\mathbb{R})$ is the canonical embedding, we find $x^{-1}\mathrm{d}x$ is just the usual

matrix representation of θ, since in that case we have for a vector field $X \in \mathscr{X}(G)$:

$$\begin{aligned}
x(g)^{-1}\mathrm{d}x|_g(X|_g) &= x(g)^{-1}\mathrm{d}_G x|_g\left(X(g)^{lm}\left.\frac{\partial}{\partial x^{lm}}\right|_g\right)\\
&= \left\{\left(g^{-1}\right)^{ik}\left(\mathrm{d}_G x|_g\right)^{kj}X(g)^{lm}\left.\frac{\partial}{\partial x^{lm}}\right|_g\right\}^{ij}\\
&= \left\{\left(g^{-1}\right)^{ik}\delta^k_l\delta^j_m X(g)^{lm}\right\}^{ij}\\
&= g^{-1}X(g)\\
&= X(e).
\end{aligned}$$

By the exponent of a matrix we can give another relation between Lie groups and their Lie algebra. Next proposition can be found in [32] as Theorems 4.6, 5.1 and 5.12.

Proposition 1.4.12. Let G be a connected and compact Lie group with Lie algebra $\mathfrak{g}$ and let $X \in \mathfrak{g}$ (where we identify $X(e)$ with X), then

1. $\mathfrak{g} = \{X \in M^n(\mathbb{C}) : \exp(tX) \in G \text{ for all } t \in \mathbb{R}\}$.

2. For $X, Y \in \mathfrak{g}$, $[X,Y] = 0$ if and only if e^{tX} and e^{sY} commute for all $s,t \in \mathbb{R}$, in which case we have $e^{X+Y} = e^X e^Y$.

3. The map $t \mapsto \exp(tX)$ is a homomorphism: $\exp(tA)\exp(sA) = \exp\Big((t+s)\Big)A$.

4. If G is connected, $\exp\mathfrak{g}$ generates G as group, i.e. all elements $g \in G$ can be written as $g = e^{X_1}e^{X_2}\dots e^{X_n}$ where $X_1,\dots,X_n \in \mathfrak{g}$.

5. If G is connected and compact, then exp is surjective: every $g \in G$ can be written as $g = \exp(X)$ for some $X \in \mathfrak{g}$.

Example 1.4.13. The simple Lie algebra $\mathfrak{u}(n) = \{X \in M^n(\mathbb{C}) : X = -X^\dagger\}$ is the Lie algebra of the compact Lie group $\mathrm{U}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A = A^\dagger\}$.

Example 1.4.14. The simple Lie algebra $\mathfrak{su}(n) = \{X \in M^n(\mathbb{C}) : X = -X^\dagger, \mathrm{tr}X = 0\}$ is the Lie algebra of the compact Lie group $\mathrm{SU}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A = A^\dagger, \det A = 1\}$.

Definition 1.4.15. For $g \in G$, we define the *adjoint map* $\mathrm{Ad}_g : G \to G$ by $h \mapsto ghg^{-1}$. In other words $\mathrm{Ad}_g = L_g \circ R_{g-1}$, and it defines a representation $\mathrm{Ad} : G \to \mathrm{Diff}(G)$ by $g \mapsto \mathrm{Ad}_g$. Furthermore, we define the map $\mathrm{ad}_g : T_hG \to T_{ghg^{-1}}G$ by $\mathrm{ad}_g = \mathrm{Ad}_{g*} = L_{g*} \circ R_{g^{-1}*}$. Hence we find that $\mathrm{ad} : G \times \mathrm{Lie}(G) \to \mathrm{Lie}(G)$, which gives by $g \mapsto \mathrm{ad}_g$ a representation $G \to GL(\mathrm{Lie}(G))$, called the *adjoint representation*. Notice that ad induces a map $G \times \mathfrak{g} \to \mathfrak{g}$, which we also denote with ad.

Remarks.

1. The map $\mathrm{ad}_g : \mathfrak{g} \to \mathfrak{g}$ for all $g \in G$ is a Lie algebra homomorphism: we have

$$[\mathrm{ad}_g X, \mathrm{ad}_g Y] = \mathrm{ad}_g[X,Y] \tag{1.4.4}$$

for $X, Y \in \mathfrak{g}$, which follows directly from $\mathrm{ad}_g X = gXg^{-1}$ and $[X,Y] = XY - YX$.

2. The representation Ad is only faithful if the center $Z(G) = \{g \in G : gh = hg \text{ for all } h \in G\} = 0$. If G is connected, but not necessarily compact, the representation ad is faithful if the center $Z(\mathfrak{g}) = \{X \in \mathfrak{g} : [X,Y] = 0 \text{ for all } Y \in \mathfrak{g}\} = 0$. Connectedness of G allows us to write any element $g \in G$ as $e^{X_1}e^{X_2}\dots e^{X_n}$ for $X_i \in \mathfrak{g}$, so if $\mathfrak{g}$ has nontrivial center with $X \in Z(\mathfrak{g}) \setminus \{0\}$, then writing $g = e^X$, we find for any $A \in \mathfrak{g}$

$$\begin{aligned} \mathrm{ad}_g(A) &= gAg^{-1} \\ &= e^X A e^{-X} \\ &= Ae^X e^{-X} \\ &= A, \end{aligned}$$

hence $\mathrm{ad}_g = \mathrm{Id}$. So it is clear that if $\mathfrak{g}$ is centerless, we always have $\mathrm{ad}_g \neq \mathrm{Id}$ for $g \neq e$, thus ad is faithful. The condition that $\mathfrak{g}$ is centerless is fulfilled if $\mathfrak{g}$ is semisimple.

Lemma 1.4.16. Let G be a connected Lie group and $\mathfrak{g}$ its Lie algebra. Then $Z(\mathfrak{g}) = \mathrm{Lie}(Z(G))$.

Proof. Let $X \in \mathrm{Lie}(Z(G))$. Then $e^{tX}g = ge^{tX}$ for all $g \in G$, so in special we have $e^{tX}e^{sY} = e^{sY}e^{tX}$ for all $Y \in \mathfrak{g}$ and $t, s \in \mathbb{R}$. By Proposition 1.4.12, we find that $[X,Y] = 0$ for all $Y \in \mathfrak{g}$, so $X \in Z(\mathfrak{g})$.

Conversely, let $X \in Z(\mathfrak{g})$. Then we have $[X,Y] = 0$ for all $Y \in \mathfrak{g}$, thus we have $e^{tX}e^{sY} = e^{tX+sY} = e^{sY}e^{tX}$ for all $Y \in \mathfrak{g}$ and $t, s \in \mathbb{R}$. Now also by Proposition 1.4.12, all elements $g \in G$ can be written as $g = e^{X_1}e^{X_2}\dots e^{X_n}$ and since $[X, X_i] = 0$ ($X \in Z(\mathfrak{g})$), we find that e^{tX} commutes with all e^{X_i}, thus $e^{tX}g = ge^{tX}$ for all $g \in G$. But this means that $e^{tX} \in Z(G)$ for all $t \in \mathbb{R}$, whence $X \in \mathrm{Lie}(Z(G))$. ☐

Lemma 1.4.17. The Maurer-Cartan form θ and $\mathrm{ad} : G \times \mathfrak{g} \to \mathfrak{g}$ satisfy the following identities

$$R_a^*\theta|_{ga} = \mathrm{ad}_{a^{-1}}\theta|_g \tag{1.4.5}$$

$$\mathrm{d}\theta = -\frac{1}{2}[\theta, \theta], \tag{1.4.6}$$

$$\mathrm{d}(\mathrm{ad}_{g^{-1}}\alpha) = \mathrm{ad}_{g^{-1}}\mathrm{d}\alpha - [\mathrm{ad}_{g^{-1}}\alpha, g^*\theta], \tag{1.4.7}$$

where $\alpha \in \Omega^1(X, \mathfrak{g})$.

Proof. We have

$$\begin{aligned} R_a^*\theta|_{ga} &= \theta|_{ga}R_{a*} \\ &= L_{(ga)^{-1}*}R_{a*} \\ &= L_{a^{-1}*}L_{g^{-1}*}R_{a*} \\ &= \mathrm{Ad}_{a^{-1}*}L_{g^{-1}*} \\ &= \mathrm{ad}_{a^{-1}} \circ \theta|_g, \end{aligned}$$

which proves the first identity. The second identity can be proven without using that G is a a matrix group, see for instance Theorem 5.3(b) of [25]. Assuming that G is a matrix group, we find

$$\begin{aligned} \mathrm{d}\theta &= \mathrm{d}(x^{-1}\mathrm{d}x) \\ &= \mathrm{d}(x^{-1}) \wedge \mathrm{d}x \\ &= -x^{-1}\mathrm{d}xx^{-1} \wedge \mathrm{d}x \\ &= -\theta \wedge \theta, \end{aligned} \tag{1.4.8}$$

where we used (A.2.5) in the third equality. It follows by (A.1.9) that this equals $-\frac{1}{2}[\theta,\theta]$. The third identity follows from

$$\begin{aligned}
\mathrm{d}(\mathrm{ad}_{g^{-1}}\alpha) &= \mathrm{d}(g^{-1}\alpha g) \\
&= \mathrm{d}(g^{-1})\wedge\alpha g + g^{-1}\mathrm{d}\alpha g - g^{-1}\alpha\wedge\mathrm{d}g && \text{[By Proposition } A.2.2] \\
&= -g^{-1}(\mathrm{d}g)g^{-1}\wedge\alpha g + g^{-1}\mathrm{d}\alpha g - g^{-1}\alpha\wedge\mathrm{d}g && \text{[By } (A.2.5)] \\
&= -g^{-1}(\mathrm{d}g)g^{-1}\wedge\alpha g + \mathrm{ad}_{g^{-1}}\mathrm{d}\alpha - g^{-1}\alpha g\wedge g^{-1}\mathrm{d}g \\
&= -g^{*}\theta\wedge\mathrm{ad}_{g^{-1}}\alpha + \mathrm{ad}_{g^{-1}}\mathrm{d}\alpha - \mathrm{ad}_{g^{-1}}\alpha\wedge g^{*}\theta \\
&= \mathrm{ad}_{g^{-1}}\mathrm{d}\alpha - [\mathrm{ad}_{g^{-1}}\alpha, g^{*}\theta] && \text{[By } (A.1.2).]
\end{aligned}$$

□

Definition 1.4.18. Let $\mathfrak{g}$ be the Lie algebra of a matrix Lie group G and let $\mathrm{ad}:\mathfrak{g}\to\mathrm{End}(\mathfrak{g})$ be the adjoint representation of $\mathfrak{g}$. Then the bilinear form

$$\kappa(X,Y) = \mathrm{tr}(\mathrm{ad}X\circ\mathrm{ad}Y), \tag{1.4.9}$$

where $X,Y\in\mathfrak{g}$ is called the *Killing form*

Next theorem can be found in [32] as Theorem 6.16.

Theorem 1.4.19. Let $\mathfrak{g}$ be a semisimple Lie algebra. Then

1. κ is negative definite on $[\mathfrak{g},\mathfrak{g}] := \{[X,Y] : X,Y\in\mathfrak{g}\}$;
2. If $\mathfrak{g}\subset\mathfrak{u}(n)$, then there is a positive $c\in\mathbb{R}$ such that $\kappa(X,Y) = c\,\mathrm{tr}(XY)$ for all $X,Y\in\mathfrak{g}$.

It is easy to see that $[\mathfrak{g},\mathfrak{g}]=\mathfrak{g}$ if $\mathfrak{g}$ is simple. Corollary 5.2 of [19] assures that $[\mathfrak{g},\mathfrak{g}]=\mathfrak{g}$ even if $\mathfrak{g}$ is semisimple. Hence we see that $-\kappa$ is an inner product on semisimple Lie algebras. Furthermore, we see that $(X,Y)\mapsto -\mathrm{tr}(XY)$ is a inner product on $\mathfrak{g}\subset\mathfrak{u}(n)$.

Finally, we state a useful fact about the center of a matrix Lie group G.

Proposition 1.4.20. Let G be a connected compact matrix Lie group with semisimple Lie algebra $\mathfrak{g}$. Then the center $Z(G)=\{g\in G : gh=hg\ \forall h\in G\}$ is discrete.

Proof. Since $\mathfrak{g}$ is semisimple, we have $Z(\mathfrak{g})=0$. Then the statement follows immediately by Lemma 1.4.16 and Proposition 1.4.12. □

1.5 Associated bundles

Definition 1.5.1. Let $P(M,G)$ be a principal bundle, F a space, and let $\rho: G\to\mathrm{Aut}(F)$ be a representation. If F is a manifold, we take $\mathrm{Aut}(F)=\mathrm{Diff}(F)$ and if F is a vector space, we let $\mathrm{Aut}(F)=\mathrm{GL}(F)$. Then we define the action $G\times F\to F$ given by $(g,f)\mapsto\rho(g)f$, which we sometimes will denote with $(g,f)\mapsto g\cdot f$.

Lemma 1.5.2. Let $P(M,G)$ be a principal bundle, F a space, and let $\rho: G\to\mathrm{Aut}(F)$ be a faithful representation. Then the action defined in previous definition is a free action.

Proof. Trivial. □

Definition 1.5.3. Let $P(M,G)$ be a principal bundle with G compact, F a space, and let $\rho: G\to\mathrm{Aut}(F)$ be a faithful representation. Then we define an action of G on $P\times F$ by $(p,f)\mapsto(pg,\rho(g)^{-1}f)$, which is free, since the actions $(g,x)\to gx$ and $(g,f)\to\rho(g)f$ are free. We denote the space of equivalence classes under this action with E_ρ. That is $E_\rho=\{[(p,g)] : p\in P, f\in F\}$, where $[(p,f)]$ is an equivalence class defined by the equivalence relation $(p,f)\tilde{(}pg,\rho(g)^{-1}f)$.

Proposition 1.5.4. E_ρ is a manifold.

Proof. This follows directly from Theorem 1.1.8 since $E_\rho = (P \times F)/G$ with G compact and the action of G on $P \times F$ is free. $\square$

Theorem 1.5.5. Let $P(M,G)$ be a fibre bundle with G compact, F a manifold, and let $\rho : G \to \mathrm{Aut}(F)$ be a faithful representation. Then the space $P \times_\rho F := E_\rho$ can be made into a fibre bundle with base space M and fibre F, called the *associated fibre bundle* of $P(M,G)$. If F is a vector space, $P \times_\rho F$ is called the *associated vector bundle*. We make $P \times_\rho F$ into a bundle by defining the projection π_ρ of $P \times_\rho F$ onto M by

$$\pi_\rho(p,v) = \pi(p), \tag{1.5.1}$$

and local trivializations $\psi_{\rho_i} : U_i \times F \to \pi_\rho^{-1}(U_i)$ by

$$\psi_{\rho_i}(x,f) = [(s_i(x), f)], \tag{1.5.2}$$

where s_i is the section with respect to the trivialization $\psi_i : U_i \times G \to \pi^{-1}(U_i)$ defined in (1.3.1). The inverse is given by

$$\psi_{\rho_i}^{-1}\Big([(p,f)]\Big) = \Big(\pi(p), \rho\Big(g_i(p)\Big)f\Big), \tag{1.5.3}$$

and the transition functions by

$$\rho_{ij}(x) = \rho\Big(g_{ij}(x)\Big). \tag{1.5.4}$$

Proof. The projection π_ρ is well defined since $\pi(p) = \pi(pg)$ implies

$$\begin{aligned} \pi_\rho(pg, \rho(g)^{-1}v) &= \pi(pg) \\ &= \pi(p) \\ &= \pi_\rho(p,v). \end{aligned}$$

The map ψ_{ρ_i} is clearly smooth, for it is the composition of smooth functions. We prove that ψ_{ρ_i} is well-defined: let $[(pa, \rho(a)^{-1}f)]$ be another representative of $[(p,f)]$, then we find

$$\begin{aligned} \psi_{\rho_i}^{-1}\Big([(pa, \rho(a)^{-1}f)]\Big) &= (\pi(pa), \rho\Big(p_i(pa)\Big)\rho(a)^{-1}f)] \\ &= \Big(\pi(p), \rho\Big(p_i(p)a\Big)\rho(a)^{-1}f)\Big) \\ &= \psi_{\rho_i}^{-1}\Big([(p,f)]\Big). \end{aligned}$$

We have indeed that ψ_{ρ_i} and $\psi_{\rho_i}^{-1}$ are inverses of each other, since by Proposition 1.3.3:

$$\begin{aligned} \psi_{\rho_i}^{-1} \circ \psi_{\rho_i}(x,f) &= \Big(\pi \circ s_i(x), \rho\Big(g_i \circ s_i(x)\Big)f\Big) \\ &= (x,e), \end{aligned}$$

while we also have

$$\begin{aligned} \psi_{\rho_i} \circ \psi_{\rho_i}^{-1}\Big([(p,f)]\Big) &= \Big[\Big(s_i \circ \pi(p), \rho\Big(g_i(p)\Big)f\Big)\Big] \\ &= \Big[\Big(pg_i(p)^{-1}, \rho\Big(g_i(p)\Big)f\Big)\Big] \\ &= [(p,f)]. \end{aligned}$$

Hence we see that ψ_{ρ_i} is clearly a diffeomorphism between $U_i \times F$ and $\pi_\rho^{-1}(U_i)$. The transition function $\rho_{ij}(x)$ for ψ_{ρ_i} can be found by looking at the transition functions for ψ_i. So, let $x \in U_{ij}$ and $f_i, f_j \in F$ such that $\psi_{\rho_i}(x, f_i) = \psi_{\rho_j}(x, f_j)$. This is equivalent with

$$\begin{aligned}
[(s_i(x), f_j)] &= [(s_j(x), f_j)] \\
&= [(s_i(x)g_{ij}(x), f_j)] && \text{[By Proposition 1.3.3]} \\
&= \Big[\Big(s_i(x), \rho\Big(g_{ij}(x)\Big) f_j\Big)\Big],
\end{aligned}$$

so we can conclude that indeed $\rho_{ij}(x) = \rho\Big(g_{ij}(x)\Big)$. ☐

Remarks.

1. Since we sometimes denote the action on the fibre $G \times F \to F$ is given by $(g, f) \mapsto \rho(g)f$ by $(g, f) \mapsto g \cdot f$, the transition function is given by $\rho_{ij}(x) = g_{ij}(x)$ in this notation.
2. In most instances, we shall take $G = \mathrm{SU}(n)$, which is a connected, centerless matrix group with semisimple Lie algebra $\mathfrak{su}(n)$. So by the second remark below definition 1.4.15, $\mathrm{SU}(n)$ has faithful representations Ad en ad.

Theorem 1.5.6. Let F be a group. Then we can make $\pi_\rho^{-1}(x)$ into a group by defining

$$[(p, v)][(p, w)] = [(p, vw)], \tag{1.5.5}$$

with $p \in P$, $v, w \in F$. If F is a vector space, then we can make $\pi_\rho^{-1}(x)$ into a vector space by defining addition respectively scalar multiplication by

$$[(p, v)] + [(p, w)] = [(p, v + w)]; \tag{1.5.6}$$
$$\lambda[(p, v)] = [(p, \lambda v)], \tag{1.5.7}$$

where $p \in P$, $v, w \in F$ and $\lambda \in \mathbb{R}$.

So if F is a vector space, the associated vector bundle is indeed a vector bundle.

Proof. The operations on $\pi_\rho^{-1}(x)$ defined above are well-defined. In order to show this for (1.5.5), let $[(p', v')] = [(p, v)]$ and $[(p', w')] = [(p, w)]$. This means that there is a $g \in G$ such that $p' = pg$, $v' = \rho(g)^{-1}v$ and $w' = \rho(g)^{-1}w$. Then we find

$$\begin{aligned}
[(p', v')][(p', w')] &= [(p', v'w')] \\
&= [(pg, \rho(g)^{-1}v\rho(g)^{-1}w)] \\
&= [(pg, \rho(g)^{-1}(vw))] && [\rho \text{ is an automorphism}] \\
&= [(p, vw)] \\
&= [(p, v)][(p, w)].
\end{aligned}$$

Equation (1.5.6) is the abelian case of (1.5.5). For (1.5.7) let p, p', v, v' as above, and let $\lambda \in \mathbb{R}$. Then

$$\begin{aligned}
\lambda[(p', v')] &= [(p', \lambda v')] \\
&= [(pg, \rho(g)^{-1}\lambda v)] \\
&= [(pg, \lambda\rho(g)^{-1}v)] \\
&= \lambda[(pg, \rho(g)^{-1}v)] \\
&= \lambda[(p, v)],
\end{aligned}$$

which proves that the operations are well-defined. If F is a group, the neutral element is given by $[(p, e)]$ and the inverse of $[(p, v)]$ is given by $[(p, v^{-1})]$. If F is a vector space, the neutral element and the inverse of $[(p, v)]$ are given by $[(p, 0)]$ respectively $[(p, -v)]$. ☐

Theorem 1.5.7. If F defined as above is a group, then the fibre $\pi_\rho^{-1}(x)$ is isomorphic to F. If F is a vector space, then the fibre is linear isomorphic to F.

Proof. In both cases, we establish the isomorphism $f : F \to \pi_\rho^{-1}(x)$ by $v \mapsto [(p, v)]$, where we have chosen a fixed $p \in \pi^{-1}(x)$. If F is a group, we have $f^{-1}[(p, e)] = \{v \in F : f(v) = e\} = \{v \in F : [(p, v)] = [(p, e)]\} = \{e\}$, f is injective. Notice that if F is a vector space, the same argument is valid, but with e replaced by 0. Furthermore, we find that f is surjective: Let $[(p, v)] \in \pi_\rho^{-1}(x)$ and let $p' \in \pi^{-1}(x)$ such that $f(w) = [(p', w)]$ for all $w \in F$. Since G works transitively on the fibre $\pi^{-1}(x)$, we have a $g \in G$ such that $p' = pg$. Now choose $w = \rho(g)^{-1}v$, then we find

$$\begin{aligned} f(w) &= [(p', w)] \\ &= [(pg, \rho(g)^{-1}v)] \\ &= [(p, g)]. \end{aligned}$$

Finally, if F has a group structure, we have $f(vw) = [(p, vw)] = [(p, v)][(p, w)] = f(v)f(w)$, so f is an isomorphism of groups. If F is a vector space, we find

$$\begin{aligned} f(\alpha v + \beta w) &= [(p, \alpha v + \beta w)] \\ &= \alpha[(p, v)] + \beta[(p, w)] \\ &= \alpha f(v) + \beta f(w), \end{aligned}$$

so f is linear. □

Example 1.5.8. The two most important associated bundles we will encounter are ad $P := P \times_{\mathrm{ad}} \mathfrak{g}$ and Ad $P := P \times_{\mathrm{Ad}} G$.

Remark. By Corollary 1.2.2 it is also possible to associate a principal bundle to a vector bundle. These operations are exactly each others inverses. P and $P \times_\rho F$ share the same coordinate neighborhoods U_i, and by the first remark below Theorem 1.5.5, they share the same overlap functions g_{ij}. The associated principal bundle of $P \times_\rho F$ is then given by $X/\sim$, where $X = \bigcup_i U_i \times G$ and where the relation $\sim$ is defined by saying that $(x, g) \in U_j \times F$ is equivalent with $(x, g_{ij}(x)g) \in U_i \times G$, whence $X/\sim$ is exactly P.

Corollary 1.5.9. Every vector bundle is the associated vector bundle of some principal bundle.

Proposition 1.5.10. Let $f_i : U_i \to F$ be a family of functions satisfying

$$f_i(x) = \rho\Big(g_{ij}(x)\Big) f_j(x) \tag{1.5.8}$$

for $x \in U_{ij}$ and let s_i be the section defined in (1.3.1). Then $f : M \to P \times_\rho F$ defined by

$$f(x) = \Big[\Big(s_i(x), f_i(x)\Big)\Big] \tag{1.5.9}$$

is a smooth section $M \to P \times_\rho F$. Conversely, a smooth section $f \in \Gamma^\infty(M, P \times_\rho F)$ induces functions $f_i : U_i \to F$ satisfying (1.5.8) such that (1.5.9) holds.

Proof. Let $f_i : U_i \to F$ be a family of functions satisfying (1.5.8). Then we define $f : M \to P \times_\rho F$ by (1.5.9). This is independent of the choice of trivialization, since

$$\begin{aligned} f(x) &= \Big[\Big(s_i(x), f_i(x)\Big)\Big] \\ &= \Big[\Big(s_i(x), \rho\Big(g_{ij}(x)\Big) f_j(x)\Big)\Big] \\ &= \Big[\Big(s_i(x)g_{ij}(x), f_j(x)\Big)\Big] \\ &= \Big[\Big(s_j(x), f_j(x)\Big)\Big]. \end{aligned}$$

By $\pi_\rho([(p,g)]) = \pi(p)$, we find that f is a section, since

$$\begin{aligned}\pi_\rho \circ f(x) &= \pi_\rho\Big(\Big[\Big(s_i(x), f_i(x)\Big)\Big]\Big)\\ &= \pi \circ s_i(x)\\ &= x.\end{aligned}$$

Conversely, let $f \in \Gamma^\infty(M, P \times_\rho F)$ and let $\psi_{\rho_i} : U_i \times F \to \pi_\rho^{-1}(U_i) \subset P \times_\rho F$ be the trivialization induced by the trivialization $\psi_i : U_i \times G \to \pi^{-1}(U_i)$. Then we define $f_i(x) = \psi_{\rho_i,x}^{-1} \circ f(x)$, which is clearly smooth, for it is the composition of smooth functions. Notice that $\pi_{\rho_i} \circ f(x) = x$ since f is a section, so $f(x) \in \pi_{\rho_i}^{-1}(x) \subset \pi_{\rho_i}^{-1}(U_i)$, whence f_i is well defined.

If $x \in U_{ij}$, we have by definition of the transition functions and (1.5.4) that the transition function $\rho\Big(g_{ij}(x)\Big)$ equals $\psi_{\rho_i}^{-1} \circ \psi_{\rho_j}$, hence we find

$$\begin{aligned}\rho\Big(g_{ij}(x)\Big) f_j(p) &= \psi_{\rho_i}^{-1} \circ \psi_{\rho_j} f_j(p)\\ &= \psi_{\rho_i}^{-1} \circ f(x)\\ &= f_i(x).\end{aligned}$$

Furthermore, we have

$$\begin{aligned}f(x) &= \psi_{\rho_i,x} \circ \psi_{\rho_i,x}^{-1} \circ f(x)\\ &= \psi_{\rho_i,x} \circ f_i(x)\\ &= \Big[\Big(s_i(x), f_i(x)\Big)\Big],\end{aligned}$$

where we used (1.5.2) in the last equality. This proves that (1.5.9) holds. □

Corollary 1.5.11. Let $P(M,G)$ be a principal bundle and $\rho : G \to \mathrm{Aut}(F)$ a faithful representation for some space F. Then a family of forms $\alpha_i \in \Omega^r(U_i, F)$

$$\alpha_i|_x = \rho\Big(g_{ij}(x)\Big) \circ \alpha_j|_x \tag{1.5.10}$$

define a form $\alpha \in \Omega^r(M, P \times_\rho F)$ and vice versa.

2 Gauge theories

2.1 Connections on principal bundles

In differential geometry the notion of a connection on a principal bundle is very important, for it gives information about how tensors are transported along a curve. Furthermore, it gives rise to a differential operator on the principal bundle, called the covariant derivative, which is an extension of the Euclidean derivative. We start by introducing the concept of the vertical subspace.

The vertical subspace

Definition 2.1.1. Let P be a fibre bundle over M, then $V_pP = \ker \pi_* \subset T_pP$ is called the *vertical subspace* of T_pP. If a vector field $X \in \mathscr{X}(P)$ satisfies $X|_p \in V_pP$ for all $p \in P$, the vector field X is called *vertical.* By writing $VP = \{V_pP : p \in P\}$, we can denote this with $X \in VP$.

Proposition 2.1.2. Let $x = \pi(p)$. Then the vertical subspace V_pP equals $T_p\Big(\pi^{-1}(x)\Big)$.

Proof. Let $f : M \to \mathbb{R}$ and $X|_p \in T_p\Big(\pi^{-1}(x)\Big)$. Since $f \circ \pi$ restricted to $\pi^{-1}(x)$ is constant, we find $\pi_*X|_pf = X|_p(f \circ \pi) = 0$, so $X|_p \in V_pP$ and we find that $T_p\Big(\pi^{-1}(x)\Big)$ can be embedded in V_pP. Since π is surjective, we find that π_* is a linear surjection, so from the Dimension Theorem for vector spaces, we find that $\dim T_xM = \dim \mathrm{im}\ \pi_* = \dim T_pP - \dim \ker \pi_*$. From Corollary 1.1.9, we know that $\dim M = \dim P - \dim G$, thus by $\dim T_pP = \dim M$, we find that $\dim \ker \pi_* = \dim G$. Whence $\pi^{-1}(x) \simeq G$ implies $\dim T_p\Big(\pi^{-1}(x)\Big) = \dim G$, we find for dimensional reasons that the embedding is also surjective. ☐

Fundamental vector fields

Let $P(M, G)$ be a principal bundle. For any $A \in \mathfrak{g}$ we can define a curve $t \mapsto \exp(tA)$ in G, which for all $p \in P$ gives rise to a curve $c_p(t) = R_{\exp(tA)}p = p\exp(tA)$. Notice that c_p depends smoothly on p. Since $c_p(0) = p$, we see that $c_p'(0)$ is an element of T_pP, which we denote with $\sigma(A)\big|_p$.

$$\sigma(A)|_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{tA} \tag{2.1.1}$$

Then $\sigma(A)$ is a vector field on P, called the *fundamental vector field associated to* A, so we have a map $\sigma : \mathfrak{g} \to \mathscr{X}(P)$. We have $\pi \circ c_p(t) = \pi(p)$ for all $t \in \mathbb{R}$, whence we find $\pi_*\, \sigma(A)\big|_p = \pi_* c_p'(0) = (\pi \circ c_p)'(0) = 0$, hence $\sigma(A)\big|_p \in V_pP$ for all $p \in P$.

Lemma 2.1.3. σ satisfies the following identities:

1. $R_{g*}\sigma(A) = \sigma(\mathrm{ad}_{g^{-1}}A)$
2. $g_{i*}\sigma(A)|_p = g_i(p)A$.

Proof. The proof goes by direct calculation. The first identity follows from

$$\begin{aligned}
(R_{g*}\sigma(A)|_p &= R_{g*}\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{tA}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g\left(pe^{tA}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{tA}g \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pgg^{-1}e^{tA}g \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pge^{t\left(\mathrm{ad}_{g^{-1}}A\right)} \\
&= \sigma(\mathrm{ad}_{g^{-1}}A)|_{pg},
\end{aligned}$$

the second identity by

$$\begin{aligned}
g_{i*}\sigma(A)|_p &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g_i\left(pe^{tA}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g_i(p)e^{tA} \\
&= g_i(p)A.
\end{aligned}$$

□

Next proposition can be found in [30] as Proposition 8.1.

Proposition 2.1.4. Let $P(M,G)$ be a principal bundle. Then we can define a map $\sigma : \mathfrak{g} \to \mathscr{X}(P)$, such that

1. We have $\sigma([A,B]) = [\sigma(A),\sigma(B)]$ for all $A,B \in \mathfrak{g}$;
2. The map $A \mapsto \sigma(A)|_p$ is a vector space isomorphism between $\mathfrak{g}$ and V_pP.

This proposition implies that V_pP is an Lie algebra isomorphic to $\mathfrak{g}$. Notice that every vertical vector field X can be written as $X = \sigma(A)$ for a unique $A \in \mathfrak{g}$.

Corollary 2.1.5. $R_{g*}V_pP = V_{pg}P$.

Proof. Let $X|_p \in V_pP$. Let $A \in \mathfrak{g}$ such that $\sigma(A)|_p = X|_p$. Then $R_{g*}X|_p = R_{g*}\sigma(A)|_p = \sigma(\mathrm{ad}_{g^{-1}}A)|_{pg}$, which is vertical. Conversely, let $X|_{pg} \in V_{pg}P$. Then there is an $B \in \mathfrak{g}$ such that $X|_{pg} = \sigma(B)|_{pg}$. Let $A \in \mathfrak{g}$ such that $B = \mathrm{ad}_{g^{-1}}A$. Then $X|_{pg} = \sigma(\mathrm{ad}_{g^{-1}}A)|_{pg} = R_{g*}\sigma(A)|_p$, and since $\sigma(A)|_p \in V_pP$, we find that $X|_{pg}$ is in $R_{g*}V_pP$. □

The Eheresmann connection

Definition 2.1.6. An *Ehresmann connection* on a principal bundle $P(M,G)$ is defined as a choice of a *horizontal subspace* H_pP which is a subspace of T_pP satisfying

1. $T_pP = V_pP \oplus H_pP$;
2. Any smooth vector field X on P can be written as the sum of two smooth vector fields $X = X^H + X^V$ such that $X^H|_p \in H_pP$ and $X^V|_p \in V_pP$ for all $p \in P$;

3. $H_{pg}P = R_{g*}H_pP$ for all $g \in G$.

The second condition guarantees that the choice of H_pP is a smooth one. The third condition is called the *equivariance condition*, and tells us, given the horizontal subspace H_pP at p, what the horizontal subspace H_qP at q in the same fiber as p is. Analogous to the definition of vertical vector fields, we define *horizontal* vector fields to be the fields $X \in \mathscr{X}(P)$ such that $X|_p \in H_pP$ for all $p \in P$. If we introduce the notation $HP = \{H_pP : p \in P\}$, we can denote this with $X \in HP$. With this terminology the second condition can be translated to 'Every vector field X can be split in a smooth way into a horizontal vector field X^H and a vertical vector field X^V.'

Example 2.1.7. Given a G-invariant Riemannian metric on P, we can define a connection by defining $H_pP = V_pP^{\perp}$.

Proposition 2.1.8. Let $x = \pi(p)$. Then $\pi_* : H_pP \to T_xM$ is an isomorphism.

Proof. This follows directly from the facts that V_pP is defined as the kernel of π_* and the decomposition $T_pP = H_pP \oplus V_pP$. ☐

Sometimes it is more convenient to describe the horizontal spaces in terms of projection operators.

Definition 2.1.9. Let HP be a connection. Then we define the projection operator H_p onto the space of horizontal vector fields by

$$H_p(X|_p) = \begin{cases} X|_p, & X|_p \in H_pP; \\ 0, & X|_p \in V_pP. \end{cases} \tag{2.1.2}$$

If no confusion is possible, we drop the subscript p and write H. We have a corresponding projection operator V onto the space of vertical vector fields along the space of horizontal vector fields given by $V = I - H$.

Proposition 2.1.10. Let P_p be a projection operator acting on T_pP. Then im P_p is an Ehresmann connection if and only if

1. $\ker P_p = V_pP$;

2. P_p depends smoothly on p;

3. $P_{pg}R_{g*} = R_{g*}P_p$.

Conversely, given horizontal spaces H_pP, the operator H satisfies above conditions.

Proof. It is clear that first two conditions for P_p are equivalent with the first two conditions of Definition 2.1.6. Assume the third condition for of the proposition for P_p holds and let $X|_p \in \text{im } P_p$. Then $R_{g*}X|_p = R_{g*}P_pX|_p = P_{pg}R_{g*}X|_p$, so $R_{g*}X|_p \in \text{im } P_{pg}$. Let $X|_p \in \ker P$, then $P_{pg}R_{g*}X|_p = R_{g*}P_pX|_p = 0$, so $R_{g*}X|_p \in \ker P_p$. So we find that $R_{g*}\text{im } P_p = \text{im } P_{pg}$, which corresponds with the third condition in the definition of the Ehresmann connection.

Conversely, from Corollary 2.1.5, we find that not not only the horizontal subspace, but also the vertical subspace satisfies an equivariance condition, whence

$$\begin{aligned} H_{pg}R_{g*}X|_p &= H_{pg}R_{g*}X|_p^H + H_{pg}R_{g*}X|_p^V \\ &= H_{pg}R_{g*}X|_p^H && [R_{g*}X|_p^V \in V_{pg}P] \\ &= R_{g*}X|_p^H && [R_{g*}X|_p^H \in H_{pg}P] \\ &= R_{g*}H_pX|_p. \end{aligned}$$

☐

Proposition 2.1.11. The space of vertical vector fields form an ideal in the Lie algebra of vector fields: Let X be an arbitrary vector field and V a vertical vector field on P. Then $[X, V]$ is vertical.

Proof. By Proposition 2.1.4 we know that $[X, V]$ is vertical if X is vertical, so assume that X is horizontal. Then $X|_{pg} = R_{g*}X|_p$, so we find $\pi_* X|_{pg} = \pi_* R_{g*} X|_p = \pi_* X|_p$, whence $\pi_* X$ is a well-defined vector field. Then $\pi_*[X, V] = [\pi_* X, \pi_* V]$ (Proposition 4.16 of [27]), and since $\pi_* V = 0$, we find $\pi_*[X, V] = 0$. Thus $[X, V]$ is vertical. ☐

The connection one-form

Definition 2.1.12. A *connection one-form* $\omega \in \Omega^1(P, \mathfrak{g})$ is a smooth Lie algebra valued one-form satisfying

1. $\omega \circ \sigma = \mathrm{Id}_{\mathfrak{g}}$,
2. $R_g^*\omega = \mathrm{ad}_{g^{-1}}\omega$ for all $g \in G$.

Remark. If we write $X_A = \sigma(A)$ for $A \in \mathfrak{g}$, condition 1 is equivalent with $\omega(X_A) = A$. In other words, ω acts as a Maurer-Cartan form on V_pP. Condition 2 means that for $X \in \mathscr{X}(P)$ we have $R_g^*\omega(X) = \omega(R_{g*}X) = g^{-1}\omega(X)g$.

Lemma 2.1.13. For all connection one-forms ω we have that $\sigma \circ \omega$ is a projection operator onto the vertical subspace.

Proof. Let ω be a connection one-form. Since $(\sigma \circ \omega)^2 = \sigma \circ (\omega \circ \sigma) \circ \omega = \sigma \circ \mathrm{Id} \circ \omega$, we see that $\sigma \circ \omega$ is a projection operator. Furthermore, ω is by definition a surjection, while σ is an isomorphism between $\mathfrak{g}$ and V_pP, so $\mathrm{im}(\sigma \circ \omega) = V_pP$. ☐

Proposition 2.1.14. The definitions of the connection one-form and the Ehresmann connection are equivalent; we have a map between the collection of connection one-forms and the collection of Ehresmann connections given by $\omega \mapsto \mathrm{Id} - \sigma \circ \omega$ with inverse $H \mapsto \sigma^{-1} \circ (\mathrm{Id} - H)$.

Proof. It is trivial that both maps are inverses of each other. Notice that σ^{-1} only acts on vertical vector fields, but whence $\mathrm{Id} - H = V$, the inverse is well-defined.

Since $\sigma \circ \omega$ is a projection, it directly follows that $H = \mathrm{Id} - \sigma \circ \omega$ is also a projection operator. Whence,

$$\begin{aligned}
\ker H_p &= \{X|_p : (\mathrm{Id} - \sigma \circ \omega)X|_p = 0\} \\
&= \{X|_p : X|_p = \sigma \circ \omega(X|_p)\} \\
&= \mathrm{im}(\sigma \circ \omega) \\
&= V_pP,
\end{aligned}$$

we find that H is a projection operator along the vertical subspace. Since H is a composition of smooth operators, it is smooth. Finally, H is equivariant, since for $X|_p \in T_pP$, we have

$$\begin{aligned}
R_{g*}H(X) &= R_{g*}(\mathrm{Id} - \sigma \circ \omega)X \\
&= R_{g*}X - R_{g*}\sigma \circ \omega(X) \\
&= R_{g*}X - \sigma\Big(\mathrm{ad}_{g^{-1}}\omega(X)\Big) && \text{[By Lemma 2.1.3]} \\
&= R_{g*}X - \sigma\Big(R_g^*\omega(X)\Big) && \text{[By Definition 2.1.12]} \\
&= \mathrm{Id} \circ R_{g*}X - \sigma \circ \omega \circ R_{g*}X \\
&= HR_{g*}(X).
\end{aligned}$$

Conversely, let H be the projection operator corresponding with the horizontal subspace H_pP and let $\omega = \sigma^{-1} \circ (\mathrm{Id} - H)$. Let $X = \sigma(A)$. Since we have $H \circ \sigma = 0$, we find $\omega(X) = \sigma^{-1}(X)$, hence $\omega \circ \sigma(A) = A$. Lastly, we have to check that $R_g^*\omega(X) = \mathrm{ad}_{g^{-1}}\omega(X)$ for all $X|_p \in T_pP$. For horizontal X the relation holds, since $\mathrm{ad}_{g^{-1}}\omega(X) = 0$, for ω kills horizontal vector fields, and also by definition of H we have

$$\begin{aligned} R_g^*\omega(X) &= R_g^*\omega(HX) \\ &= \omega(R_{g*}HX) \\ &= \omega(HR_{g*}X) \\ &= 0, \end{aligned}$$

so both sides are clearly zero. For vertical X, take an $A \in \mathfrak{g}$ such that $X = \sigma(A)$. Then

$$\begin{aligned} R_g^*\omega(X) &= \omega \circ R_{g*}X \\ &= \omega \circ R_{g*}\sigma(A) \\ &= \omega \circ \sigma(\mathrm{ad}_{g^{-1}}A) && \text{[By Lemma 2.1.3]} \\ &= \mathrm{ad}_{g^{-1}}A \\ &= \mathrm{ad}_{g^{-1}}\omega(X), \end{aligned}$$

hence we see that the second property also holds for vertical X, which finishes the proof. □

Corollary 2.1.15. Let HP be the Ehresmann connection corresponding to the connection one-form ω. Then $\omega(X) = 0$ if X is horizontal.

Proof. Given ω, we have the horizontal projector $H = \mathrm{Id} - \sigma \circ \omega$. Then for any vector field X, we have $\omega(HX) = \omega\Big(\mathrm{Id}(X) - \sigma \circ \omega(X)\Big) = \omega(X) - \mathrm{Id} \circ \omega(X) = 0$. □

Proposition 2.1.16. Any principal bundle $P \to M$ has a connection one-form ω.

Proof. For any coordinate neighborhood U_i let $\omega_i = g_i^*\theta$, then ω_i is a connection one-form on the trivial bundle $\pi^{-1}(U_i) \to U_i$. We will postpone the proof that ω_i is indeed a connection one-form until section 2.6, where it is given in Proposition 2.6.3. Let $\{\lambda_i\}$ be a partition of unity subordinate the covering $\{U_i\}$ of M, then we define $\omega = \sum_i \lambda_i\omega_i$. It is easy to see that ω satisfies $R_g^*\omega = \mathrm{ad}_{g^{-1}} \circ \omega$ and $\omega \circ \sigma = \mathrm{Id}$, since the ω_i satisfy the same identities. Hence we see that ω is a connection on $\pi^{-1}(U_i)$ for all coordinate neighborhoods U_i. But since this is valid for all coordinate neighborhoods, it follows that ω is a connection on whole P. □

The gauge potential

For mathematicians P might be more interesting than M, since latter space is a building stone of the more complex former space. For physicists however, the space M is more interesting, since it forms the spacetime on which the gauge theories are defined. Therefore we are interested in describing connections in terms of functions on M.

Definition 2.1.17. Let s_i be the smooth section associated to the local trivializations ψ_i. The local connection form or *gauge potential* is defined as

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

Proposition 2.1.18. Let U_i be a coordinate neighborhood. Then the restriction of the connection one-form ω to $\pi^{-1}(U_i)$ agrees with ω_i defined as

$$\omega_i|_p = \mathrm{ad}_{g_i(p)^{-1}} \circ \pi^*\mathcal{A}_i|_{\pi(p)} + g_i^*\theta|_{g_i(p)}, \tag{2.1.3}$$

with θ the Maurer-Cartan form, and $x = \pi(p)$.

For the proof of the proposition, we need a lemma.

Lemma 2.1.19. Let $s : U \to P$ be a local section and $p \in \text{im } s$. Then $(s \circ \pi)_* : T_pP \to T_pP$ is a projection operator such that $T_pP = \text{im}(s \circ \pi)_* \oplus V_pP$.

Proof. We first remark that for every projection operator $P : X \to X$, with X a vector space, we have $X = \text{im } P \oplus \ker P$. Then since $\pi \circ s = \text{Id}$, we have $\pi_* \circ s_* = \text{Id}$, so s_* is injective, for it has a left inverse. Whence

$$(s \circ \pi)_*^2 = s_* \circ (\pi_* \circ s_*) \circ \pi_* = (s \circ \pi)_*,$$

we find that $(s \circ \pi)_*$ is a projection operator. From the injectivity of s_* we find that $\ker(s \circ \pi)_* = \ker \pi_* = V_pP$, which proves the lemma. □

Proof of Proposition 2.1.18. Now, let $x \in U_i$ and let $p = s_i(x)$, so by Proposition 1.3.3 we have $g_i \circ s_i(x) = e$. By Lemma 2.1.19, we find for all $X \in \mathscr{X}(P)$ that $X = s_{i*} \circ \pi_*(X) + \sigma(A)$ for a unique vertical vector $\sigma(A)$. Then we find

$$\begin{aligned}
\omega_i|_p(X|_p) &= \text{ad}_{g_i(p)^{-1}} \circ \pi^* \mathcal{A}_i|_{\pi(p)}(X|_p) + g_i^*\theta|_{g_i(p)}(X|_p) \\
&= \text{ad}_{g_i(p)^{-1}} \circ \pi^* s_i^* \omega|_{s_i \circ \pi(p)}(X|_p) + \theta|_{g_i(p)} g_{i*}(X|_p) \\
&= \text{ad}_e \circ \pi^* s_i^* \omega|_p(X|_p) + \theta|_{g_i(p)} g_{i*}(s_{i*} \circ \pi_*(X|_p) + \sigma(A)|_p) \\
&= \omega|_p\Big(s_{i*} \circ \pi_*(X|_p)\Big) + \theta|_{g_i(p)} g_{i*}\sigma(A)|_p && [\text{Since } (g_i \circ s_i)_* = 0] \\
&= \omega|_p\Big(s_{i*} \circ \pi_*(X|_p)\Big) + \theta|_{g_i(p)} g_i(p) A && [\text{By Lemma 2.1.3}] \\
&= \omega|_p\Big(s_{i*} \circ \pi_*(X|_p)\Big) + L_{g_i(p)^{-1}} g_i(p) A && [\text{By Definition of } \theta] \\
&= \omega|_p\Big(s_{i*} \circ \pi_*(X|_p)\Big) + A \\
&= \omega|_p\Big(s_{i*} \circ \pi_*(X|_p)\Big) + \omega|_p \sigma(A)|_p && [\omega \circ \sigma = \text{Id}_{\mathfrak{g}}] \\
&= \omega|_p(X|_p).
\end{aligned}$$

Hence we find that in the point $p = s_i(x)$ we have $\omega_i|_p = \omega|_p$. In order to prove that the identity holds for all $p \in \pi^{-1}(U_i)$, we show that ω_i transforms in the same way as ω under the right action of G. First we notice that ad_a for $a \in G$ is a pullback, which commutes with pushforwards. Then we find

$$\begin{aligned}
R_g^*\omega_i|_{pg} &= \text{ad}_{g_i(pg)^{-1}} R_g^* \pi^* s_i^* \omega|_{pg} + R_g^* g_i^* \theta|_{g_i(p)g} \\
&= \text{ad}_{(g_i(p)g)^{-1}} R_g^* \pi^* s_i^* \omega|_{pg} + g_i^* R_g^* \theta|_{g_i(pg)} && [\text{Equivariance of } g_i] \\
&= \text{ad}_{g^{-1} g_i(p)^{-1}} \pi^* s_i^* \omega|_p + g_i^* R_g^* \theta|_{g_i(pg)} && [\pi \circ R_g = \pi] \\
&= \text{ad}_{g^{-1}} \text{ad}_{g_i(p)^{-1}} \pi^* s_i^* \omega|_p + g_i^* \text{ad}_{g^{-1}} \theta|_{g_i(p)} && [\text{By (1.4.5)}] \\
&= \text{ad}_{g^{-1}}\Big(\text{ad}_{g_i(p)^{-1}} \pi^* s_i^* \omega|_p + g_i^* \theta|_{g_i(p)}\Big) \\
&= \text{ad}_{g^{-1}} \circ \omega_i|_p.
\end{aligned}$$

□

Corollary 2.1.20. Let $\mathcal{A}_i \in \Omega^1(U_i, \mathfrak{g})$. Then the $\mathcal{A}_i$ define a connection one-form ω by (2.4.7) if and only if on overlaps U_{ij} the $\mathcal{A}_i$ satisfy

$$\begin{aligned}
\mathcal{A}_i|_x &= \text{ad}_{g_{ij}(x)} \mathcal{A}_j|_x + g_{ji}^* \theta|_{g_{ji}(x)} \\
&= g_{ij}(x) \mathcal{A}_j|_x g_{ij}(x)^{-1} + g_{ji}(x)^{-1} \text{d} g_{ji}|_x.
\end{aligned} \tag{2.1.4}$$

Proof. Let ω be a connection one-form, then we have $\omega_i = \omega_j$ on U_{ij}, so we find

$$\begin{aligned}
\mathcal{A}_i|_x &= s_i^*\omega_i|_{s_i(x)}\\
&= s_i^*\omega_j|_{s_i(x)}\\
&= s_i^*\left(\mathrm{ad}_{g_j(p)^{-1}}\circ\pi^*\mathcal{A}_j|_{\pi\circ s_i(x)} + g_j^*\theta|_{g_j\circ s_i(x)}\right)\\
&= \mathrm{ad}_{g_j(p)^{-1}} s_i^*\circ\pi^*\mathcal{A}_j|_x + s_i^*g_j^*\theta|_{g_j\circ s_i(x)}\\
&= \mathrm{ad}_{g_j(s_i(x))^{-1}}(\pi\circ s_i)^*\mathcal{A}_j|_x + (g_j\circ s_i)^*\theta|_{g_j\circ s_i(x)}\\
&= \mathrm{ad}_{g_{ji}(x)^{-1}}\mathcal{A}_j|_x + g_{ji}^*\theta|_{g_{ji}(x)} && \text{[By (1.3.6)]}\\
&= \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j|_x + g_{ji}^*\theta|_{g_{ji}(x)}.
\end{aligned}$$

Conversely, let $\mathcal{A}_i$ satisfies (2.1.4) and define ω_i by (2.1.3). We shall show that the ω_i agree on overlaps, and so define a global one-form, which satisfies the conditions of a connection one-form. By Proposition 1.3.3 we have $g_i\circ s_i(x) = e$, so $(g_i\circ s_i)^* = 0$. Furthermore, we have $s_i^*\circ\pi^* = (\pi\circ s_i)^* = \mathrm{Id}^* = \mathrm{Id}$, whence on overlaps U_{ij}

$$\begin{aligned}
s_i^*\omega|_{s_i(x)} &= \mathrm{ad}_{(g_i\circ s_i(x))^{-1}}\circ s_i^*\pi^*\mathcal{A}_i|_{\pi\circ s_i(x)} + s_i^*g_i^*\theta|_{g_i\circ s_i(x)}\\
&= \mathrm{ad}_e\circ\mathcal{A}_i|_x + (g_i\circ s_i)^*\theta|_{g_i\circ s_i(x)}\\
&= \mathcal{A}_i|_x\\
&= \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j|_x + g_{ji}^*\theta|_{g_{ji}(x)} && \text{[By (2.1.4)]}\\
&= \mathrm{ad}_{g_j(s_i(x))^{-1}}(\pi\circ s_i)^*\mathcal{A}_j|_x + (g_j\circ s_i)^*\theta|_{g_j\circ s_i(x)}\\
&= \mathrm{ad}_{g_j(p)^{-1}} s_i^*\circ\pi^*\mathcal{A}_j|_x + (s_i^*\circ g_j^*\theta)|_{g_j\circ s_i(x)}\\
&= s_i^*\left(\mathrm{ad}_{g_j(p)^{-1}}\circ\pi^*\mathcal{A}_j|_{\pi\circ s_i(x)} + g_j^*\theta|_{g_j\circ s_i(x)}\right)\\
&= s_i^*\omega_j|_{s_i(x)}.
\end{aligned}$$

Since s_i and thus s_i^* is injective, we find $\omega_i = \omega_j$ on U_{ij}, so the ω_i define a global one-form ω. We check that this is a connection one-form. Firstly, we check if $\omega_i\circ\sigma = \mathrm{Id}_{\mathfrak{g}}$.

$$\begin{aligned}
\omega_i|_p\circ\sigma(A)|_p &= \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)}\sigma(A)|_p + g_i^*\theta|_{g_i(p)}\sigma(A)|_p\\
&= \mathrm{ad}_{g_i(p)^{-1}}\mathcal{A}_i|_{\pi(p)}\pi_*\sigma(A)|_p + \theta|_{g_i(p)}g_{i*}\sigma(A)|_p\\
&= \theta|_{g_i(p)}g_{i*}\sigma(A)|_p && [\sigma(A)\in V_pP = \ker\pi_*]\\
&= \theta|_{g_i(p)}g_i(p)A && \text{[Lemma 2.1.3]}\\
&= L_{g_i(p)^{-1}}g_i(p)A\\
&= A.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
R_g^*\omega_i|_{pg} &= \mathrm{ad}_{g_i(pg^{-1})}\circ R_g^*\pi^*\mathcal{A}_i|_{\pi(pg)} + R_g^*g_i^*\theta|_{g_i(pg)}\\
&= \mathrm{ad}_{g_i(pg)}\circ R_g^*\pi^*\mathcal{A}_i|_{\pi(pg)} + R_g^*g_i^*\theta|_{g_i(pg)}.
\end{aligned}$$

First we remark that by Proposition 1.1.7 we have $R_g^*\pi^* = (\pi\circ R_g)^* = \pi^*$ and $g_i(pg) = g_i(p)g$, or equivalently $g_i\circ R_g = R_g\circ g_i$ (we need both formulations), so we find

$$R_g^*\omega_i|_{pg} = \mathrm{ad}_{(g_i(p)g)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(pg)} + g_i^*R_g^*\theta|_{g_i(p)g}.$$

By (1.4.5), we find $R_g^*\theta|_{g_i(p)g} = \mathrm{ad}_{g^{-1}}\theta|_{g_i(p)}$, whence

$$\begin{aligned}
R_g^*\omega_i|_{pg} &= \mathrm{ad}_{g^{-1}g_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + \mathrm{ad}_{g^{-1}}g_i^*\theta|_{g_i(p)}\\
&= \mathrm{ad}_{g^{-1}}\left(\mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + g_i^*\theta|_{g_i(p)}\right)\\
&= \mathrm{ad}_{g^{-1}}\omega_i|_p,
\end{aligned}$$

which proves that ω is a connection one-form. ∎

Last proposition shows that we can define connections by (2.1.4) without referring to the space P, we only need a covering $\{U_i\}$ of M on which we define the $\mathcal{A}_i$, so the U_i do not even have to be coordinate neighborhoods of a principal bundle. This allows us to define connections on other bundles than principal bundles.

Definition 2.1.21. Let E be a vector bundle over M with structure group G. If $\mathcal{A}_i$ is a collection on $\mathfrak{g}$-valued one-forms on U_i satisfying (2.1.4), then we call the family $\{\mathcal{A}_i\}$ a connection on E.

On a principal bundle P, to summarize, we have four equivalent descriptions of a connection on P:

1. A horizontal distribution $HP \subset TP$ satisfying $R_{g*}H_pP = H_{pg}P$;
2. A projection operator P acting on TP satisfying $P_{pg}R_{g*} = R_{g*}P_p$;
3. A one-form $\omega \in \Omega^1(P, \mathfrak{g})$ satisfying $\omega \circ \sigma = \mathrm{Id}_{\mathfrak{g}}$ and $R_g^*\omega = \mathrm{ad}_{g^{-1}} \circ \omega$;
4. A family of one-forms $\mathcal{A}_i \in \Omega^1(U_i, \mathfrak{g})$ satisfying $\mathcal{A}_i|_x = \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j|_x + g_{ji}^*\theta|_{g^{ji}(x)}$.

The space of connections

Definition 2.1.22. We define the following two subsets of $\Omega^1(P, \mathfrak{g})$.

$$\begin{aligned} \mathscr{A} &= \{\omega \in \Omega^1(P, \mathfrak{g}) : R_g^*\omega = \mathrm{ad}_{g^{-1}} \circ \omega,\ \omega \circ \sigma = \mathrm{Id}\} \\ \mathscr{V} &= \{\tau \in \Omega^1(P, \mathfrak{g}) : R_g^*\tau = \mathrm{ad}_{g^{-1}} \circ \tau,\ \tau \circ \sigma = 0\} \end{aligned}$$

The space $\mathscr{A}$ is exactly the space of all connection one-forms and is therefore called the *space of connections*. It is not immediately clear that $\mathscr{A}$ is non-empty, but we shall prove this below.

Note that $\mathscr{A}$ is not a vector space, since the sum of two connection one-forms ω and ω' do not satisfy the second condition anymore: $(\omega + \omega') \circ \sigma = 2\ \mathrm{Id}$. However, $\mathscr{A}$ is a so called affine space, which we will introduce immediately.

Definition 2.1.23. Let A be a set and V a vector space with an action $V \times A \to A$ denoted with $(v, a) \mapsto v + a$. Then A is called an *affine space over* V if the action satisfies the following properties:

1. $0 + a = a$ for all $a \in A$;
2. $v + (w + a) = (v + w) + a$ for all $v, w \in V$ and $a \in A$;
3. For all $a \in A$ the map $V \to A$ given by $v \mapsto v + a$ is a bijection.

Lemma 2.1.24. If A is a subset of some vector space W such that $V = \{a' - a : a, a' \in A\}$ is a subspace of W, then A is an affine space over V.

Proof. Trivial. □

Since $\mathscr{V} = \{\omega' - \omega : \omega, \omega' \in \mathscr{A}\}$, it is clear that $\mathscr{A}$ is an affine space modelled on $\mathscr{V}$. The space of connection $\mathscr{A}$ turns out to be infinite-dimensional, but later on, we shall consider quotients of subspaces of $\mathscr{A}$, which are finite-dimensional under certain conditions. It will turn out that these quotients, called *moduli spaces*, are geometric invariants, which makes them interesting.

We can also define $\mathscr{A}$ in terms of an affine space modelled on the space of ad P-valued one-forms.

Proposition 2.1.25. Identify $\mathscr{A}$ with the space of families of gauge fields $\mathcal{A}_i \in \Omega^1(U_i, \mathfrak{g})$ satisfying (2.1.4) by Corollary 2.1.20. Then $\mathscr{A}$ is an affine space over $\Omega^1(M, \mathrm{ad}\ P)$.

Proof. Let $\{\mathcal{A}_i\}$ be a family of gauge fields in $\mathscr{A}$ and let $\tau \in \Omega^1(M, \mathrm{ad}\ P)$. By Corollary 1.5.11 we find that there are unique forms $\tau_i \in \Omega^1(U_i, \mathfrak{g})$ such that

$$\tau_i|_x = \mathrm{ad}_{g_{ij}(x)} \circ \tau_j|_x, \tag{2.1.5}$$

Then we define the action $\Omega^1(M, \mathrm{ad}P) \times \mathscr{A} \to \mathscr{A}$ by

$$\Big(\tau, \{\mathcal{A}_i\}\Big) \mapsto \{\tau_i + \mathcal{A}_i\}.$$

From (2.1.4) and (2.1.5) it follows that $\tau_i + \mathcal{A}_i$ satisfies (2.1.4), so the action is well-defined. Then it is easy to see that this action satisfies the axioms for an affine space. $\square$

2.2 The curvature of a connection

Horizontal forms and the exterior derivative

In Paragraph 2.1 we introduced the projection operator $H_p : T_pP \to H_pP$, and showed that it is equivariant: $H \circ R_{g*} = R_{g*} \circ H$. We use the notation $H^\vee$ for the dual of H, so if $\beta \in \Omega^k(P, F)$, we have $(H^\vee\beta)(X_1, \ldots, X_k) = \beta(HX_1, \ldots, HX_k)$.

Definition 2.2.1. A form $\alpha \in \Omega^k(P, F)$ with F a vector space is called *horizontal* if

$$H^\vee\alpha = \alpha, \tag{2.2.1}$$

or equivalently $\alpha(HX_1, \ldots, HX_k) = \alpha(X_1, \ldots, X_k)$, where $X_i \in T_pP$.

If α is horizontal, we do not have necessarily that $\mathrm{d}\alpha$ is also horizontal. In order to fix this we make the following definition:

Definition 2.2.2. Let $\alpha \in \Omega^k(P, F)$, then we define the *exterior covariant derivative* by $\mathrm{d}_\omega\alpha = H^\vee\mathrm{d}\alpha$, or equivalently by

$$\mathrm{d}_\omega\alpha(X_1, \ldots, X_{k+1}) = \mathrm{d}\alpha(X_1^H, \ldots, X_{k+1}^H), \tag{2.2.2}$$

where $X_1, \ldots, X_{k+1} \in T_pP$.

We use the subscript ω, since the derivative depends on H and thus on ω.

The curvature 2-form

A distribution $\mathscr{D}$ of $\mathscr{X}(M)$ is called *integrable* if for $X, Y \in \mathscr{D}$, we have $[X, Y] \in \mathscr{D}$. It turns out that integrable distributions have some nice properties. For instance, if the subset of holomorphic vector fields of an almost complex manifold is integrable, the manifold is complex. For principal fibre bundles, the question wether the horizontal subspace is integrable is also important, since connections with an integrable horizontal subspace are so-called *flat*, a notion which will be discussed later on. We first make the following definition:

Definition 2.2.3. Let $\omega \in \Omega^1(P, \mathfrak{g})$ be the connection one-form for a connection $HP \subset TP$. The 2-form $\Omega = \mathrm{d}_\omega\omega \in \Omega^2(P, \mathfrak{g})$ is called the *curvature (2-form)* of the connection.

Proposition 2.2.4. The curvature satisfies the following identities:

$$H^\vee\Omega = \Omega; \tag{2.2.3}$$
$$R_g^*\Omega = \mathrm{ad}_{g^{-1}}\Omega; \tag{2.2.4}$$
$$\Omega = \mathrm{d}\omega + \frac{1}{2}[\omega,\omega]; \tag{2.2.5}$$
$$\mathrm{d}_\omega\Omega = 0. \tag{2.2.6}$$

Equation (2.2.5) is called the *Cartan structure equation*. Equation (2.2.6) is called the *Bianchi identity*.

Proof. This first identity follows directly from the definition of Ω. For the second, let $X, Y \in \mathscr{X}(P)$. Then

$$\begin{aligned}
R_g^*\Omega(X,Y) &= \Omega(R_{g*}X, R_{g*}Y) \\
&= H^\vee\mathrm{d}\omega(R_{g*}X, R_{g*}Y) \\
&= \mathrm{d}\omega(HR_{g*}X, HR_{g*}Y) \\
&= \mathrm{d}\omega(R_{g*}HX, R_{g*}HY) && \text{[By Proposition 2.1.10]} \\
&= R_g^*\mathrm{d}\omega(HX, HY) \\
&= \mathrm{d}R_g^*\omega(HX, HY) \\
&= \mathrm{ad}_{g^{-1}}\mathrm{d}\omega(HX, HY) && \text{[By Definition 2.1.12]} \\
&= \mathrm{ad}_{g^{-1}}\Omega(X,Y).
\end{aligned}$$

By (A.1.10) we see (2.2.5) is equivalent with

$$\mathrm{d}\omega(HX, HY) = \mathrm{d}\omega(X,Y) + [\omega(X), \omega(Y)] \tag{2.2.7}$$

for all vector fields X, Y on P. We show that this equation holds by considering three different cases.

1. X and Y are horizontal. In this case (2.2.7) holds, since $\omega(X) = \omega(Y) = 0$ and $HX = X$, $HY = Y$.

2. X and Y are vertical. In this case the left-hand side of (2.2.7) is zero. Let $A, B \in \mathfrak{g}$ such that $\sigma(A) = X$ and $\sigma(B) = Y$, then we find

$$\begin{aligned}
\mathrm{d}\omega(X,Y) &= X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) && \text{[By Proposition } A.2.3] \\
&= X(\omega\circ\sigma(B)) - Y(\omega\circ\sigma(A)) - \omega([\sigma(A),\sigma(B)]) \\
&= X(\omega\circ\sigma(B)) - Y(\omega\circ\sigma(A)) - \omega\circ\sigma([A,B]) && \text{[By Proposition 2.1.4]} \\
&= XB - YA - [A,B] && [\omega\circ\sigma = \mathrm{Id}_\mathfrak{g}] \\
&= -[A,B] && [A, B \text{ are constant}] \\
&= -[\omega\circ\sigma(A), \omega\circ\sigma(B)] && [\omega\circ\sigma = \mathrm{Id}_\mathfrak{g}] \\
&= -[\omega(X), \omega(Y)],
\end{aligned}$$

which proves that the right-hand side of (2.2.7) is also zero.

3. X is horizontal and Y is vertical. Write $Y = \sigma(B)$ for some $B \in \mathfrak{g}$. Then we find (2.2.7) reduces to

$$\begin{aligned}
0 &= \mathrm{d}\omega(X, \sigma(B)) \\
&= X(\omega\circ\sigma(B)) - \sigma(B)(\omega(X)) - \omega([X,\sigma(B)]) && \text{[By Proposition } A.2.3] \\
&= XB - Y\omega(X) - \omega([X,\sigma(B)]) \\
&= XB - \omega([X,\sigma(B)]) && [X \text{ is horizontal}] \\
&= -\omega([X,\sigma(B)]) && [B \text{ is constant}],
\end{aligned}$$

so we have to show that the commutator of a horizontal and a vertical vector field is horizontal in order to show that (2.2.7) holds. We shall not prove this but refer to Lemma 8.15 of [30] for a proof, which is quite simple if one is familiar with the concept of a Lie derivative.

We prove the Bianchi identity by using the Cartan structure equation.

$$\begin{aligned}
\mathrm{d}_\omega\Omega &= H^\vee \mathrm{d}\left(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]\right) \\
&= H^\vee\left(\frac{1}{2}[\mathrm{d}\omega,\omega] - \frac{1}{2}[\omega,\mathrm{d}\omega]\right) && \text{[By Proposition } A.2.2\text{]} \\
&= H^\vee[\mathrm{d}\omega,\omega] && \text{[By } (A.1.8)\text{]} \\
&= [H^\vee \mathrm{d}\omega, H^\vee\omega] \\
&= 0.
\end{aligned}$$

□

The gauge field-strength

As with gauge potentials, which were introduced in order to describe connections in terms of one-forms on subsets of M, we to describe curvature in terms of two-forms on subsets of M. As in the case of the gauge potential, this can be done by pulling back via the canonical sections s_i.

Definition 2.2.5. We define the *gauge field-strength* $\mathcal{F}_i \in \Omega^2(U_i,\mathfrak{g})$ by $\mathcal{F}_i = s_i^*\Omega$.

Sometimes we shall drop the subscript i and write $\mathcal{F}$ instead of $\mathcal{F}_i$ if it is clear in which patch we work.

Proposition 2.2.6. $\mathcal{F}_i$ satisfies the following identities:

$$\mathcal{F}_i = H^\vee\mathcal{F}_i; \tag{2.2.8}$$

$$\mathcal{F}_i = \mathrm{d}\mathcal{A}_i + \frac{1}{2}[\mathcal{A}_i, A_i]; \tag{2.2.9}$$

$$0 = \mathrm{d}\mathcal{F}_i + [\mathcal{A}_i,\mathcal{F}_i] \tag{2.2.10}$$

Equation (2.2.10) is called the *Bianchi identity* for $\mathcal{F}_i$. Notice that by (A.1.9) equation (2.2.9) is equivalent to

$$\mathcal{F}_i = \mathrm{d}\mathcal{A}_i + \mathcal{A}_i\wedge\mathcal{A}_i. \tag{2.2.11}$$

Proof of Proposition 2.2.6. For the first identity let $X, Y \in \mathscr{X}(U_i)$, then

$$\begin{aligned}
H^\vee\mathcal{F}_i(X,Y) &= H^\vee s_i^*\Omega(X,Y) \\
&= H^\vee s_i^*\mathrm{d}\omega(HX,HY) \\
&= s_i^*\mathrm{d}\omega(HX,HY) \\
&= \mathcal{F}_i(X,Y).
\end{aligned}$$

The second identity follows from the Cartan structure equation (2.2.5).

$$\begin{aligned}
\mathcal{F}_i &= s_i^*\Omega \\
&= s_i^*\left(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]\right) \\
&= \mathrm{d}s_i^*\omega + \frac{1}{2}[s_i^*\omega, s_i^*\omega] \\
&= \mathrm{d}\mathcal{A}_i + \frac{1}{2}[\mathcal{A}_i,\mathcal{A}_i].
\end{aligned}$$

Finally, we find

$$\begin{aligned}
\mathrm{d}\mathcal{F}_i + [\mathcal{A}_i, \mathcal{F}_i] &= s_i^*\mathrm{d}\Omega + [s_i^*\omega, s_i^*\Omega] \\
&= s_i^*\left(\mathrm{d}\Omega + [\omega, \Omega]\right) \\
&= s_i^*\left(\mathrm{d}\left(\mathrm{d}\omega + \frac{1}{2}[\omega,\omega]\right) + \left[\omega, \mathrm{d}\omega + \frac{1}{2}[\omega,\omega]\right]\right) && \text{[By (2.2.5)]} \\
&= s_i^*\left(\frac{1}{2}[\mathrm{d}\omega,\omega] - \frac{1}{2}[\omega,\mathrm{d}\omega] + [\omega,\mathrm{d}\omega] + \left[\omega, \frac{1}{2}[\omega,\omega]\right]\right) && \text{[By Proposition } A.2.2\text{]} \\
&= s_i^*\left(-[\omega,\mathrm{d}\omega] + [\omega,\mathrm{d}\omega] + [\omega, \omega\wedge\omega]\right) && \text{[By (}A.1.8\text{) and (}A.1.9\text{)]} \\
&= s_i^*\left(\omega\wedge\omega\wedge\omega - \omega\wedge\omega\wedge\omega\right) && \text{[By (}A.1.2\text{)]} \\
&= 0,
\end{aligned}$$

which proves the Bianchi identity. □

Remark. Since we can express $\mathcal{F}_i$ fully in terms of $\mathcal{A}_i$, we can also define field-strengths for connections on vector bundles by (2.2.11).

Equation (2.2.9) allows us to describe $\mathcal{F}_i$ in terms of $\mathcal{A}_i$. If we drop the subscript i for the moment, write $\mathcal{A} = \mathcal{A}_\mu \mathrm{d}x^\mu$ and $\mathcal{F} = \frac{1}{2}\mathcal{F}_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$. Then

$$\begin{aligned}
\frac{1}{2}\mathcal{F}_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu &= \mathcal{F} \\
&= \mathrm{d}\mathcal{A} + \frac{1}{2}[\mathcal{A}\,\mathcal{A}] \\
&= \mathrm{d}(\mathcal{A}_\nu \mathrm{d}x^\nu) + \frac{1}{2}[\mathcal{A}_\mu \mathrm{d}x^\mu, \mathcal{A}_\nu \mathrm{d}x^\nu] \\
&= \partial_\mu \mathcal{A}_\nu \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu + \frac{1}{2}[\mathcal{A}_\mu, \mathcal{A}_\nu] x^\mu \wedge x^\nu,
\end{aligned}$$

where the second term in the last line follows from (A.1.6) by writing $\alpha = \beta = \mathcal{A}$ with $\alpha^i = \beta^i = \mathcal{A}^i_\mu T_i$. We can replace the first term with $\frac{1}{2}(\partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu)\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, where we used that $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$ is antisymmetric in μ and ν. With this we obtain (2.2.9) in local coordinates:

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]. \tag{2.2.12}$$

Proposition 2.2.7. $\mathcal{F}_i$ satisfies the following identity

$$\mathcal{F}_i|_x = \mathrm{ad}_{g_{ij}(x)} \mathcal{F}_j|_x. \tag{2.2.13}$$

Proof.

$$
\begin{aligned}
\mathcal{F}_i|_x &= \mathrm{d}\mathcal{A}_i|_x + \frac{1}{2}[\mathcal{A}_i, \mathcal{A}_i]|_x && \text{[By (2.2.9)]}\\
&= \mathrm{d}\Big(\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j|_x + g_{ji}^*\theta|_{g_{ji}(x)}\Big) + \frac{1}{2}\Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j\Big]\Big|_x\\
&\quad + \frac{1}{2}\Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, g_{ji}^*\theta\Big]\Big|_x + \frac{1}{2}\Big[g_{ji}^*\theta, \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j\Big]\Big|_x + \frac{1}{2}\Big[g_{ji}^*\theta, g_{ji}^*\theta\Big]\Big|_{g_{ji}(x)} && \text{[By (2.1.4)]}\\
&= \mathrm{d}\Big(\mathrm{ad}_{g_{ji}(x)^{-1}}\mathcal{A}_j|_x\Big) + \mathrm{d}\Big(g_{ji}^*\theta|_{g_{ji}(x)}\Big) + \frac{1}{2}\Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j\Big]\Big|_x\\
&\quad + \Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, g_{ji}^*\theta\Big]\Big|_x + \frac{1}{2}\Big[g_{ji}^*\theta, g_{ji}^*\theta\Big]\Big|_{g_{ji}(x)} && \text{[By } (A.1.8)\text{]}\\
&= \mathrm{d}\Big(\mathrm{ad}_{g_{ji}(x)^{-1}}\mathcal{A}_j|_x\Big) + \frac{1}{2}\Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j\Big]\Big|_x + \Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, g_{ji}^*\theta\Big]\Big|_x && \text{[By (1.4.6)]}\\
&= \mathrm{ad}_{g_{ji}(x)^{-1}}\mathrm{d}\mathcal{A}_j|_x + \frac{1}{2}\Big[\mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j, \mathrm{ad}_{g_{ij}(x)}\mathcal{A}_j\Big]\Big|_x && \text{[By (1.4.7)]}\\
&= \mathrm{ad}_{g_{ij}(x)}\mathrm{d}\mathcal{A}_j|_x + \mathrm{ad}_{g_{ij}(x)}\frac{1}{2}\Big[\mathcal{A}_j, \mathcal{A}_j\Big]\Big|_x && \text{[By (1.4.4)]}\\
&= \mathrm{ad}_{g_{ij}(x)}\mathcal{F}_j|_x.
\end{aligned}
$$

□

Corollary 2.2.8. The $\mathcal{F}_i$ define a global 2-form $\mathcal{F}_\mathcal{A} \in \Omega^2(M, \mathrm{ad}\ P)$.

Proof. This follows directly from Corollary 1.5.11. □

2.3 The covariant derivative

Since the base space M of a principal bundle $P(M, G)$ is more interesting for physicists than the total space P, it is important to translate the notion of the exterior covariant derivative to a derivative on forms of M, called the covariant derivative. This translation will be done as follows. Firstly, given a vector space F and a representation $\rho : G \to \mathrm{GL}(F)$, we will define the subspace of so-called basic forms on P, and prove that these forms are in bijection with forms on the associated vector bundle $P \times_\rho F$. The most interesting case will be $F = \mathfrak{g}$ and $\rho = \mathrm{ad}$. By this bijection, we can find in a natural way an exterior covariant derivative on the vector bundle. The representations on coordinate patches of this exterior covariant derivative will be our covariant derivative. We will also denote the induced representation of $\mathfrak{g}$ by $\rho : \mathfrak{g} \to \mathfrak{gl}(F)$.

Basic forms

Definition 2.3.1. Let F be a vector space such that $\rho : G \to \mathrm{GL}(F)$ is a representation. A form $\bar{\alpha} \in \Omega^k(P, F)$ is called *invariant* if

$$R_g^*\bar{\alpha}|_{pg} = \rho\Big(g^{-1}\Big) \circ \bar{\alpha}|_p \tag{2.3.1}$$

for all $g \in G$. If $\bar{\alpha}$ is both horizontal and invariant, it is called a *basic*. The subspace of $\Omega^k(P, F)$ of basic forms is denoted with $\Omega^k_G(P, F)$.

Definition 2.3.2. We define $\bar{H}_i : \Omega^k_G(P, F) \to \Omega^k(U_i, F)$ and $\bar{I}_i : \Omega^k(U_i, F) \to \Omega^k_G(\pi^{-1}(U_i), F)$ by

$$\bar{H}_i(\bar{\alpha}) = s_i^*\bar{\alpha}; \tag{2.3.2}$$

$$\bar{I}_i(\alpha_i)|_p = \rho\Big(g_i(p)^{-1}\Big) \circ \pi^*\alpha_i|_{\pi(p)}. \tag{2.3.3}$$

Proposition 2.3.3.

1. For all $\bar{\alpha} \in \Omega^k_G(P,F)$, the forms $\bar{H}_i(\bar{\alpha}) \in \Omega^k(U_i,F)$ are the representations of a form $\bar{H}(\bar{\alpha}) \in \Omega^k(M, P \times_\rho F)$.

2. If $\alpha \in \Omega^k(M, P \times_\rho F)$ is represented by $\alpha_i \in \Omega^k(U_i,F)$, the forms $\bar{I}_i(\alpha_i)$ define a global form $\bar{I}(\alpha)$ on $\Omega^k(P,F)$.

3. The maps $\bar{H}_i$ restricted to $\Omega^k_G(\pi^{-1}(U_i),F)$ and $\bar{I}_i$ are each other's inverses.

4. The map $\bar{H}$ is an isomorphism with inverse $\bar{I}$.

Proof.

1. Let $\bar{\alpha} \in \Omega^k_G(P,F)$. We shall first show that $\bar{H}_i(\bar{\alpha})$ satisfy (1.5.10), and thus are representations of a form $\bar{H}(\bar{\alpha}) \in \Omega^k(M, P \times_\rho F)$ by Corollary 1.5.11. We start by remarking that by (1.3.8), we have $R_{g_{ji}(x)} \circ s_j(x) = s_i(x)$ for $x \in U_{ij}$, so

$$\begin{aligned}
\bar{H}_i(\bar{\alpha})|_x &= s_i^* \bar{\alpha}|_{s_i(x)} \\
&= (R_{g_{ji}(x)} \circ s_j)^* \bar{\alpha}|_{s_j(x)g_{ji}(x)} \\
&= s_j^* R^*_{g_{ji}(x)} \bar{\alpha}|_{s_j(x)g_{ji}(x)} \\
&= s_j^* \rho\Big(g_{ji}(x)^{-1}\Big) \circ \bar{\alpha}|_{s_j(x)} && \text{[By (2.3.1)]} \\
&= \rho\Big(g_{ij}(x)\Big) \circ \bar{H}_j(\bar{\alpha})|_x.
\end{aligned}$$

2. Given $\alpha \in \Omega^k(M, P \times_\rho F)$, we will prove that $\bar{\alpha}_i$ obtained by (2.3.3) is the restriction of a global form $\bar{\alpha} \in \Omega^k(P,F)$ by showing that the $\bar{\alpha}_i$ agree on overlaps U_{ij}.

$$\begin{aligned}
\bar{I}_i(\alpha_i)|_p &= \rho\Big(g_i(p)^{-1}\Big) \circ \pi^* \alpha_i|_{\pi(p)} \\
&= \rho\Big((g_{ij}(x) g_j(p))^{-1}\Big) \circ \pi^* \alpha_i|_{\pi(p)} && \text{[By Proposition 1.1.3]} \\
&= \rho\Big(g_j(p)^{-1} g_{ij}(x)^{-1}\Big) \circ \pi^* \rho\Big(g_{ij}(x)\Big) \circ \alpha_j|_{\pi(p)} && \text{[By (1.5.10)]} \\
&= \rho\Big(g_j(p)^{-1}\Big) \circ \pi^* \circ \alpha_j|_{\pi(p)} && [\rho \text{ is a representation}] \\
&= \bar{I}_j(\alpha_j)|_p.
\end{aligned}$$

3. We will first show that $\bar{H}_i \circ \bar{I}_i = \mathrm{Id}$. Let α_i be a representing function of $\alpha \in \Omega^k(M, P \times_\rho F)$. Then

$$\begin{aligned}
\bar{H}_i \circ \bar{I}_i(\alpha_i)|_x &= s_i^* \rho\Big(g_i(s_i(p))^{-1}\Big) \circ \pi^* \alpha_i|_{\pi \circ s_i(x)} \\
&= \rho(e) \circ (\pi \circ s_i)^* \alpha_i|_{\pi \circ s_i(x)} && \text{[By (1.3.5)]} \\
&= \alpha_i|_x && [\pi \circ s_i = \mathrm{Id.}]
\end{aligned}$$

Conversely, let $\bar{\alpha}_i$ be the restriction of $\bar{\alpha} \in \Omega^k_G(P,G)$ to $\pi^{-1}(U_i)$. By (1.3.7) we have $s_i \circ \pi(p) = R_{g_i(p)^{-1}} p$, so we find

$$\begin{aligned}
\bar{I}_i \circ \bar{H}_i(\bar{\alpha}_i)|_p &= \rho\Big(g_i(p)^{-1}\Big) \circ \pi^* s_i^* \bar{\alpha}_i|_{s_i \circ \pi(p)} \\
&= \rho\Big(g_i(p)^{-1}\Big) \circ (s_i \circ \pi)^* \bar{\alpha}_i|_{p g_i(p)^{-1}} \\
&= \rho\Big(g_i(p)^{-1}\Big) R^*_{g_i(p)^{-1}} \bar{\alpha}_i|_{p g_i(p)^{-1}} \\
&= \bar{\alpha}_i|_p && \text{[By (2.3.1).]}
\end{aligned}$$

4. Follows directly from (1)-(3) and the fact that $\bar{H}_i$ and $\bar{I}_i$ are clearly linear in their arguments.

$\square$

Corollary 2.3.4. Let E be an arbitrary vector bundle over M with fibre F. Then every $\alpha \in \Omega^k(M, E)$ can be represented by $\alpha_i \in \Omega^k(M, F)$.

Proof. By Corollary 1.5.9 we find that there is a principal bundle P such that $E = P \times_\rho F$. Then the statement follows directly from previous proposition. $\square$

The exterior covariant derivative on an associated vector bundle

Proposition 2.3.5. Let $\bar{\alpha} \in \Omega^k_G(P, F)$. Then we have

$$\mathrm{d}_\omega \bar{\alpha} = \mathrm{d}\bar{\alpha} + \rho(\omega) \wedge \bar{\alpha} \in \Omega^{k+1}_G(P, F). \tag{2.3.4}$$

Proof. Let $X_1, \ldots, X_{k+1}$ be vector fields on P. Then we have by

$$\begin{aligned}
\mathrm{d}_\omega \bar{\alpha}(X_1, \ldots, X_{k+1}) &= \mathrm{d}\alpha(HX_1, \ldots, HX_{k+1}) \\
&= \sum_{i=1}^{k+1} (-1)^{i-1} HX_i\Big(\bar{\alpha}(HX_1, \ldots, \widehat{HX}_i, \ldots, HX_{k+1})\Big) \\
&\quad + \sum_{i<j} (-1)^{i+j} \bar{\alpha}\Big([HX_i, HX_j], HX_1, \ldots, \widehat{H}X_i, \ldots, \widehat{H}X_j \ldots, HX_{k+1}\Big),
\end{aligned}$$

where we used (A.2.4). Since $\bar{\alpha}$ is horizontal, we can replace in the first term the factor $\bar{\alpha}(HX_1, \ldots, \widehat{HX}_i, \ldots, HX_{k+1})$ by $\bar{\alpha}(X_1, \ldots, \widehat{H}_i, \ldots, X_{k+1})$. In the second term, we replace all occurrences of HX_i by $X_i - X_i^V$, where X_i^V denotes the vertical component of X_i. Since $\bar{\alpha}$ is horizontal, we see that all terms containing X_i^V vanish, also the terms containing commutators $[X_i^V, X_j^V]$, $[X_i, X_j^V]$ or $[X_i^V, X_j]$, since these commutators are vertical by Proposition 2.1.11. So we find

$$\begin{aligned}
\mathrm{d}_\omega \bar{\alpha}(X_1, \ldots, X_{k+1}) =& \sum_{i=1}^{k+1} (-1)^{i-1} HX_i\Big(\bar{\alpha}(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\Big) \\
&+ \sum_{i<j} (-1)^{i+j} \bar{\alpha}\Big([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j \ldots, X_{k+1}\Big) \\
=& \mathrm{d}\bar{\alpha}(X_1, \ldots, X_{k+1}) - \sum_{i=1}^{k+1} (-1)^{i-1} X_i^V \Big(\bar{\alpha}(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\Big),
\end{aligned}$$

where we once again used (A.2.4). Since X_i^V is the vertical component of X_i and the horizontal projection operator H satisfies $H = \mathrm{Id} - \sigma \circ \omega$, we find $X_i^V = \sigma \circ \omega(X_i)$. Then if we write $g(t) = e^{t\omega(X_i)}$, we find by definition of σ and the fact that $\bar{\alpha}$ is invariant:

$$\begin{aligned}
X_i^V \bar{\alpha} &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} R^*_{g(t)} \bar{\alpha} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \rho(g(t)^{-1}) \bar{\alpha} && \text{[By (2.3.1)]} \\
&= -\rho\Big(\omega(X_i)\Big)\bar{\alpha},
\end{aligned}$$

hence we find

$$\mathrm{d}_\omega \bar{\alpha}(X_1, \ldots, X_{k+1}) = \mathrm{d}\bar{\alpha}(X_1, \ldots, X_{k+1}) + \sum_{i=1}^{k+1} (-1)^{i-1} \rho\Big(\omega(X_i)\Big)\Big(\bar{\alpha}(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1})\Big).$$

Since $\bar{\alpha}$ is totally antisymmetric, the second term equals

$$\frac{1}{k!}\sum_{P\in S_{k+1}} \operatorname{sign}(P)\rho\Big(\omega(X_{P(1)})\Big)\bar{\alpha}(X_{P(2)},\ldots,\widehat{X}_{P(1)},\ldots,X_{P(k+1)}),$$

where P is a permutation. By definition of the wedge product (A.1.1), we see that this equals $\rho(\omega)\wedge\bar{\alpha}(X_1,\ldots,X_{k+1})$. In order to show that $\mathrm{d}_\omega\bar{\alpha}$ is basic, we remark that we already have that $\mathrm{d}_\omega\bar{\alpha}$ is horizontal by definition of d_ω. We show that $\mathrm{d}_\omega\bar{\alpha}$ is invariant:

$$\begin{aligned}
R_g^*\mathrm{d}_\omega\bar{\alpha} &= R_g^* H^\vee \mathrm{d}\bar{\alpha} \\
&= H^\vee R_g^*\mathrm{d}\bar{\alpha} && \text{[By Proposition 2.1.10]} \\
&= H^\vee \mathrm{d}R_g^*\bar{\alpha} && [R_g^* \text{ and d commute}] \\
&= H^\vee \mathrm{d}\Big(\rho(g^{-1})\circ\bar{\alpha}\Big) && \text{[By (2.3.1)]} \\
&= \rho(g^{-1})\circ \mathrm{d}_\omega\bar{\alpha}.
\end{aligned}$$

□

Next corollary shows that in contrary to d, we do not have d_ω^2.

Corollary 2.3.6. Let $\bar{\alpha}\in\Omega_G^k(P,F)$. Then $\mathrm{d}_\omega^2\bar{\alpha}=\rho(\Omega)\wedge\bar{\alpha}$.

Proof. By direct calculation, we find

$$\begin{aligned}
\mathrm{d}_\omega^2\bar{\alpha} &= \mathrm{d}_\omega(\mathrm{d}\bar{\alpha}+\rho(\omega)\wedge\bar{\alpha}) \\
&= H^\vee\mathrm{d}(\mathrm{d}\bar{\alpha}+\rho(\omega)\wedge\bar{\alpha}) \\
&= H^\vee(\rho(\mathrm{d}\omega)\wedge\bar{\alpha}-\rho(\omega)\wedge\mathrm{d}\bar{\alpha}) && \text{[By Proposition } A.2.2] \\
&= \rho(H^\vee\mathrm{d}\omega)\wedge\bar{\alpha} && [\omega \text{ kills } H^\vee\mathrm{d}\bar{\alpha}] \\
&= \rho(\Omega)\wedge\bar{\alpha}.
\end{aligned}$$

□

Now we are able to define the exterior covariant derivative on $\Omega^k(M,P\times_\rho F)$.

Definition 2.3.7. Let $\alpha\in\Omega^k(M,P\times_\rho F)$ and let $\bar{\alpha}\in\Omega_G^k(P,F)$ be the corresponding basic form on P. Then we define $\mathrm{d}_{\mathcal{A}}\alpha$ as the corresponding form of $\mathrm{d}_\omega\bar{\alpha}$ in $\Omega^{k+1}(M,P\times_\rho F)$.

The covariant derivative

Definition 2.3.8. We define *covariant derivative* of a form $\alpha_i\in\Omega^k(U_i,F)$ as

$$\mathrm{d}_{\mathcal{A}_i}\alpha_i=\mathrm{d}\alpha_i+\rho(\mathcal{A}_i)\wedge\alpha_i\in\Omega^{k+1}(U_i,F). \tag{2.3.5}$$

With the machinery developed in previous paragraphs, we can see that this is not an arbitrary definition. Next proposition shows that the covariant derivative is a derivative on $\Omega^k(U_i,F)$ which is compatible with the connection.

Proposition 2.3.9. Let $\alpha\in\Omega^k(M,P\times_\rho F)$ and let $\alpha_i\in\Omega^k(U_i,F)$ be the corresponding representatives. Then $\mathrm{d}_{\mathcal{A}_i}\alpha_i$ is exactly the representative $(\mathrm{d}_{\mathcal{A}}\alpha)_i$ of $\mathrm{d}_{\mathcal{A}}\alpha$ in $\Omega^{k+1}(U_i,F)$.

Proof. By (2.3.2) and Proposition 2.3.3, we have to show that $\mathrm{d}_{\mathcal{A}_i}\alpha_i = s_i^*\mathrm{d}_\omega\bar{\alpha}$ if $\bar{\alpha} \in \Omega^k_G(P, F)$ corresponds with α. By direct calculation:

$$\begin{aligned}\mathrm{d}_{\mathcal{A}_i}\alpha &= \mathrm{d}\alpha_i + \rho(\mathcal{A}_i) \wedge \alpha_i \\ &= \mathrm{d}s_i^*\bar{\alpha} + \rho(s_i^*\omega) \wedge s_i^*\bar{\alpha} \\ &= s_i^*(\mathrm{d}\bar{\alpha} + \rho(\omega) \wedge \bar{\alpha}) \\ &= s_i^*\mathrm{d}_\omega\bar{\alpha}.\end{aligned}$$

□

Corollary 2.3.10. If $\alpha_i \in \Omega^k(U_i, F)$ are the representatives of $\alpha \in \Omega^k(M, P \times_\rho F)$, we have

$$\mathrm{d}_{\mathcal{A}_i}\alpha_i|_x = \rho\Big(g_{ij}(x)\Big) \circ \mathrm{d}_{\mathcal{A}_j}\alpha_j|_x. \tag{2.3.6}$$

Proof. Since the $\mathrm{d}_{\mathcal{A}_i}\alpha_i$ are representatives of $\mathrm{d}_{\mathcal{A}}\alpha \in \Omega^k(M, P \times_\rho F)$ by definition, they automatically satisfy above identity by Corollary 1.5.11. □

Sometimes it is convenient to know what $\mathrm{d}_{\mathcal{A}}$ is in coordinates.

Proposition 2.3.11. If we drop for a moment the subscript i, let

$$(\mathrm{d}_{\mathcal{A}})_\mu = \partial_\mu + \rho(\mathcal{A}_\mu) \tag{2.3.7}$$

Then we have $\mathrm{d}_{\mathcal{A}}\alpha = (\mathrm{d}_{\mathcal{A}})_\mu \mathrm{d}x^\mu \wedge \alpha$ for $\alpha \in \Omega^k(U, F)$.

Proof. Let $\alpha \in \Omega^1(U, F)$. Then we find by direct calculation

$$\begin{aligned}\mathrm{d}_{\mathcal{A}}\alpha &= \mathrm{d}\mathcal{A} + \rho(\mathcal{A}) \wedge \alpha \\ &= \mathrm{d}\alpha_{\nu_1\ldots\nu_k}\mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_k} + \rho(\mathcal{A}_\mu) \wedge \alpha_{\nu_1\ldots\nu_k}\mathrm{d}x^\mu \wedge \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_k} \\ &= \partial_\mu\alpha_{\nu_1\ldots\nu_k}\mathrm{d}x^\mu \wedge \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_k} + \rho(\mathcal{A}_\mu) \wedge \alpha_{\nu_1\ldots\nu_k}\mathrm{d}x^\mu \wedge \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_k} \\ &= (\mathrm{d}_{\mathcal{A}})_\mu\alpha_{\nu_1\ldots\nu_k}\mathrm{d}x^\mu \wedge \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_k} \\ &= (\mathrm{d}_{\mathcal{A}})_\mu\mathrm{d}x^\mu \wedge \alpha.\end{aligned}$$

□

Remark. We often write $(\mathrm{d}_A)_\mu = \partial_\mu + \mathcal{A}_\mu$ if it is clear which representation we use.

Proposition 2.3.12. We have

$$\mathrm{d}^2_{\mathcal{A}_i}\alpha_i = \rho(\mathcal{F}_i) \wedge \alpha_i \tag{2.3.8}$$

for all $\alpha_i \in \Omega^k(U_i, F)$.

Proof. By direct calculation:

$$\begin{aligned}\mathrm{d}^2_{\mathcal{A}_i}\alpha_i &= \mathrm{d}_{\mathcal{A}_i}(\mathrm{d}\alpha_i + \rho(\mathcal{A}_i) \wedge \alpha \\ &= \mathrm{d}(\mathrm{d}\alpha_i + \rho(\mathcal{A}_i) \wedge \alpha_i) + \rho(\mathcal{A}_i) \wedge (\mathrm{d}\alpha_i + \rho(\mathcal{A}_i) \wedge \alpha_i) \\ &= \rho(\mathrm{d}\mathcal{A}_i) \wedge \alpha_i - \rho(\mathcal{A}_i) \wedge \mathrm{d}\alpha_i + \rho(\mathcal{A}_i) \wedge \mathrm{d}\alpha_i + \rho(\mathcal{A}_i \wedge \mathcal{A}_i) \wedge \alpha_i \quad \text{[By Proposition } A.2.2] \\ &= \rho(\mathrm{d}\mathcal{A}_i + \mathcal{A}_i \wedge \mathcal{A}_i) \wedge \alpha_i \\ &= \rho(\mathcal{F}_i) \wedge \alpha_i \quad \text{[By (2.2.9).]}\end{aligned}$$

□

We also have a local expression for $\mathrm{d}^2_{\mathcal{A}_i}$.

Proposition 2.3.13. Forgetting the subscript i for the moment, we have

$$(\mathrm{d}_{\mathcal{A}}^2)_{\mu\nu} = \rho(\mathcal{F}_{\mu\nu}) = \Big[(\mathrm{d}_{\mathcal{A}})_\mu, (\mathrm{d}_{\mathcal{A}})_\nu\Big]. \tag{2.3.9}$$

Proof. Let $\alpha \in \Omega^k(U, F)$. Then we find by direct calculation

$$\begin{aligned}
\Big[(\mathrm{d}_{\mathcal{A}})_\mu, (\mathrm{d}_{\mathcal{A}})_\nu\Big]\alpha &= \Big(\partial_\mu + \rho(\mathcal{A}_\mu)\Big)\Big(\partial_\nu + \rho(\mathcal{A}_\nu)\Big) \\
&= \partial_\mu\partial_\nu\alpha - \partial_\nu\partial_\mu\alpha + \partial_\mu\Big(\rho(\mathcal{A}_\nu)\alpha\Big) - \rho(\mathcal{A}_\nu)\partial_\mu\alpha \\
&\quad - \partial_\nu\Big(\rho(\mathcal{A}_\mu)\alpha\Big) + \rho(\mathcal{A}_\mu)\partial_\nu\alpha + \rho(\mathcal{A}_\mu)\rho(\mathcal{A}_\nu)\alpha - \rho(\mathcal{A}_\nu)\rho(\mathcal{A}_\mu)\alpha \\
&= \Big(\partial_\mu\rho(\mathcal{A}_\nu)\Big)\alpha - \Big(\partial_\nu\rho(\mathcal{A}_\mu)\Big)\alpha + \rho(\mathcal{A}_\mu)\rho(\mathcal{A}_\nu)\alpha - \rho(\mathcal{A}_\nu)\rho(\mathcal{A}_\mu)\alpha \\
&= \rho\Big(\partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]\Big)\alpha,
\end{aligned}$$

where we used that ∂_μ and ∂_ν commute and the fact that ρ is a representation, so a homomorphism of algebras. By (2.2.12) we find that $\Big[(\partial_{\mathcal{A}})_\mu, (\partial_{\mathcal{A}})_\nu\Big] = \rho(\mathcal{F}_{\mu\nu})$. By previous proposition we have $(\mathrm{d}_{\mathcal{A}}^2)_{\mu\nu} = \rho(\mathcal{F}_{\mu\nu})$, which completes the proof. ☐

Since the adjoint bundle is the most interesting associated bundle, we assign a special symbol for the covariant derivative on ad P.

Definition 2.3.14. For $F = \mathfrak{g}$ and $\rho = \mathrm{ad}$ we write $\mathcal{D}_{\mathcal{A}_i}$ instead of $\mathrm{d}_{\mathcal{A}_i}$.

Lemma 2.3.15. Let $\eta \in \Omega^n(U_i, \mathfrak{g})$ and $\alpha_i \in \Omega^k(U_i, \mathfrak{g})$. Then we have $\mathrm{ad}(\eta) \wedge \alpha_i = [\eta, \alpha_i]$.

Proof. Let $\{T_i\}$ is a basis of $\mathfrak{g}$. Then if we use that $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is the induced representation of $\mathfrak{g}$, we find

$$\begin{aligned}
\mathrm{ad}(\eta) \wedge \alpha_i &= \mathrm{ad}(T_a)T_b\eta^a \wedge \alpha_i^b \\
&= [T_a, T_b]\eta^a \wedge \alpha_i^b \\
&= [\eta, \alpha_i],
\end{aligned}$$

☐

Proposition 2.3.16. We have the following identities for $\mathcal{D}_{\mathcal{A}_i}$:

$$\mathcal{D}_{\mathcal{A}_i}\alpha_i = \mathrm{d}\alpha_i + [\mathcal{A}_i, \alpha_i] \tag{2.3.10}$$
$$\mathcal{D}_{\mathcal{A}_i}^2\alpha_i = [\mathcal{F}_i, \alpha_i], \tag{2.3.11}$$

where $\alpha_i \in \Omega^k(U_i, \mathfrak{g})$.

Proof. This follows directly from previous lemma and the identities (2.3.5) and (2.3.8). ☐

Notice that equation (2.3.10) by (A.1.2) is equivalent to

$$\mathcal{D}_{\mathcal{A}_i}\alpha_i = \mathrm{d}\alpha_i + \mathcal{A}_i \wedge \alpha_i - (-1)^k\alpha_i \wedge \mathcal{A}_i. \tag{2.3.12}$$

As a corollary, we can give an alternative formulation of the Bianchi identity (2.2.10) in terms of the covariant derivative.

Corollary 2.3.17. We have

$$\mathcal{D}_{\mathcal{A}_i}\mathcal{F}_i = 0. \tag{2.3.13}$$

2.4 The gauge group

In this section, we give a definition of the gauge group. This is an important object in mathematics, since it forms the group of automorphisms of a fibre bundle: If Φ is an element of the gauge group of a principal bundle $P(M,G)$, then also $\Phi(P)(M,G)$ is a fibre bundle. The gauge group is also important for physicists, since it turns out to be the group of transformations leaving the Yang-Mills equations invariant.

Gauge transformations

Definition 2.4.1. Let $P(M,G)$ be a principal bundle. Then a diffeomorphism $\Phi : P \to P$ is called a *gauge transformation* if it satisfies the following two properties:

1. Φ preserves fibres. In other words, the diagram

$$\begin{tikzcd} P \arrow[rr, "\Phi"] \arrow[dr, "\pi"'] & & P \arrow[dl, "\pi"] \\ & M & \end{tikzcd}$$

   commutes. This means $\pi \circ \Phi = \pi$;

2. Φ is G-equivariant: $\Phi \circ R_g = R_g \circ \Phi$, or equivalently, $\Phi(pg) = \Phi(p)g$ for all $g \in G$.

With respect to the composition of maps, the set of gauge transformations is a group called the *gauge group*.

An alternative description of the gauge group

There are equivalent ways to define the gauge group, which are important, since alternative descriptions can give us more insight into what a gauge group is. In order to explore them, we we need the following definition, which can be made since $\Phi(p)$ and p are in the same fibre.

Definition 2.4.2. Let $\psi_i : U_i \times G \to \pi^{-1}(U_i)$ be a collection of local trivializations. Then we define the following maps between function spaces:

1. $\widehat{H}_i : \mathcal{G} \to C^\infty\Big(\pi^{-1}(U_i), G\Big)$ by $\widehat{H}_i(\Phi)(p) = g_i(p)^{-1} g_i\Big(\Phi(p)\Big)$,

2. $\widetilde{H}_i : \mathcal{G} \to C^\infty\Big(\pi^{-1}(U_i), G\Big)$ by $\widetilde{H}_i(\Phi)(p) = g_i\Big(\Phi(p)\Big) g_i(p)^{-1}$.

With these maps, we shall see that that the gauge group is isomorphic as group to two other function spaces. The first of these spaces is introduced in next definition.

Definition 2.4.3. $C^\infty_{\mathrm{Ad}}(P,G)$ is the space of function $\widehat{f} \in C^\infty(P,G)$ satisfying

$$\widehat{f}(pg) = g^{-1} \widehat{f}(p) g \tag{2.4.1}$$

for all $p \in P$ and $g \in G$, which can be made into a group by pointwise multiplication.

Before we prove that an isomorphism exists, we state three lemmas.

Lemma 2.4.4. The $\widehat{H}_i(\Phi)$ define a global function $\widehat{H}(\Phi) \in C^\infty_{\mathrm{Ad}}(P,G)$.

Proof. First we show that the $\widehat{H}_i(\Phi)$ satisfy (2.4.1). Let $g \in G$, then we have

$$\begin{aligned}\widehat{H}_i(\Phi)(pg) &= g_i(pg)^{-1} g_i\big(\Phi(pg)\big)\\ &= \big(g_i(p)g\big)^{-1} g_i\big(\Phi(p)\big)g && \text{[Equivariance of } \Phi \text{ and } g_i]\\ &= g^{-1}\widehat{H}_i(\Phi)(p)g.\end{aligned}$$

Then if $p \in \pi^{-1}(U_{ij})$ and $x = \pi(p)$, we find

$$\begin{aligned}\widehat{H}_i(\Phi)(p) &= g_i(p)^{-1} g_i\big(\Phi(p)\big)\\ &= \big(g_{ij}(x)g_j(p)^{-1}\big) g_{ij}(x) g_j\big(\Phi(p)\big) && [\Phi(p) \in \pi^{-1}(x)]\\ &= g_j(p)^{-1} g_j\big(\Phi(p)\big)\\ &= \widehat{H}_j(\Phi)(p)\end{aligned}$$

hence the $\widehat{H}_i(\Phi)$ agree on overlaps, which guarantees that $\widehat{H}(\Phi)$ is well-defined. $\widehat{H}(\Phi)$ satisfies (2.4.1) since the $\widehat{H}_i(\Phi)$ do. □

Lemma 2.4.5. Let $\Phi \in \mathcal{G}$, Then $\widehat{H}(\Phi)$ satisfies $\Phi(p) = p\widehat{H}(\Phi)(p)$.

Proof. Let $p \in P$ and U_i the coordinate neighborhood such that $p \in \pi^{-1}(U_i)$. Since $\Phi(p)$ and p are in the same fibre and the right action of G on P is free and transitive, there is a unique $g_p \in G$ such that $\Phi(p) = pg_p$. Then by the equivariance of g_i we have $g_i(p)g_p = g_i(pg_p) = g_i\big(\Phi(p)\big)$. Multiplying on the left with $g_i(p)^{-1}$ gives now $g_p = g_i(p)^{-1} g_i\big(\Phi(p)\big) = \widehat{H}(\Phi)(p)$. □

Lemma 2.4.6. Let $\widehat{\phi} \in C^\infty_{\mathrm{Ad}}(P, G)$. Then $\Phi : P \to P$ given by $p \mapsto p\widehat{\phi}(p)$ is a gauge transformation, which satisfies $\widehat{H}(\Phi) = \widehat{\phi}$.

Proof. We shall first prove that Φ defined as in the Lemma is a gauge transformation. Firstly, Φ preserves fibres since $\pi \circ \Phi(p) = \pi\big(p\widehat{\phi}(p)\big) = \pi(p)$, where we used the G-invariance of π. Secondly, Φ is G-equivariant, since

$$\begin{aligned}\Phi(pg) &= pg\widehat{\phi}(pg)\\ &= p\widehat{\phi}(p) && \text{[By (2.4.1)]}\\ &= \Phi(p)g.\end{aligned}$$

Now let U_i be the coordinate neighborhood such that $p \in \pi^{-1}(U_i)$. Then we have

$$\begin{aligned}\widehat{H}(\Phi)(p) &= g_i(p)^{-1} g_i\big(\Phi(p)\big)\\ &= g_i(p)^{-1} g_i\big(p\widehat{\phi}(p)\big)\\ &= \widehat{\phi}(p) && \text{[Equivariance of } g_i],\end{aligned}$$

which finishes the proof. □

Theorem 2.4.7. $\widehat{H} : \mathcal{G} \to C^\infty_{\mathrm{Ad}}(P, G)$ is an isomorphism.

Proof. Lemma 2.4.4 assures that $\widehat{H}$ maps gauge transformations into $C^\infty_{\mathrm{Ad}}(P, G)$, while Lemma 2.4.6 assures that $\widehat{H}$ is surjective. Injectivity can be proved as follows. Let $\widehat{\phi}_1, \widehat{\phi}_2 \in C^\infty_{\mathrm{Ad}}(P, G)$ and let $\Phi_1, \Phi_2 \in \mathcal{G}$ such that by Lemma 2.4.6 we have $\widehat{H}(\Phi_i) = \widehat{\psi}_i$. Then $p\widehat{\psi}_1(p) = p\widehat{\psi}_2(p)$ and by Lemma 2.4.5 we find $\Phi_1(p) = \Phi_2(p)$.

What remains to prove is that $\widehat{H}$ is a homomorphism, so $p \in P$. Then we find

$$\begin{aligned} p\widehat{H}(\Phi \circ \Psi)(p) &= \Phi \circ \Psi(p) && \text{[Lemma 2.4.5]} \\ &= \Phi\Big(p\widehat{H}(\Psi)(p)\Big) && \text{[Lemma 2.4.5]} \\ &= \Phi(p)\widehat{H}(\Psi)(p) && \text{[Equivariance of } \Phi] \\ &= p\widehat{H}(\Phi)(p)\widehat{H}(\Psi)(p) && \text{[Lemma 2.4.5]} \\ &= p\Big(\widehat{H}(\Phi)\widehat{H}(\Psi)\Big)(p) && \text{[Def. multiplication on } C^\infty_{\mathrm{Ad}}(P,G)], \end{aligned}$$

and since the right action of G on P is free, we find $\widehat{H}(\Phi \circ \Psi)(p) = \Big(\widehat{H}(\Phi)\widehat{H}(\Psi)\Big)(p)$. □

Convention. From now on we shall denote $\widehat{H}(\Phi)$ with $\widehat{\phi}$.

A second alternative description of the gauge group

It is also possible to describe gauge transformations in terms of functions $U_i \to G$ which are constant on the fibres. We shall see that the maps $\widetilde{H}_i$ will play a role in this. This description turns out to be the the most important of the two alternative descriptions, since it describes the gauge group in terms of an associated vector bundle.

Lemma 2.4.8. Let Φ be a gauge transformation. Then Φ and $\widetilde{H}_i$ satisfy

1. $\widetilde{H}_i(\Phi)(p) = \mathrm{Ad}_{g_i(p)}\widehat{H}(\Phi)$,
2. $\Phi(p) = p\ \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{H}_i(\Phi)(p)$.

Proof.

1. Follows directly from the definitions of $\widetilde{H}_i$ and $\widehat{H}$.
2. Follows from 1. and Lemma 2.4.5.

□

Proposition 2.4.9. $\widetilde{H}_i$ is a homomorphism

Proof. From the definitions of $\widehat{H}_i$ and $\widetilde{H}_i$ we find $\widetilde{H}_i(\Phi)(p) = \mathrm{Ad}_{g_i(p)}\widehat{H}_i(\Phi)(p)$. Since $\widehat{H}_i$ is a homomorphism, we find

$$\begin{aligned} \widetilde{H}_i(\Psi \circ \Phi)(p) &= \mathrm{Ad}_{g_i(p)}\widehat{H}_i(\Psi \circ \Phi)(p) \\ &= \mathrm{Ad}_{g_i(p)}\Big(\widehat{H}_i(\Psi)\widehat{H}_i(\Phi)\Big)(p) \\ &= \mathrm{Ad}_{g_i(p)}\widehat{H}_i(\Psi)(p)\widehat{H}_i(\Phi)(p) \\ &= \mathrm{Ad}_{g_i(p)}\widehat{H}_i(\Psi)(p)\mathrm{Ad}_{g_i(p)}\widehat{H}_i(\Phi)(p) \\ &= \widetilde{H}_i(\Psi)(p)\widetilde{H}_i(\Phi)(p) \\ &= \Big(\widetilde{H}_i(\Psi)\widetilde{H}_i(\Phi)\Big)(p), \end{aligned}$$

□

Lemma 2.4.10. Let $\widetilde{\phi}_i : \pi^{-1}(U_i) \to G$ satisfy

$$\widetilde{\phi}_i(pg) = \widetilde{\phi}_i(p) \tag{2.4.2}$$

$$\mathrm{Ad}_{g_{ij}(x)}\widetilde{\phi}_j(p) = \widetilde{\phi}_i(p) \qquad [p \in \pi^{-1}(U_{ij})]. \tag{2.4.3}$$

Then Φ defined by $\Phi(p) = p\ \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{\phi}_i(p)$ is the unique gauge transformation which satisfies $\widetilde{H}_i(\Phi) = \widetilde{\phi}_i$.

Proof. Let $\widehat{\phi}_i(p) = \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{\phi}_i(p)$. Then $\widehat{\phi}_i$ is globally defined on P, since on overlaps $\pi^{-1}(U_{ij})$ we have

$$\begin{aligned}
\widehat{\phi}_i(p) &= \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{\phi}_i(p) \\
&= \mathrm{Ad}_{g_i(p)^{-1}}\mathrm{Ad}_{g_{ij}(x)}\widetilde{\phi}_j(p) && \text{[By (2.4.3)]} \\
&= \mathrm{Ad}_{g_{ji}(x)g_i(p)^{-1}}\widetilde{\phi}_j(p) \\
&= \mathrm{Ad}_{g_j(p)^{-1}}\widetilde{\phi}_j(p) \\
&= \widehat{\phi}_j(p)
\end{aligned}$$

so we can define $\widehat{\phi} = \widehat{\phi}_i$. Furthermore $\widehat{\phi} \in C^\infty_{\mathrm{Ad}}(P, G)$, since

$$\begin{aligned}
\widehat{\phi}(pg) &= \mathrm{Ad}_{g_i(pg)^{-1}}\widetilde{\phi}_i(pg) \\
&= \mathrm{Ad}_{g^{-1}g_i(p)^{-1}}\widetilde{\phi}_i(p) && \text{[By (2.4.2)]} \\
&= g^{-1}\mathrm{Ad}_{g_i(p)^{-1}}\widetilde{\phi}_i(p)g \\
&= g^{-1}\widehat{\phi}(p)g.
\end{aligned}$$

From Lemma 2.4.6 we find that $\Phi(p) = p\widehat{\phi}(p)$ is a gauge transformation, such that $\widehat{H}(\Phi) = \widehat{\phi}$. By Lemma 2.4.8, we have $\widehat{H}(\Phi)(p) = \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{H}_i(\Phi)(p)$, whence $\widetilde{H}_i(\Phi) = \widetilde{\phi}_i$. Uniqueness follows from the fact that the maps $q \mapsto \mathrm{Ad}_{g_i(p)^{-1}}q$ and $\widehat{H}$ are bijective. □

Lemma 2.4.11. Let Φ be a gauge transformation. Then $\widetilde{H}_i(\Phi)$ satisfies (2.4.2) and (2.4.3).

Proof. By definition of $\widetilde{H}_i(\Phi)$ we find

$$\begin{aligned}
\widetilde{H}_i(\Phi)(pg) &= g_i\Big(\Phi(pg)\Big)g_i(pg)^{-1} \\
&= g_i\Big(\Phi(p)\Big)g_i(p)^{-1} && \text{[Equivariance of } \Phi \text{ and } g_i\text{]} \\
&= \widetilde{H}_i(\Phi)(p),
\end{aligned}$$

which proves that $\widetilde{H}_i(\Phi)$ satisfies (2.4.2). From

$$\begin{aligned}
\mathrm{Ad}_{g_{ij}(x)}\widetilde{H}_j(\Phi)(p) &= \mathrm{Ad}_{g_{ij}(x)}g_j\Big(\Phi(p)\Big)g_j(p)^{-1} \\
&= g_{ij}(x)g_j\Big(\Phi(p)\Big)\Big(g_{ij}(x)g_j(p)\Big)^{-1} \\
&= g_i\Big(\Phi(p)\Big)g_i(p)^{-1} && [p, \Phi(p) \in \pi^{-1}(x)] \\
&= \widetilde{H}_i(\Phi)(p),
\end{aligned}$$

we see that $\widetilde{H}_i(\Phi)$ also satisfies (2.4.3). □

To summarize, we found that a gauge transformation can be described by a family of functions $\widetilde{\phi}_i : \pi^{-1}(U_i) \to G$ satisfying the two identities (2.4.2) and (2.4.3). The first identity allows us to apply Proposition (1.3.5), so we can find a function $\phi : U_i \to G$ satisfying $\phi \circ \pi = \widetilde{\phi}$, which gives rise the next proposition.

Proposition 2.4.12. Let $H_i : \mathcal{G} \to C^\infty(U_i, G)$ be given by $H_i(\Phi) = g_i \circ \Phi \circ s_i$. Then

1. $H_i(\Phi) \circ \pi = \widetilde{H}_i(\Phi)$,
2. H_i is a homomorphism.

Proof. First identity follows by direct calculation:

$$\begin{aligned} H_i(\Phi)\circ\pi(p) &= g_i\circ\Phi\circ s_i\circ\pi(p) \\ &= g_i\circ\Phi\Big(pg_i(p)^{-1}\Big) && \text{[By (1.3.7)]} \\ &= g_i\circ\Phi(p)g_i(p)^{-1} && \text{[Equivariance of } \Phi \text{ and } g_i] \\ &= \widetilde{H}_i(\Phi)(p). \end{aligned}$$

The second identity follows from the first and the fact that $\widetilde{H}_i$ is a homomorphism, since then we find

$$\begin{aligned} H_i(\Phi\circ\Psi)(x) &= H_i(\Phi\circ\Psi)\circ\pi\circ s_i(x) && [\pi\circ s_i = \mathrm{Id}_{U_i}] \\ &= \widetilde{H}_i(\Phi\circ\Psi)\Big(s_i(x)\Big) \\ &= \widetilde{H}_i(\Phi)\Big(s_i(x)\Big)\widetilde{H}_i(\Psi)\Big(s_i(x)\Big) \\ &= \Big(H_i(\Phi)\circ\pi\circ s_i(x)\Big)\Big(H_i(\Psi)\circ\pi\circ s_i(x)\Big) \\ &= \Big(H_i(\Phi)(x)\Big)\Big(H_i(\Psi)(x)\Big) \\ &= \Big(H_i(\Phi)H_i(\Psi)\Big)(p). \end{aligned}$$

□

Proposition 2.4.13. Let $\phi_i : U_i \to G$ be smooth. Then ϕ_i satisfies

$$\mathrm{Ad}_{g_{ij}(x)}\phi_j(x) = \phi_i(x) \tag{2.4.4}$$

for $x \in U_{ij}$ if and only if $\Phi(p) = p\ \mathrm{Ad}_{g_i(p)^{-1}}\phi_i\circ\pi(p)$ is a gauge transformation satisfying $H_i(\Phi) = \phi_i$.

Proof. Let ϕ satisfy $\mathrm{Ad}_{g_{ij}(x)}\phi_j(x) = \phi_i(x)$. Then $\widetilde{\phi}_i := \phi_i\circ\pi$ satisfies (2.4.2) and (2.4.3), where the first follows from the G-invariance of π, so by Lemma 2.4.10 we find that $\Phi(p) = \mathrm{Ad}_{g_i(p)^{-1}}\phi_i\circ\pi(p) =$ is a gauge transformation, such that $\widetilde{H}_i(\Phi) = \phi_i\circ\pi$. By Proposition 2.4.12, we find that the left-hand side equals $H_i(\Phi)\circ\pi$, so that if we let both sides act on $s_i(x)$, we find $H_i(\Phi)(x) = \phi_i(x)$.

Conversely, let Φ given by $p\mapsto p\ \mathrm{Ad}_{g_i(p)^{-1}}\phi_i\circ\pi(p)$. By Lemma 2.4.8 we find that $\Phi(p) = p\ \mathrm{Ad}_{g_i(p)^{-1}}\widetilde{H}_i(\Phi)(p)$, so if we compare both expressions, we find from the transitivity of the right action that $\phi_i\circ\pi = \widetilde{H}_i(\Phi)$. From Lemma (2.4.11) we know that $\widetilde{H}_i(\Phi)$, and so $\phi_i\circ\pi$, satisfies (2.4.3), and since $\pi(p) = x$, this implies $\mathrm{Ad}_{g_{ij}(x)}\phi_j(x) = \phi_i(x)$. □

The construction of the homomorphisms H_i helps us to show the existence of an isomorphism between $\mathcal{G}$ and $\Gamma^\infty(M, \mathrm{Ad}\ P)$. Before we can prove this, we have to show that $\Gamma^\infty(M, \mathrm{Ad}\ P)$ is a group.

Lemma 2.4.14. The function space $\Gamma^\infty(M, \mathrm{Ad}\ P)$ can be made into a group.

Proof. By Theorem 1.5.6, we know that $\pi_{\mathrm{Ad}}^{-1}(x) \subset \mathrm{Ad}\ P$ has a group structure given by $[(p,g)][(p,h)] = [(p,gh)]$. For any section $s \in \Gamma^\infty(M, \mathrm{Ad}\ P)$, we have by definition of a section that $s(x) \in \pi_{\mathrm{Ad}}^{-1}(x)$, so we can define multiplication of sections s_1 and s_2 by $s_1 s_2(x) = s_1(x)s_2(x)$, where in the right-hand side the multiplication on $\pi_{\mathrm{Ad}}^{-1}(x)$ is used. □

Remarks.

1. Let $s_i : U_i \to P$ be the local sections corresponding to the local trivializations ψ_i and let $f_i : U_i \to G$ be smooth. Then $s \in \Gamma^\infty(M, \mathrm{Ad}\ P)$ given by $s(x) = \Big[\Big(s_i(x), f_i(x)\Big)\Big]$ is well-defined if and only if f_i satisfies (2.4.4), since we necessarily have

$$\begin{aligned}
\Big[\Big(s_i(x), f_i(x)\Big)\Big] &= s(x) \\
&= \Big[\Big(s_j(x), f_j(x)\Big)\Big] \\
&= \Big[\Big(s_i(x)g_{ij}(x), f_j(x)\Big)\Big] \\
&= \Big[\Big(s_i(x), \mathrm{Ad}_{g_{ij}(x)} f_j(x)\Big)\Big]
\end{aligned}$$

2. The identity $e(x)$ of $\Gamma^\infty(M, \mathrm{Ad}\ P)$ is represented by $(s_i(x), e)$.

We can now prove next theorem.

Theorem 2.4.15. Let $H : \mathcal{G} \to \Gamma^\infty(M, \mathrm{Ad}\ P)$ be given by $H(\Phi)(x) = \Big[\Big(s_i(x), H_i(\Phi)(x)\Big)\Big]$, where $s_i : U_i \to P$ is the local section corresponding to the local trivialization $\psi_i : U_i \times G \to P$. Then H is an isomorphism of groups.

Proof. We first show that H is injective. Consider the set $H^{-1}\Big(e(x)\Big) = \{\Phi \in \mathcal{G} : H(\Phi)(x) = e(x)\}$. This is equal to $\Big\{\Phi \in \mathcal{G} : \Big[\Big(s_i(x), g_i \circ \Phi \circ s_i(x)\Big)\Big] = [(s_i(x), e)]\Big\}$, from which we find that $H^{-1}\Big(e(x)\Big) = \{\Phi \in G : g_i \circ \Phi \circ s_i(x) = e \text{ for all } i \in I\}$. By equation (1.3.5) we find that $g_i^{-1}(e) = s_i(x)$, whence $H^{-1}\Big(e(x)\Big) = \{\Phi \in \mathcal{G} : \Phi \circ s_i(x) = s_i(x) \text{ for all } i \in I\} = \{\mathrm{Id}\}$.

For surjectivity we first remark that by Proposition 1.5.10 every section $s \in \Gamma^\infty(M, \mathrm{Ad}\ P)$ can be written as $\Big[\Big(s_i(x), f_i(x)\Big)\Big]$ for some function $f_i : U_i \to G$ satisfying $f_i(x) = \mathrm{Ad}_{g_{ij}(x)} f_j(x)$. Hence by Proposition 2.4.13 we find that there is a gauge transformation F, such that $H_i(F) = f_i$. With other words, we find $s(x) = \Big[\Big(s_i(x), H_i(F)(x)\Big)\Big]$, which is $H(F)(x)$, so H is surjective. Finally, by

$$\begin{aligned}
H(\Psi)H(\Phi)(x) &= H(\Psi)(x)H(\Phi)(x) \\
&= \Big[\Big(s_i(x), H_i(\Psi)(x)\Big)\Big]\Big[\Big(s_i(x), H_i(\Phi)(x)\Big)\Big] \\
&= \Big[\Big(s_i(x), H_i(\Psi)(x)H_i(\Phi)(x)\Big)\Big] \\
&= \Big[\Big(s_i(x), H_i(\Psi \circ \Phi)(x)\Big)\Big] \\
&= H(\Psi \circ \Phi)(x),
\end{aligned}$$

we find that H is a homomorphism. $\square$

The center of the gauge group

Definition 2.4.16. Fix a $g \in G$. Then we define $\Phi_g : P \to P$ and $\widehat{\phi}_g : P \to G$ by $\Phi_g = R_g$ and $\widehat{\phi}_g(p) = g$ for all $p \in P$.

The following proposition related the center of $\mathcal{G}$ with the center of G.

Lemma 2.4.17. Φ_g is a gauge transformation if and only if $g \in Z(G)$. If Φ_g is a gauge transformation, $\widehat{\phi}_g$ is the corresponding element in $C^\infty_{\mathrm{Ad}}(P, G)$.

Proof. By Proposition 1.1.7 we have $\pi \circ \Phi_g = \pi$. Let $g \in G$. Since we have $R_g \circ R_a = R_{ga} = R_{ag} = R_a \circ R_g$ for all $a \in G$ if and only if $g \in Z(G)$, we find

$$\Phi_g \circ R_a = R_a \circ \Phi_g$$

if and only if $g \in Z(G)$. So Φ_g satisfies the equivariance axiom and is therefore a gauge transformation if and only if $g \in Z(G)$. By Theorem 2.4.7, we find that $\widehat{\phi}_g$ is the corresponding element in $C^\infty_{\mathrm{Ad}}(P, G)$. □

Proposition 2.4.18. We have $\Phi_g \in Z(\mathcal{G})$ if $g \in Z(G)$. Equivalently, if we identify $\mathcal{G}$ with $C^\infty_{\mathrm{Ad}}(P, G)$, we have $\widehat{\phi}_g \in Z(\mathcal{G})$ if $g \in Z(G)$.

Proof. Let $g \in Z(G)$. Let $\Phi \in \mathcal{G}$ be an arbitrary other gauge transformation. Then we have by the equivariance of Φ

$$\Phi_g \circ \Phi(p) = \Phi(p)g = \Phi(pg) = \Phi \circ \Phi_g(p),$$

thus $\Phi_g \in Z(\mathcal{G})$. By Theorem 2.4.7 it follows that $\widehat{\phi}_g \in Z(\mathcal{G})$ if we identify by the same theorem $\mathcal{G}$ with $C^\infty_{\mathrm{Ad}}(P, G)$. □

We see that $\{\Phi_g : g \in Z(G)\} \subset Z(\mathcal{G})$. Next theorem shows that under certain circumstances we have equality.

Theorem 2.4.19. Let G be a connected compact matrix group with semisimple Lie algebra $\mathfrak{g}$ and let M be connected. Then we have the following isomorphisms:

$$\begin{aligned} Z(\mathcal{G}) &= Z(G) \\ &= \Gamma^\infty(M, P \times_{\mathrm{Ad}} Z(G)) \\ &= \{\Phi_g : g \in Z(G)\} \\ &= \{\widehat{\phi}_g : g \in Z(G)\}. \end{aligned}$$

Proof. The second equality follows from Theorem 2.4.15 and the identity $\mathrm{Ad}\ P = P \times_{\mathrm{Ad}} G$. We prove the first equality by proving the equality $Z(G) = \Gamma^\infty(M, P \times_{\mathrm{Ad}} Z(G))$. Let $\phi_i \in C^\infty(U_i, G)$ be a family representing a section $s \in \Gamma^\infty(M, P \times_{\mathrm{Ad}} Z(G))$. Since all $\phi_i(x) \in Z(G)$, we have $\phi_i(x) = \mathrm{Ad}_{g_{ij}(x)}\phi_j(x) = \phi_j(x)$. So the ϕ_i define a global function $\phi : M \to Z(G)$. Since we assumed that M is connected, we find by Proposition 1.4.20 that ϕ must be constant. So we find that ϕ does not depend on its argument, thus we have $\phi \in Z(G)$.

Conversely, for $g \in Z(G)$ define $\phi \equiv g$ be constant and let $\phi_i = \phi|_{U_i}$. Then since $g \in Z(G)$, we have $\phi_i(x) = \phi_j(x) = \mathrm{Ad}_{g_{ij}(x)}\phi_j(x)$, which defines a unique element $s \in \Gamma^\infty(M, P \times_{\mathrm{Ad}} Z(G))$.

We prove the third equality by showing that if $\Phi \in Z(\mathcal{G})$, we must have $\Phi = \Phi_g$ for some $g \in Z(G)$. So let $\Phi \in Z(\mathcal{G})$. By Proposition 2.4.13 we find maps $\phi_i : U_i \to G$ such that $\Phi(p) = p\ \mathrm{Ad}_{g_i(p)^{-1}}\phi_i \circ \pi(p)$ and $H_i(\Phi) = \phi_i$. By the isomorphism H in Theorem 2.4.15, we see that these ϕ_i define an element $H(\Phi) \in \Gamma^\infty(M, \mathrm{Ad}\ P)$ and since we assumed that $\Phi \in Z(\mathcal{G})$, we see that $H(\Phi) \in \Gamma^\infty(M, P \times_{\mathrm{Ad}} G)$. Hence we see by the proof of first equality that the ϕ_i must be constant; $\phi_i \equiv g$ with $g \in Z(G)$. Then by $\Phi(p) = p\mathrm{Ad}_{g_i(p)^{-1}}\phi_i \circ \pi(p)$, we see $\Phi(p) = pg$ for all $p \in P$.

Last equality simply follows from the third. □

The action of the gauge group on the connection, curvature and field-strength

The gauge group acts in a natural way on connections in the following way. Let HP be a connection and $\Phi : P \to P$ be a gauge transformation. We can let Φ act on TP by $X \mapsto \Phi_*^{-1}X$. In particular, this defines an action of Φ on HP. We denote the gauge transformed horizontal space by $H_p^\Phi P := \Phi_*^{-1} H_{\Phi(p)}P$. The corresponding projection operator is given by $H_p^\Phi = \Phi_*^{-1} H_{\Phi(p)} \Phi_*$.

Lemma 2.4.20. $H_p^\Phi : T_pP \to T_pP$ is a projection operator with image $H_p^\Phi P$.

Proof. First we show that H_p^Φ is a projection operator:

$$(H_p^\Phi)^2 = \Phi_*^{-1} H_{\Phi(p)} \Phi_* \Phi_*^{-1} H_{\Phi(p)} \Phi_* = \Phi_*^{-1} H_{\Phi(p)} H_{\Phi(p)} \Phi_* = \Phi_*^{-1} H_{\Phi(p)} \Phi_* = H_p^\Phi.$$

Secondly, we have

$$\begin{aligned}
\text{im } H_p^\Phi &= \Phi_*^{-1} H_{\Phi(p)} \Phi_* T_p P \\
&= \Phi_*^{-1} H_{\Phi(p)} T_{\Phi(p)} P && [\Phi \text{ is a diffeomorphism}] \\
&= \Phi_*^{-1} H_{\Phi(p)} P \\
&= H_p^\Phi P.
\end{aligned}$$

□

Before we check that $H_p^\Phi P$ is indeed a connection, we will show that the vertical space is invariant under the action of Φ.

Lemma 2.4.21. We have $\Phi_* \sigma(A) = \sigma(A)$ for every $\Phi \in \mathcal{G}$ and $A \in \mathfrak{g}$.

Proof. By direct calculation

$$\begin{aligned}
\Phi_* \sigma(A)|_p &= \Phi_* \left(\left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} p e^{tA} \right) \\
&= \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} \Phi\left(p e^{tA} \right) \\
&= \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} \Phi(p) e^{tA} && [\text{Equivariance of } \Phi] \\
&= \sigma(A)|_{\Phi(p)}
\end{aligned}$$

□

Proposition 2.4.22. $H^\Phi P$ is connection on P.

Proof. We prove this by checking that H^Φ satisfies the properties in Proposition 2.1.10. Firstly, by X is vertical $\Leftrightarrow$ $\Phi_*(X)$ is vertical $\Leftrightarrow$ $H\Phi_*(X) = 0 \Leftrightarrow H^\Phi(X) = \Phi_*^{-1} H \Phi_*(X) = 0$, we see that $\ker H^\Phi = V_p P$. Furthermore, since H^Φ is composed of smooth functions, it is smooth. Finally, the equivariance condition for Φ can be translated into $\Phi \circ R_g = R_g \circ \Phi$, whence $R_{g*} \circ \Phi_* = \Phi_* \circ R_{g*}$, which leads to

$$R_{g*} H_p^\Phi = R_{g*} \Phi_*^{-1} H_p \Phi_* = \Phi_*^{-1} R_{g*} H_p \Phi_* = \Phi_*^{-1} H_{pg} R_{g*} \Phi_* = \Phi_*^{-1} H_{pg} \Phi_* R_{g*} = H_{pg}^\Phi R_{g*},$$

which concludes our proof. □

Given the action of the gauge group on a connection, we can derive the induced action on the corresponding one-form.

Proposition 2.4.23. Let $\omega^\Phi = \Phi^* \omega$ be the result of the action of Φ on ω. Then ω^Φ is the connection one-form, which corresponds with the Ehresmann connection $H^\Phi P$.

Proof. Let $H^\Phi = \Phi_*^{-1} H \Phi_*$ be the projection operator onto the gauge transformed horizontal subspace. The corresponding connection one-form is given by $\omega^\Phi = \sigma^{-1} \circ (\mathrm{Id} - H^\Phi)$. We shall show that $\omega^\Phi = \Phi^*\omega$ by direct calculation.

$$\begin{aligned}
\omega^\Phi &= \sigma^{-1}(\mathrm{Id} - \Phi_*^{-1}H\Phi_*) \\
&= \sigma^{-1}\circ\Phi_*^{-1}(\mathrm{Id} - H)\Phi_* \\
&= (\Phi_*\circ\sigma)^{-1}(\mathrm{Id} - H)\Phi_* \\
&= \sigma^{-1}(\mathrm{Id} - H)\Phi_* && \text{[By Lemma 2.4.21]} \\
&= \omega\Phi_* \\
&= \Phi^*\omega.
\end{aligned}$$

□

Proposition 2.4.24. Let $\Phi : P \to P$ be a gauge transformation acting on a connection one-form ω. Then the gauge potential $\mathcal{A}_i^\Phi$ corresponding to the gauge transformed connection ω^Φ satisfies

$$\begin{aligned}
\mathcal{A}_i^\Phi|_x &= \mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{A}_i|_x + \phi_i^*\theta|_{\phi_i(x)} \\
&= \phi_i(x)^{-1}\mathcal{A}_i|_x\phi_i(x) + \phi_i(x)^{-1}\mathrm{d}\phi_i|_x, \qquad (2.4.5)
\end{aligned}$$

where $\phi_i = H_i(\Phi)$ as defined in Proposition 2.4.12.

Proof. We first note that the connection one-forms ω and ω^Φ are given at p by

$$\omega|_p = \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + g_i^*\theta|_{g_i(p)} \qquad (2.4.6)$$

$$\omega^\Phi|_p = \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\mathcal{A}_i^\Phi|_{\pi(p)} + g_i^*\theta|_{g_i(p)}, \qquad (2.4.7)$$

with $x = \pi(p)$. We have $\omega^\Phi|_p = \Phi^*\omega|_q$ with $q = \Phi(p)$, from which we can deduce a relation between $\mathcal{A}_i$ and $\mathcal{A}_i^\Phi$. First, notice that

$$\widetilde{\phi}_i(p) = g_i(\Phi(p))g_i(p)^{-1} = g_i(q)g_i(p)^{-1},$$

implying $g_i(q)^{-1} = g_i(p)^{-1}\widetilde{\phi}_i(p)^{-1}$. Then we find

$$\begin{aligned}
\omega^\Phi|_p &= \Phi^*\omega|_{\Phi(p)} \\
&= \mathrm{ad}_{g_i(q)^{-1}}\circ\Phi^*\pi^*\mathcal{A}_i|_{\pi\circ\Phi(p)} + \Phi^*g_i^*\theta|_{g_i\circ\Phi(p)} \\
&= \mathrm{ad}_{g_i(q)^{-1}}\circ(\pi\circ\Phi)^*\mathcal{A}_i|_{\pi\circ\Phi(p)} + (g_i\circ\Phi)^*\theta|_{g_i\circ\Phi(p)} \\
&= \mathrm{ad}_{g_i(q)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + (g_i\circ\Phi)^*\theta|_{g_i\circ\Phi(p)} && [\pi\circ\Phi = \pi] \\
&= \mathrm{ad}_{g_i(p)^{-1}}\mathrm{ad}_{\widetilde{\phi}_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + (g_i\circ\Phi)^*\theta|_{g_i\circ\Phi(p)},
\end{aligned}$$

where we used in the third equality that the pullback is a contravariant functor. Now, since $(g_i\circ\Phi)(p) = g_i(q) = \widetilde{\phi}_i(p)g_i(p)$, we find

$$\begin{aligned}
(g_i\circ\Phi)^*\theta|_{g_i\circ\Phi(p)} &= g_i(p)^{-1}\widetilde{\phi}_i(p)^{-1}\mathrm{d}\Big(\widetilde{\phi}_i(p)g_i(p)\Big) && [f^*\theta|_{f(p)} = f(p)^{-1}\mathrm{d}f(p)] \\
&= g_i(p)^{-1}\mathrm{d}g_i(p) + g_i(p)^{-1}\widetilde{\phi}_i(p)^{-1}\mathrm{d}\Big(\widetilde{\phi}_i(p)\Big)g_i(p) \\
&= g_i^*\theta|_{g_i(p)} + \mathrm{ad}_{g_i(p)^{-1}}\circ\widetilde{\phi}_i^*\theta|_{\widetilde{\phi}_i(p)} \\
&= g_i^*\theta|_{g_i(p)} + \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\phi_i^*\theta|_{\phi_i\circ\pi(p)} && [\widetilde{\phi}_i = \phi_i\circ\pi].
\end{aligned}$$

Collecting these results, we find

$$\begin{aligned}
\omega^\Phi|_p &= \mathrm{ad}_{g_i(p)^{-1}}\mathrm{ad}_{\widetilde{\phi}_i(p)^{-1}}\circ\pi^*\mathcal{A}_i|_{\pi(p)} + \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\phi_i^*\theta|_{\phi_i\circ\pi(p)} + g_i^*\theta|_{g_i(p)} \\
&= \mathrm{ad}_{g_i(p)^{-1}}\circ\pi^*\Big(\mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{A}_i|_{\pi(p)} + \phi_i^*\theta|_{\phi_i\circ\pi(p)}\Big) + g_i^*\theta|_{g_i(p)} && [\widetilde{\phi}_i(p) = \phi_i(x)].
\end{aligned}$$

Comparing this result with (2.4.7) gives the desired expression for $\mathcal{A}_i^\Phi$. □

We can also derive how the curvature and the gauge field-strength transform under a gauge transformation

Proposition 2.4.25. Let $\Phi : P \to P$ be a gauge transformation acting on a connection HP. Then the curvature Ω^Φ and gauge-field strength $\mathcal{F}_i^\Phi$ corresponding to the gauge transformed connection $H^\Phi P$ satisfy

$$\Omega^\Phi = \Phi^*\Omega, \tag{2.4.8}$$

$$\mathcal{F}_i^\Phi|_x = \mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{F}_i|_x, \tag{2.4.9}$$

where $\phi_i = H_i(\Phi)$ is defined as in Proposition 2.4.12.

Proof. By definition, we have $\Omega = \mathrm{d}_\omega\omega = H^\vee \mathrm{d}\omega$. So we have $\Omega(X,Y) = \mathrm{d}\omega(HX, HY)$, from which we find a formula for Ω^Φ by replacing ω by ω^Φ and H by H^Φ. Hence, we find

$$\begin{aligned}
\Omega^\Phi(X,Y) &= \mathrm{d}\omega^\Phi(H^\Phi X, H^\Phi Y)\\
&= \mathrm{d}\Phi^*\omega(H^\Phi X, H^\Phi Y)\\
&= \Phi^*\mathrm{d}\omega(H^\Phi X, H^\Phi Y)\\
&= \Phi^*\mathrm{d}\omega(\Phi_*^{-1}H\Phi_* X, \Phi_*^{-1}H\Phi_* Y)\\
&= \Phi^* H^\vee(\Phi^{-1})^*\Phi^*\mathrm{d}\omega(X,Y)\\
&= \Phi^* H^\vee \mathrm{d}\omega(X,Y)\\
&= \Phi^*\Omega(X,Y).
\end{aligned}$$

The proof of the second identity is similar to the proof of Proposition 2.2.7, but with $\mathcal{F}_i$ replaced by $\mathcal{F}_i^\Phi$, $g_{ji} = g_{ij}^{-1}$ replaced by ϕ_i and $\mathcal{F}_j$ replaced by $\mathcal{F}_i$. $\square$

Convention. Instead of ω^Φ, $\mathcal{A}_i^\Phi$ etcetera, we shall sometimes write $\Phi(\omega)$, $\Phi(\mathcal{A}_i)$ etcetera.

Next theorem summarizes all results about gauge transformations acting on other objects.

Theorem 2.4.26. Let $\{H_pP : p \in P\}$ be a connection and let H, ω, $\mathcal{A}_i$, Ω and $\mathcal{F}_i$ be the corresponding projection operator, connection one-form, gauge field, curvature respectively gauge field-strength, then a gauge transformation Φ acts on these object in the following way:

1. $\Phi(H_pP) = \Phi_*^{-1}H_{\Phi(p)}P$;
2. $\Phi(H)_p = \Phi_*^{-1}H_{\Phi(p)}\Phi_*$;
3. $\Phi(\omega) = \Phi^*\omega$;
4. $\Phi(\mathcal{A}_i)|_x = \mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{A}_i|_x + \phi_i^*\theta|_{\phi_i(x)}$;
5. $\Phi(\Omega) = \Phi^*\Omega$;
6. $\Phi(\mathcal{F}_i)|_x = \mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{F}_i|_x$,

where $\phi_i = H_i(\Phi)$ is defined as in Proposition 2.4.12.

2.5 Parallel transport and holonomy

Definition 2.5.1. Let $P(M,G)$ be a principal bundle with connection HP and let $\gamma : [0,1] \to M$ be a piecewise smooth curve in M. A piecewise smooth curve $\tilde{\gamma} : [0,1] \to P$ is called *horizontal* if $\tilde{\gamma}'(t) \in H_{\tilde{\gamma}(t)}P$. The curve $\tilde{\gamma}$ is called a *lift* of γ if $\pi \circ \tilde{\gamma} = \gamma$. If $\tilde{\gamma}$ is both a lift of γ and horizontal, we call $\tilde{\gamma}$ a *horizontal lift* of γ.

Notice that this definition implies that $\omega\Big(\tilde{\gamma}'(t)\Big) = 0$ for all $t \in [0,1]$. A full proof of next theorem can be found in [30] as Proposition 8.7.

Theorem 2.5.2. Let $\gamma : [0,1] \to M$ be a piecewise smooth curve and let $p_0 \in \pi^{-1}\Big(\gamma(0)\Big)$. Then there exists a unique horizontal lift $\tilde{\gamma}(t)$ in P such that $\tilde{\gamma}(0) = p_0$.

Proof (sketch). The lift may not be smooth, but piecewise smooth, we can restrict us to a smooth curve γ in a coordinate neighborhood U_i. A global piecewise smooth lift can then be obtained by gluing the curves in overlapping the coordinate neighborhoods together. So assume $\gamma : [0,1] \to U_i$ and let ψ_i be a local trivialization corresponding to U_i. Then we see that $\tilde{\gamma}$ has to satisfy $\psi_i(\gamma(t), \alpha(t)) = \tilde{\gamma}(t)$ for some $\alpha : [0,1] \to G$. Since $\tilde{\gamma}(0) = p_0$, we find that $\alpha(0)$ has to satisfy $\psi_i(x_0, \alpha(0)) = p_0$, where $x_0 = \gamma(0) = \pi(p_0)$. Since $\tilde{\gamma}'(t)$ has to be horizontal, we must have $\omega(\tilde{\gamma}'(t)) = 0$. One can find that this is equivalent with $\mathrm{ad}_{\alpha(t)^{-1}}\mathcal{A}_i(\gamma'(t)) + (\alpha^*\theta)(\gamma'(0)) = 0$, which is in matrix form $\alpha'(t) + \mathcal{A}_i(\gamma'(0))\alpha(t) = 0$. This is a first-order ordinary differential equation for α, which has a unique solution given the initial condition $\psi_i(x_0, \alpha(0)) = p_0$. ☐

Lemma 2.5.3. Let $\tilde{\gamma}$ be the horizontal lift of $\gamma : [0,1] \to M$ with $\tilde{\gamma}(0) = p$. Then if $\bar{\gamma}$ is another horizontal lift of γ with $\bar{\gamma}(0) = pg$ for some $g \in G$, than we have $\bar{\gamma} = R_g\tilde{\gamma}$.

Proof. Using Proposition 1.1.7, we see that $\pi\Big(R_g \circ \tilde{\gamma}(t)\Big) = \pi\Big(\tilde{\gamma}(t)\Big) = \gamma(t)$, so $R_g \circ \tilde{\gamma}$ is a lift of $\gamma(t)$. Furthermore, we have $(R_g \circ \tilde{\gamma})'(t) = R_{g*}\tilde{\gamma} \in H_{\tilde{\gamma}(t)g}P$, thus $R_g \circ \tilde{\gamma}$ is a horizontal lift. By previous theorem it follows that $R_g \circ \tilde{\gamma}$ must be equal to $\bar{\gamma}$. ☐

Definition 2.5.4. Let $p \in P$ and let $\gamma : [0,1] \to M$ be a smooth piecewise curve that is closed. Thus we have $\gamma(0) = \pi(p)$. Then we define the *parallel transport* $\tau_\gamma : \pi^{-1}\Big(\gamma(0)\Big) \to \pi^{-1}\Big(\gamma(1)\Big)$ of p along γ by $\tau_\gamma(p) = \tilde{\gamma}(1)$, where $\tilde{\gamma}$ is the horizontal lift of γ with starting point $\tilde{\gamma}(0) = p$.

Proposition 2.5.5. Parallel transport is equivariant: $R_g \circ \tau_\gamma = \tau_\gamma \circ R_g$ for all $g \in G$ and all curves $\gamma : [0,1] \to M$.

Proof. We have $R_g \circ \tau_\gamma(p) = \tilde{\gamma}(1)g$. Let $\bar{\gamma}$ be given by $R_g \circ \tilde{\gamma}$. Then $\bar{\gamma}$ is the horizontal lift of γ with starting point pg, so $\tau_\gamma(pg) = \bar{\gamma}(1)$. But this is also precisely $\tau_\gamma \circ R_g(p) = \tilde{\gamma}(1)g$. ☐

Notice that if γ is a closed curve such that $\gamma(0) = x_0$ with horizontal lift $\tilde{\gamma}$, we have $\tilde{\gamma}(0), \tilde{\gamma}(1) \in \pi^{-1}(x_0)$. In other words, for closed curves with $\gamma(0) = \pi(p)$ we have $\tau_\gamma(p) = pg$ for some $g \in G$. Notice that this g is unique, since the action is free.

Definition 2.5.6. Let $\gamma_1, \gamma_2 : [0,1] \to M$ be piecewise smooth closed curves. Then we define the path product $\gamma_1 * \gamma_2$ by

$$\gamma_1 * \gamma_2(t) = \begin{cases} \gamma_1(2t), & 0 \le t \le \frac{1}{2}; \\ \gamma_2(2t-1), & \frac{1}{2} \le t \le 1. \end{cases}$$

It follows directly from the definition that $\gamma_1 * \gamma_2$ is a piecewise smooth closed curve.

Proposition 2.5.7. Let $\gamma_1, \gamma_2 : [0,1] \to M$ be piecewise smooth closed curves. Then if $\gamma = \gamma_1 * \gamma_2$ is the path product of γ_1 and γ_2, we have $\tau_\gamma = \tau_{\gamma_2} \circ \tau_{\gamma_1}$.

Proof. Let $g_1 \in G$ be the element such that $\tau_{\gamma_1}(p) = \tilde{\gamma}_1(1) = pg_1$. Then a horizontal lift $\tilde{\gamma}$ of γ is given by

$$\tilde{\gamma}(t) = \begin{cases} \tilde{\gamma}_1(2t), & 0 \leq t \leq \frac{1}{2}; \\ \tilde{\gamma}_2(2t-1)g_1, & \frac{1}{2} \leq t \leq 1, \end{cases}$$

where $\tilde{\gamma}_i$ is the horizontal lift of γ_i. Note that the factor g_1 is essential to make the lift continuous at $t = \frac{1}{2}$. Then by previous proposition we find $\tau_\gamma(p) = \tilde{\gamma}(1) = \tau_{\gamma_2}(p)g_1 = \tau_{\gamma_2}(pg_1) = \tau_{\gamma_2} \circ \tau_{\gamma_1}(p)$. □

A direct consequence of this proposition is that if there are $g_1, g_2 \in G$ and piecewise smooth closed curves $\gamma_1, \gamma_2 : [0,1] \to M$ such that $\tau_{\gamma_i}(p) = pg_i$, we can find a curve $\gamma : [0,1] \to M$ such that $\tau_\gamma(p) = pg_2g_1$. This is exactly the curve $\gamma = \gamma_1 * \gamma_2$. This shows that the holonomy group defined directly below is a group, since it is a subset of G closed under multiplication.

Definition 2.5.8. Let $C_p(M)$ be the group of piecewise smooth closed curves $\gamma : [0,1] \to M$ such that $\gamma(0) = \gamma(1) = \pi(p)$. Then we define the *holonomy group* of a connection ω at the point p as $\mathrm{Hol}_p(\omega) = \{g \in G : \tau_\gamma(p) = pg \text{ for } \gamma \in C_p(M)\}$. Let $C_p^0(M)$ be the subgroup of $C_p(M)$ of curves homotopic to 0, then we define *restricted holonomy group* of ω at p as $\mathrm{Hol}_p^0(\omega) = \{g \in G : \tau_\gamma(p) = pg \text{ for } \gamma \in C_p^0(M)\}$.

Proposition 2.5.9.

1. We have $\mathrm{Hol}_{pa}(\omega) = a^{-1}\mathrm{Hol}_p(\omega)a$.

2. Let $p, q \in P$ be connected to each other by a horizontal curve $\tilde{\gamma}$ (not necessarily a lift). Thus we have $\tilde{\gamma}(0) = p$ and $\tilde{\gamma}(1) = q$. Then $\mathrm{Hol}_p(\omega) \simeq \mathrm{Hol}_q(\omega)$.

Proof.

1. Let $\gamma : [0,1] \to M$ be a curve such that $\tau_\gamma(p) = pg$. By Proposition 2.5.5 we see that this is equivalent to $\tau_\gamma(pa) = \tau_\gamma(p)a = pga = pa(a^{-1}ga)$. Hence we find that $g \in \mathrm{Hol}_p(\omega)$ if and only if $a^{-1}ga \in \mathrm{Hol}_{pa}(\omega)$, which is equivalent with stating that $\mathrm{Hol}_{pa}(\omega) = a^{-1}\mathrm{Hol}_p(\omega)a$.

2. Let $p \sim q$ denote the equivalence relation that there is a horizontal curve $\tilde{\gamma}$ that connects p and q, thus we have $\tilde{\gamma}(0) = p$ and $\tilde{\gamma}(1) = q$. It follows immediately that $\sim$ is an equivalence relation. Notice that $g \in \mathrm{Hol}_p(\omega)$ if and only if $p \sim pg$. Now, since $\sim$ is an equivalence relation, it is transitive. So if $p \sim q$, we have $p \sim pg$ if and only if $q \sim qg$. So we see that if $p \sim q$, we have $\mathrm{Hol}_p(\omega) = \mathrm{Hol}_q(\omega)$.

□

Corollary 2.5.10. Let M be connected. Then for all point $p, q \in P$ we have $\mathrm{Hol}_p(\omega) = \mathrm{Hol}_q(\omega)$.

Proof. Since M is connected and all manifolds are locally path-connected, it follows that M is path-connected. So there is a path $\gamma : [0,1] \to M$ such that $\gamma(0) = \pi(p)$ and $\gamma(1) = \pi(q)$. Then let $\tilde{\gamma}$ be the lift of γ starting in p. Since $\tilde{\gamma}$ ends in $\pi^{-1}(\pi(q))$, we find that there is an $a \in G$ such that $\tilde{\gamma}(1) = qa$. With other words, $p \sim qa$, where $\sim$ is the equivalence relation introduced in the proof of previous proposition. Hence we find by the second part of previous proposition that $\mathrm{Hol}_p(\omega) = \mathrm{Hol}_{qa}(\omega)$. By the first part of the proposition, we see that $\mathrm{Hol}_{qa}(\omega)$ and $\mathrm{Hol}_q(\omega)$ are conjugated in G, thus isomorphic, whence $\mathrm{Hol}_p(\omega) = \mathrm{Hol}_q(\omega)$. □

Convention. For connected M we are allowed by previous corollary to drop the subscript p and write $\mathrm{Hol}(\omega)$.

Definition 2.5.11. Let M be connected. Then a connection is said to be *irreducible* if the holonomy group is exactly G. The space of irreducible connections is denoted with $\mathscr{A}_o$.

Next proposition relates horizontal lifts of paths to the horizontal lifts with respect to a gauge transformed connection.

Proposition 2.5.12. Let $\tilde{\gamma}$ the lift of $\gamma : [0,1] \to M$ with respect to the connection HP with starting point $\tilde{\gamma}(0) = p$ and let Φ be a gauge transformation. Then the lift of γ with respect to the connection $H^\Phi P$ is given by $\tilde{\gamma}^\Phi = \Phi^{-1} \circ \tilde{\gamma}$ and has starting point $\Phi^{-1}(p)$.

Proof. Since Φ^{-1} is a gauge transformation, we have $\pi \circ \Phi^{-1} = \pi$, hence we find $\pi \circ \tilde{\gamma}^\Phi = \pi \circ \tilde{\gamma} = \gamma$. Since $\tilde{\gamma}$ is horizontal with respect to HP, we have $\omega(\tilde{\gamma}'(t)) = 0$. Then we find

$$\begin{aligned} \omega^\Phi\Big((\tilde{\gamma}^\Phi)'(t)\Big) &= \Phi^*\omega\Big((\Phi^{-1} \circ \gamma)'(t)\Big) && \text{[By Theorem 2.4.26]} \\ &= \Phi^*\omega\Big(\Phi_*^{-1}\tilde{\gamma}'(t)\Big) && \text{[By definition of the pushforward]} \\ &= \omega\Big(\tilde{\gamma}'(t)\Big) \\ &= 0, \end{aligned}$$

so $\tilde{\gamma}^\Phi$ is horizontal with respect to $H^\Phi P$. Finally, we have $\tilde{\gamma}^\Phi(0) = \Phi^{-1} \circ \tilde{\gamma}(0) = \Phi^{-1}(p)$. $\square$

Proposition 2.5.13. Let Φ be a gauge transformation and let τ_γ^Φ denote the parallel transport along the curve $\gamma : [0,1] \to M$ with respect to the connection $H^\Phi P$. Then we have

$$\Phi \circ \tau_\gamma = \tau_\gamma^{\Phi^{-1}} \circ \Phi. \tag{2.5.1}$$

Proof. By previous proposition we have that $\tilde{\gamma}^\Phi$ starts in $\Phi^{-1}(p)$ if $\tilde{\gamma}$ starts in p. Then we find that τ_γ^Φ acting on $\Phi^{-1}(p)$ is given by $\tilde{\gamma}^\Phi(1)$, just like τ_γ acting on p is given by $\tilde{\gamma}(1)$. Then we find

$$\begin{aligned} \tau_\gamma^\Phi \circ \Phi^{-1}(p) &= \tilde{\gamma}^\Phi(1) \\ &= \Phi^{-1} \circ \tilde{\gamma}(1) \\ &= \Phi^{-1} \circ \tau_\gamma(p), \end{aligned}$$

and the statement follows by interchanging Φ and Φ^{-1}. $\square$

Corollary 2.5.14. Let Φ be a gauge transformation that leaves a connection ω invariant: $\Phi(\omega) = \omega$. Then we have $\tau_\gamma^\Phi = \tau_\gamma$ and in particular for all loops γ,

$$\Phi \circ \tau_\gamma = \tau_\gamma \circ \Phi. \tag{2.5.2}$$

Definition 2.5.15. We denote the isotropy group of a connection ω, $\{\Phi \in \mathcal{G} : \Phi(\omega) = \omega\}$ by Γ_ω.

Proposition 2.5.16. Let ω be an irreducible connection and let G be a connected and compact matrix group with semisimple Lie algebra $\mathfrak{g}$. Then $\Gamma_\omega = Z(G)$.

Proof. Write $\Phi \in \Gamma_\omega$ as $\Phi(p) = p\widehat{\phi}(p)$ with $\widehat{\phi} \in C^\infty_{\mathrm{Ad}}(P, G)$. Then equation (2.5.2) is equivalent with

$$\tau_\gamma(p)\widehat{\phi} \circ \tau_\gamma(p) = \tau_\gamma(p)\widehat{\phi}(p)$$

for all closed piecewise smooth curves $\gamma : [0,1] \to M$, where we used the equivariance of τ_γ proved in Proposition 2.5.5 in order to obtain the right-hand side. Since the right action is free, we find that $\widehat{\phi} \circ \tau_\gamma = \widehat{\phi}$. If we fix a $p \in P$, we have $\tau_\gamma(p) = pg$ for some $g \in G$, so that we obtain $\widehat{\phi}(pg) = \widehat{\phi}(p)$. By Definition 2.4.3 we find that this implies

$$g^{-1}\widehat{\phi}(p)g = \widehat{\phi}(p). \tag{2.5.3}$$

Now, since ω is irreducible, we see that we can find paths $\gamma : [0,1] \to M$ such that (2.5.3 holds for all $g \in G$. Hence $\widehat{\phi}(p) \in Z(G)$. By Proposition 1.4.20 $Z(G)$ is discrete, so $\widehat{\phi}$ must be the constant function $p \mapsto a \in Z(G)$. So $\Phi = R_a$ with $a \in Z(G)$. Notice that Φ is indeed a gauge transformation by Lemma 2.4.17. So we see that $Z(G) \mapsto \Gamma_\omega$ given by $a \mapsto R_a$ is a surjection, and since this map is clearly injective, it is a bijection. □

Proposition 2.5.17. Let M be connected and let G be a connected compact matrix Lie group with semisimple Lie algebra $\mathfrak{g}$. Then the group $\tilde{\mathcal{G}} = \mathcal{G}/Z(\mathcal{G})$ acts in a free way on the space of irreducible connections $\mathscr{A}^+$.

Proof. Let ω be an irreducible connection and let $\Phi \in \mathcal{G}$ be a gauge transformation that leaves ω invariant. Then since ω is irreducible, we have $\mathrm{Hol}(\omega) = G$, so for all $p \in P$ and all $g \in G$ there is a closed path $\gamma : [0,1] \to M$ such that $\tau_\gamma(p) = pg$. Write $\Phi(p) = p\widehat{\phi}(p)$ with $\widehat{\phi} \in C^\infty_{\mathrm{Ad}}(P,G)$. Then for all $p \in P$ we find

$$\begin{aligned}
p\widehat{\phi}(p)g &= \Phi(p)g \\
&= \Phi(pg) && \text{[By the equivariance of } \Phi] \\
&= \Phi \circ \tau_\gamma(p) \\
&= \tau_\gamma \circ \Phi(p) && \text{[By Corollary 2.5.14]} \\
&= \tau_\gamma\Big(p\widehat{\phi}(p)\Big) \\
&= \tau_\gamma(p)\widehat{\phi}(p) && \text{[By the equivariance of } \tau_\gamma] \\
&= pg\widehat{\phi}(p),
\end{aligned}$$

so since the right action on P is free, we find that $g\widehat{\phi}(p) = \widehat{\phi}(p)g$ for all $g \in G$. Hence we see that $\widehat{\phi}(p) \in Z(G)$ for all $p \in P$. Now, by Proposition 1.4.20, we find that $Z(G)$ is discrete, so we must have that $\widehat{\phi}$ is constant. Then by Theorem 1.4.20 we see that $\Phi \in Z(\mathcal{G})$, which concludes the proof. □

Finally, we will state the Ambrose-Singer Theorem, which we will not prove, but refer to Theorem II.8.1 of [22] instead.

Theorem 2.5.18 (Ambrose-Singer Theorem)**.** Let $P(M,G)$ be a principal bundle with M connected. Let ω be a connection one-form on P. Then the Lie algebra $\mathrm{Lie}\Big(\mathrm{Hol}_p(\omega)\Big)$ is a Lie subalgebra of γ spanned by the elements of the form

$$\Omega_q(X,Y) \qquad X, Y \in H_qP,$$

where $q \in P$ is a point which can be connected with p by a horizontal curve.

2.6 Flat connections

Definition 2.6.1. A connection is called *flat* if $\Omega = 0$.

The concept flat connection is important, since flat connections have the nice property that the space of horizontal vector fields is integrable. Moreover, in physics the vacuum state of a system corresponds with $\mathcal{F}_\mathcal{A} = 0$, which is the case if the connection is flat. So in order to understand the vacuum, it it important to understand flat connections. We start by introducing the concept op the canonical flat connection on a trivial bundle.

Definition 2.6.2. Let $P = M \times G$ be the trivial bundle and let $g : P \to G$ be the projection on G. Then the connection $H_pP = \ker g_*$ is called the *canonical flat connection*.

To show that this is indeed a connection, we remark $\psi : M \times G \to P$ given by $\psi^{-1}(p) = \big(\pi(p), g(p)\big)$ is a global section. Hence $\pi_* \times g_* : T_pP \to T_{\pi(p)}M \oplus T_{g(p)}G$ is an isomorphism. We have by definition $V_pP = \ker \pi_*$, thus $V_pP \simeq T_{g(p)}G$. Likewise, we have $H_pP = \ker g_* \simeq T_{\pi(p)}M$, so we see that $T_pP = H_pP \oplus V_pP$. Since π and g are smooth, this decomposition is also smooth. By Proposition 1.1.7 we have $R_a \circ g = g \circ R_a$, thus we have $g_*R_{a*}X = R_{a*}g_*X = 0$ if and only if $g_*X = 0$ for all vector field X on P. So we see that X is horizontal if and only if $R_{a*}X$ is horizontal, thus we have $R_{a*}H_pP = H_{pa}P$, which proves that H_pP is a connection.

Proposition 2.6.3. Let $P = M \times G$ be trivial and let $g : P \to G$ be the projection on G. Then the connection one-form corresponding to the canonical flat connection is $\omega = g^*\theta$.

Proof. Let $\omega = g^*\theta$. We first remark that since g is the projection onto G, we find that g_* is the projection onto T_gG, which is the vertical subspace. If X is a horizontal vector field, we have by definition $g_*X = 0$. Then

$$\omega(X) = g^*\theta(X) = \theta(g_*X) = 0,$$

so we have indeed that ω kills al the horizontal vector fields. Let X be vertical, which is the case if and only if there is a $A \in \mathfrak{g}$ such that $X = \sigma(A)$. Then we find

$$\begin{aligned}
\omega|_p(X|_p) &= g^*\theta|_{g(p)}(\sigma(A)|_p) \\
&= \theta|_{g(p)}g_*(\sigma(A)|_p) \\
&= \theta|_{g(p)}g(p)A && \text{[By Lemma 2.1.3]} \\
&= L_{g(p)^{-1}}g(p)A \\
&= A,
\end{aligned}$$

so $\omega \circ \sigma = \mathrm{Id}_{\mathfrak{g}}$. Furthermore, we have $R_a^*\omega = \mathrm{ad}_{a^{-1}}\omega$, since

$$\begin{aligned}
R_a^*\omega &= R_a^*g^*\theta \\
&= g^*R_a^*\theta && \text{[By Proposition 1.1.7]} \\
&= g^*\mathrm{ad}_{a^{-1}}\theta, && \text{[By (1.4.5)]}
\end{aligned}$$

so ω is indeed a connection one-form. ☐

Proposition 2.6.4. Let $P(M, G)$ be a principal bundle with a connection HP. Then the connection is flat if and only if the connection satisfies one of the following equivalent conditions:

1. $[X, Y]$ is horizontal for all horizontal vector fields X and Y;
2. For all $p \in P$ there is a neighborhood $U \subset M$ of $\pi(p)$ and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times G$ such that $\phi_*(HP)$ is the canonical flat connection in $U \times G$;
3. For all $p \in P$ there is a neighborhood $U \subset M$ of $\pi(p)$ and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times G$ such that the connection one-form ω restricted to $\pi^{-1}(U)$ equals $\phi^*g^*\theta$, where g is the projection of $\pi^{-1}(U)$ onto G;

Proof. We start by proving that $\Omega = 0$ is equivalent with (1). First, we find by (2.2.3) that $\Omega(X, Y) = 0$ if either X or Y is vertical. So we only have to prove that $\Omega = 0$ if we assume that both X and Y are horizontal. If we do so, we find

$$\begin{aligned}
\Omega(X, Y) &= \mathrm{d}\omega(HX, HY) \\
&= \mathrm{d}\omega(X, Y) \\
&= X\omega(Y) - Y\omega(X) - \omega[X, Y] && \text{[By } (A.2.3)\text{]} \\
&= -\omega([X, Y])
\end{aligned}$$

so we see that $\Omega = 0$ if and only if $\omega([X, Y]) = 0$, which is only the case if and only if $[X, Y]$ is horizontal. Secondly, we prove the equivalence between (2) and (3). We start by remarking that $\phi^{-1*}\omega$ corresponds with the connection $\phi_*(HP)$, since the first acts in a natural way on the latter: $\phi^{-1*}\omega(\phi_* X) = \omega\phi_*^{-1}\phi_* X = \omega(X)$. Then by previous proposition $\phi_*(HP)$ is canonical flat if and only if $\phi^{-1*}\omega = g^*\theta$, which is the case if and only if $\omega = \phi^* g^*\theta$. We omit the proof that $\Omega = 0$ implies (3) and refer to Theorem II.9.1 of [22], since the Ambrose-Singer Theorem, which we have not proven either, is essential for the proof. However, we can prove that (3) implies that $\Omega = 0$. So assume (3) and write $f = g \circ \phi$, so that we have $\omega = f^*\theta$ on U. Then by (1.4.8), we find that $\Omega = H^\vee\omega = H^\vee f^*\theta = -H^\vee f^*(\theta \wedge \theta)$ on U, hence we find $\Omega = -H^\vee(\omega \wedge \omega)$. Since the ω kills horizontal vectors, we find that $\Omega = 0$ of U. Since we can find such a neighborhood U of $\pi(p)$ for all $p \in P$, it follows that $\Omega = 0$ on the whole bundle. □

Definition 2.6.5. We call a gauge field $\mathcal{A}_i$ on U_i *pure gauge* if there is a smooth function $g : U_i \to G$ such that $\mathcal{A}_i = g^{-1}\mathrm{d}g$.

Corollary 2.6.6. $\mathcal{F}_i = 0$ if and only if $\mathcal{A}_i$ is pure gauge.

Proof. Let $\mathcal{F}_i = 0$. Since $\mathcal{F}_i = s_i^*\Omega$ and s_i is injective, we find that $\Omega = 0$ on the trivial bundle $U_i \times G$. By previous proposition there is some function $f : \pi^{-1}(U_i) \to G$ such that $\omega = f^*\theta$. Since $\mathcal{A}_i = s_i^*\omega$, we find $\mathcal{A}_i = g^*\theta$ with $g = f \circ s_i$. Hence by Proposition 1.4.11 we find $\mathcal{A}_i = g^{-1}\mathrm{d}g$.

Conversely, let $\mathcal{A}_i = g^{-1}\mathrm{d}g$ for some function $g : U_i \to G$. If we restrict ourselves to $U_i \times G$, we have by Proposition 2.4.13 we have a gauge transformation $\Phi : U_i \times G \to U_i \times G$ such that $H_i(\Phi) = g$. By Proposition 2.4.24 we see that $\mathcal{A}_i|_x = g(x)^{-1}A_i^\Phi|_x g(x) + g^{-1}\mathrm{d}g|_x$ with $\mathcal{A}_i^\phi = 0$. By (2.2.9) we have $\mathcal{F}_i^\Phi = 0$. Whence by Proposition 2.4.25 we have $\mathcal{F}_i^\Phi|_x = \mathrm{ad}_{g(x)^{-1}}\mathcal{F}_i|_x$, we must have $\mathcal{F}_i = 0$. □

3 Instantons and the Topology of Principal Bundles

3.1 Invariant polynomials

Definition 3.1.1. Let $\mathfrak{g}$ be the Lie algebra of a matrix Lie group G. Then a n-linear polynomial $P : \mathfrak{g}^n \to \mathbb{C}$ of degree n is called *symmetric* if

$$P(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_n) = P(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_n) \tag{3.1.1}$$

for all $i, j \in \{1, \ldots, n\}$ and $X_1, \ldots, X_n \in \mathfrak{g}$. A symmetric polynomial P of degree n is called *symmetric invariant* if

$$P(\mathrm{ad}_g X_1, \ldots, \mathrm{ad}_g X_n) = P(X_1, \ldots, X_n). \tag{3.1.2}$$

A *invariant polynomial* P_n of degree n is defined as a symmetric invariant polynomial P with all its entries equal:

$$P_n(X) = P(\underbrace{X, \ldots, X}_{n}). \tag{3.1.3}$$

Example 3.1.2. If $A \in \mathfrak{g}$, then $P_n(A) = \mathrm{tr}(A^n)$ is an invariant polynomial. Indeed, $P_n(A) = P(A, \ldots, A)$ with

$$P(A_1, \ldots, A_n) = \frac{1}{n!} \sum_{P \in S_n} \mathrm{tr}(A_{P(1)} A_{P(2)} \ldots A_{P(n)}), \qquad A_1, \ldots, A_n \in \mathfrak{g},$$

which is symmetric by definition and invariant by the cyclic properties of the trace.

Remark. Since a symmetric invariant polynomial is by definition linear in all its variables, it follows that the degree of each variable in every term is at most one. Furthermore, it is easy to see that we can decompose every symmetric invariant polynomial in a sum of invariant symmetric polynomials such that each variable occurs exactly once in every term.

Lemma 3.1.3. Let P be a symmetric invariant polynomial of degree n and let $A, A_1, \ldots, A_n \in \mathfrak{g}$. Then we have

$$\sum_{k=1}^{n} P(A_1, \ldots, A_{k-1}, [A, A_k], A_{k+1}, \ldots, A_n) = 0.$$

Proof. Let $g(t) \in G$ be generated by $A \in \mathfrak{g}$: $g(t) e^{tA}$. Then we have $P(\mathrm{ad}_{g(t)} A_1, \ldots, \mathrm{ad}_{g(t)} A_n) = P(A_1, \ldots, A_n)$. Since the right-hand side does not depend on t, both left- and right-hand side are zero when we differentiate it. Furthermore, we have $\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{ad}_{g(t)} A_k = \frac{\mathrm{d}}{\mathrm{d}t} e^{tA} A_k e^{-tA} = [A, \mathrm{ad}_{g(t)} A_k]$, whence we find

$$\begin{aligned}
\sum_{k=1}^{n} P(A_1, \ldots, [A, A_k], \ldots, A_n) &= \sum_{k=1}^{n} P\left(\mathrm{ad}_{g(t)} A_1, \ldots, \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{ad}_{g(t)} A_k, \ldots, \mathrm{ad}_{g(t)} A_n\right)\Bigg|_{t=0} \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} P(\mathrm{ad}_{g(t)} A_1, \ldots, \mathrm{ad}_{g(t)} A_n) \\
&= 0.
\end{aligned}$$

□

We can extent the definition of an invariant polynomial to a polynomial of $\mathfrak{g}$-valued p-forms.

Definition 3.1.4. Let $P : \mathfrak{g}^n \to \mathbb{C}$ a symmetric invariant polynomial. If $\alpha_1, \ldots, \alpha_n \in \Omega^p(U_i, \mathfrak{g})$, we define

$$P(\alpha_1, \ldots, \alpha_n) = \eta_1 \wedge \eta_2 \wedge \ldots \wedge \eta_n P(A_1, A_2, \ldots, A_n), \tag{3.1.4}$$

where $A_i \in \mathfrak{g}$ and $\eta_i \in \Omega^p(U_i)$ such that $A_i \otimes \eta_i = \alpha_k$ (no summation over i). Given P, we define an invariant polynomial P_n of degree n for $\mathfrak{g}$-valued p-forms by

$$P_n(\alpha) = P(\alpha, \ldots, \alpha) = \eta \wedge \eta \wedge \ldots \wedge \eta P(A, A, \ldots, A), \tag{3.1.5}$$

where $A \in \mathfrak{g}$ and $\eta \in \Omega^p(U_i)$ such that $A \otimes \eta = \alpha$.

Example 3.1.5. The most important example of an invariant polynomial of degree n for $\mathfrak{g}$-valued p-forms is $\mathrm{tr}(\alpha^n)$, where α^n is recursively defined by $\alpha^n = \alpha^{n-1} \wedge \alpha$. This polynomial is invariant, since it is the extension for p-forms of the polynomial given in Example 3.1.2. The importance of this example follows from the fact that the action of an instanton is $\int_M \mathrm{tr}(\mathcal{F}^2)$, as we will see below.

Proposition 3.1.6. Let P_n be an invariant polynomial. Then $P_n(\mathcal{F}_i)$ defines a global $2n$-form and is gauge invariant.

Proof. On overlaps U_{ij}, we have $\mathcal{F}_i|_x = \mathrm{ad}_{g_{ij}(x)} \mathcal{F}_j|_x$ by Proposition 2.2.7, so $P_n(\mathcal{F}_i) = P_n(\mathcal{F}_j)$ on U_{ij} by the invariance of P_n. Similarly, we have $\mathcal{F}_i^\Phi|_x = \mathrm{ad}_{\phi_i(x)^{-1}} \mathcal{F}_i|_x$ for all gauge transformation Φ, so we find $P_n(\mathcal{F}_i^\Phi) = P_n(\mathcal{F}_i)$. □

Since the $P_n(\mathcal{F}_i)$ define a global $2n$-form, we often drop the subscript and write $P_n(\mathcal{F})$.

Lemma 3.1.7. Let P be a symmetric invariant polynomial of degree n and let $\alpha, \alpha_1, \ldots, \alpha_n$ be $\mathfrak{g}$-valued forms of degree $p, p_1, \ldots, p_n$. Then we have

$$\sum_{k=1}^{n} (-1)^{p(p_1 + \ldots + p_{k-1})} P(\alpha_1, \ldots, \alpha_{k-1}, [\alpha, \alpha_k], \alpha_{k+1}, \ldots, \alpha_n) = 0.$$

Proof. Call the left-hand side of above identity S. If we write $\alpha = A \otimes \eta$ and $\alpha_i = A_i \otimes \eta_i$, we find by Lemma 3.1.3

$$\begin{aligned} S &= \sum_{k=1}^{n} (-1)^{p(p_1 + \ldots + p_{k-1})} \eta_1 \wedge \ldots \wedge \eta_{k-1} \wedge \eta \wedge \eta_k \wedge \eta_{k+1} \wedge \ldots \wedge \eta_n P(A_1, \ldots, [A, A_k], \ldots, A_n) \\ &= \sum_{k=1}^{n} \eta \wedge \eta_1 \wedge \ldots \wedge \eta_{k+1} \wedge \ldots \wedge \eta_n P(A_1, \ldots, [A, A_k], \ldots, A_n) \\ &= 0. \end{aligned}$$

□

Theorem 3.1.8 (Chern-Weil)**.** Let P_n be an invariant polynomial of degree n and $\mathcal{F} \in \Omega^2(U, \mathfrak{g})$ be the curvature 2-form on U (for convenience we drop the subscript i). Then $P_n(\mathcal{F})$ satisfies

1. $\mathrm{d}P_n(\mathcal{F}) = 0$.
2. If $\mathcal{F}$ and $\mathcal{F}'$ are curvature 2-forms corresponding to different gauge fields $\mathcal{A}$ and $\mathcal{A}'$. Then the difference $P_n(\mathcal{F}') - P_n(\mathcal{F})$ is gauge invariant and globally exact.

We first state a lemma.

Lemma 3.1.9. Let P be a symmetric invariant polynomial of degree n such that $P_n(X) = P(X, \dots, X)$ and let $\alpha_1, \dots, \alpha_n$ be $\mathfrak{g}$-valued forms of degree $p_1, \dots, p_n$. Then we have

$$\mathrm{d}P(\alpha_1, \dots, \alpha_n) = \sum_{k=1}^{n} (-1)^{p_1+\dots+p_{k-1}} P(\alpha_1, \dots, \alpha_{k-1}, \mathrm{d}\alpha_k, \alpha_{k+1}, \dots, \alpha_n).$$

Proof. If we write $\alpha_i = A_i \otimes \eta_i$ with $A_i \in \mathfrak{g}$ and η_i an ordinary p_i-form, we find using implicitly an induction step

$$\begin{aligned}
\mathrm{d}P(\alpha_1, \dots, \alpha_n) &= \mathrm{d}(\eta_1 \wedge \dots \eta_n) P(A_1, \dots, A_n) \\
&= \Big(\mathrm{d}\eta_1 \wedge \eta_2 \wedge \dots \wedge \eta_n + (-1)^{p_1} \eta_1 \wedge \mathrm{d}(\eta_2 \wedge \dots \eta_n)\Big) P(A_1, \dots, A_n) \\
&= \sum_{k=1}^{n} (-1)^{p_1+\dots+p_{k-1}} \eta_1 \wedge \dots \eta_{k-1} \wedge \mathrm{d}\eta_k \wedge \eta_{k+1} \wedge \dots \eta_n P(A_1, \dots, A_n) \\
&= \sum_{k=1}^{n} (-1)^{p_1+\dots+p_{k-1}} P(\alpha_1, \dots, \alpha_{k-1}, \mathrm{d}\alpha_k, \alpha_{k+1}, \dots, \alpha_n).
\end{aligned}$$

□

Proof of Theorem 3.1.8.

1. Let P_n be defined by the symmetric invariant polynomial P of degree n. We first assume every variable of P occurs exactly once in every term of P. We shall prove then that $\mathrm{d}P(\mathcal{F}, \dots, \mathcal{F}) = 0$. If we use Lemma 3.1.9 with $p_i = 2$, we find

$$\begin{aligned}
\mathrm{d}P(\mathcal{F}, \dots, \mathcal{F}) &= \sum_{k=1}^{n} P(\mathcal{F}, \dots, \mathrm{d}\mathcal{F}, \dots, \mathcal{F}) \\
&= \sum_{k=1}^{n} P(\mathcal{F}, \dots, \mathrm{d}\mathcal{F}, \dots, \mathcal{F}) + \sum_{k=1}^{n} (\mathcal{F}, \dots, [\mathcal{A}, \mathcal{F}], \dots, \mathcal{F}) \\
&= \sum_{k=1}^{n} P(\mathcal{F}, \dots, \mathcal{D}_{\mathcal{A}}\mathcal{F}, \dots, \mathcal{F}) \\
&= 0.
\end{aligned}$$

where we used Lemma 3.1.7 in the second equality, and in the last equality the Bianchi identity (2.3.13) and the fact that all variables of P occur in each term of P. The general case follows from the fact that d is linear and each symmetric invariant polynomial can be written as the sum of symmetric invariant polynomials with every variable occuring exactly once in a each term.

2. Firstly, we write $\tau = \mathcal{A}' - \mathcal{A}$ and $\mathcal{A}_t = \mathcal{A} + t\tau$, hence $\mathcal{A}_0 = \mathcal{A}$ and $\mathcal{A}_1 = \mathcal{A}'$. Then by (2.2.11), we have

$$\begin{aligned}
\mathcal{F}_t &= \mathrm{d}\mathcal{A} + \mathcal{A}_t \wedge \mathcal{A}_t \\
&= \mathcal{F} + t(\mathrm{d}\tau + \mathcal{A} \wedge \tau + \tau \wedge \mathcal{A}) + t^2 \tau \wedge \tau \\
&= \mathcal{F} + t\mathcal{D}_A \tau + t^2 \tau \wedge \tau, \qquad (3.1.6)
\end{aligned}$$

where we used (2.3.12) in the last equality. Then we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_t &= \mathcal{D}_{\mathcal{A}}\tau + 2t\tau \wedge \tau \\
&= \mathrm{d}\tau + \mathcal{A} \wedge \tau + t\tau \wedge \tau + \tau \wedge \mathcal{A} + \tau \wedge t\tau \\
&= \mathrm{d}\tau + \mathcal{A}_t \wedge \tau + \tau \wedge \mathcal{A}_t \\
&= \mathcal{D}_{\mathcal{A}_t}\tau.
\end{aligned}$$

Now, since P is both linear in all its arguments and totally symmetric, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}P_n(\mathcal{F}_t) = nP\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_t, \underbrace{\mathcal{F}_t, \ldots, \mathcal{F}_t}_{n-1}\right) = nP(\mathcal{D}_{\mathcal{A}_t}\tau, \mathcal{F}_t, \ldots, \mathcal{F}_t).$$

Now assume that every variable of P occurs exactly once in every term. Then

$$\begin{aligned}
\mathrm{d}P(\tau, \mathcal{F}_t, \ldots, \mathcal{F}_t) &= P(\mathrm{d}\tau, \mathcal{F}_t, \ldots, \mathcal{F}_t) - \sum_{k=2}^{n} P(\tau, \mathcal{F}_t, \ldots, \mathrm{d}\mathcal{F}_t, \ldots \mathcal{F}_t) \\
&= P(\mathrm{d}\tau + [\mathcal{A}_t, \tau], \mathcal{F}_t, \ldots, \mathcal{F}_t) - \sum_{k=2}^{n} P(\tau, \mathcal{F}_t, \ldots, \mathrm{d}\mathcal{F}_t + [\mathcal{A}_t, \mathcal{F}_t], \ldots \mathcal{F}_t) \\
&= P(\mathcal{D}_{\mathcal{A}_t}\tau, \mathcal{F}_t \ldots, \mathcal{F}_t) - \sum_{k=2}^{n} P(\tau, \mathcal{F}_t, \ldots, \mathcal{D}_{\mathcal{A}_t}\mathcal{F}_t, \ldots, \mathcal{F}_t) \\
&= P(\mathcal{D}_{\mathcal{A}_t}\tau, \mathcal{F}_t \ldots, \mathcal{F}_t),
\end{aligned}$$

where we used Lemma 3.1.9 in the first equality, Lemma 3.1.7 in the second, and the Bianchi identity (2.3.13) in the fourth. Since every symmetric invariant polynomial can be written as the sum of symmetric invariant polynomial such that every variable occurs exactly once in every term, we find that this identity holds for all symmetric invariant polynomials. Hence we find $\frac{\mathrm{d}}{\mathrm{d}t}P_n(\mathcal{F}_t) = \mathrm{d}nP(\tau, \mathcal{F}_t, \ldots, \mathcal{F}_t)$. Now, if we integrate this from $t = 0$ to $t = 1$, we find $\mathrm{d}P_n(\mathcal{F}') - P_n(\mathcal{F}) = \mathrm{d}Q_{2n-1}(\mathcal{A}', \mathcal{A})$, where we defined the *transgression* $Q_{2n-1}(\mathcal{A}', \mathcal{A})$ as

$$Q_{2n-1}(\mathcal{A}', \mathcal{A}) = n\int_0^1 P(\mathcal{A}' - \mathcal{A}, \underbrace{\mathcal{F}_t, \ldots, \mathcal{F}_t}_{n-1})\mathrm{d}t.$$

Now, if we once again add subscripts and write $\mathcal{A} = \mathcal{A}_i$ and $\mathcal{A}' = \mathcal{A}_i$, we see by (2.1.5) that $\mathcal{A}'_i|_x - \mathcal{A}_i|_x = \tau_i|_x = \mathrm{ad}_{g_{ij}(x)}\tau_j|_x = \mathrm{ad}_{g_{ij}(x)}(\mathcal{A}'_j|_x - \mathcal{A}_j|_x)$. Furthermore, we have $\mathcal{F}_{it}|_x = \mathrm{ad}_{ij}(x)\mathcal{F}_{jt}|_x$ by Proposition 2.2.7, and since P is invariant, we find that the $Q_{2n-1}(\mathcal{A}'_i, \mathcal{A}_i)$ agree on overlaps U_{ij} and so define a global function. So we find that $P_n(\mathcal{F}') - P_n(\mathcal{F})$ is globally exact. Furthermore, by Theorem 2.4.26, we find that $\mathcal{A}'^{\Phi}|_x - \mathcal{A}^{\Phi}|_x = \mathrm{ad}_{\phi_i(x)^{-1}}(\mathcal{A}'|_x - \mathcal{A}|_x)$ and $\mathcal{F}_t|_x = \mathrm{ad}_{\phi_i(x)^{-1}}\mathcal{F}_t|_x$, so that we find once again by the invariance of P that $Q_{2n-1}(\mathcal{A}', \mathcal{A})$ is gauge-invariant.

□

The Chern-Weil Theorem has some important implications. First, since $P_n(\mathcal{F})$ is a closed $2n$-form if P_n is an invariant polynomial, we see that $P_n(\mathcal{F})$ is a represent of a (de Rham) cohomology class $[P_n(\mathcal{F})] \in H^{2n}(M)$. Since $P_n(\mathcal{F})$ differs from $P_n(\mathcal{F}')$ by an exact form $\mathrm{d}Q_{2n-1}(\mathcal{A}', \mathcal{A})$, we see for a manifold M without boundary by Stokes' Theorem that

$$\int_M P_n(\mathcal{F}') - \int_M P_n(\mathcal{F}) = \int_M \mathrm{d}Q_{2n-1}(\mathcal{A}', \mathcal{A}) = \int_{\partial M} Q_{2n-1}(\mathcal{A}', \mathcal{A}) = 0,$$

so $([P_n(\mathcal{F})], M) := \int_M P_n(\mathcal{F})$ is a quantity which does not depend on the choice of connection, hence it is an invariant of the bundle. The Chern-Weil Theorem has another corollary: it allows us to prove that $P_n(\mathcal{F})$ is locally exact without referring to Poincaré's Lemma.

Corollary 3.1.10. Let P_n be an invariant polynomial. Then locally on a patch U we have $P_n(\mathcal{F}) = \mathrm{d}Q_{2n-1}(\mathcal{A}, 0)$, where $Q_{2n-1}(\mathcal{A}, 0)$ is called the *Chern-Simons form* of $P_n(\mathcal{F})$. Notice that $\mathrm{d}Q_{2n-1}(\mathcal{A}, 0) = n\int_0^1 P(\mathcal{A}, \mathcal{F}_t^{n-1})\mathrm{d}t$ with $\mathcal{A}_t = t\mathcal{A}$ and $\mathcal{F}_t = \mathrm{d}\mathcal{A}_t + \mathcal{A}_t \wedge \mathcal{A}_t = t\mathrm{d}\mathcal{A} + t^2\mathcal{A}\wedge\mathcal{A}$.

Proof. Since we work on a patch, the bundle $\pi^{-1}(U)(U,G)$ is trivial, and so we can define a flat connection $\mathcal{A}'$ on U. Since we always can find a gauge transformation Φ such that $(\mathcal{A}')^\Phi = 0$, we may assume that $\mathcal{A}' = 0$. This implies that $\mathcal{F}' = 0$ and from the Chern-Weil Theorem, we find that on U that $P_n(\mathcal{F}) = P_n(\mathcal{F}) - P_n(\mathcal{F}') = \mathrm{d}Q_{2n-1}(\mathcal{A},\mathcal{A}') = \mathrm{d}Q_{2n-1}(\mathcal{A},0)$. ☐

Example 3.1.11. Let $\mathrm{n}(\mathcal{F}) = \frac{1}{8\pi^2}\mathrm{tr}(\mathcal{F}\wedge\mathcal{F})$. Then we have $\mathrm{n}(\mathcal{F}) = P(\mathcal{F},\mathcal{F})$ with $P(A,B) = \frac{1}{8\pi^2}\mathrm{tr}(AB)$, so we find Chern-Simons form $K(\mathcal{A},0)$ of $\mathrm{n}(\mathcal{F})$ by

$$\begin{aligned}
K(\mathcal{A},0) &= 2\int_0^1 P(\mathcal{A},\mathcal{F}_t)\mathrm{d}t \\
&= \frac{2}{8\pi^2}\int_0^1 \mathrm{tr}\Big(\mathcal{A}\wedge(t\mathrm{d}\mathcal{A} + t^2\mathcal{A}\wedge\mathcal{A})\Big)\mathrm{d}t \\
&= \frac{1}{8\pi^2}\mathrm{tr}\left(\mathcal{A}\wedge\mathrm{d}\mathcal{A} + \frac{2}{3}\mathcal{A}\wedge\mathcal{A}\wedge\mathcal{A}\right) \\
&= \frac{1}{8\pi^2}\mathrm{tr}\left(\mathcal{A}\wedge\mathcal{F} - \frac{1}{3}\mathcal{A}\wedge\mathcal{A}\wedge\mathcal{A}\right),
\end{aligned}$$

using (2.2.11) in last equality.

3.2 Chern classes

Definition 3.2.1. Let E be a complex vector bundle (for instance the associated vector bundle of a principal bundle) with $k = \dim\ G$ and let $\mathcal{F}$ be the curvature 2-form of a connection on E. Then we define the *total Chern class* by

$$\mathrm{c}(\mathcal{F}) = \det\left(I + \frac{i\mathcal{F}}{2\pi}\right). \tag{3.2.1}$$

Since $\mathcal{F}$ is a two-form, $\mathrm{c}(\mathcal{F})$ is a direct sum of forms of even degrees,

$$\mathrm{c}(\mathcal{F}) = 1 + \mathrm{c}_1(\mathcal{F}) + \mathrm{c}_2(\mathcal{F}) + \ldots \tag{3.2.2}$$

where $\mathrm{c}_i(\mathcal{F}) \in \Omega^{2i}(M,\mathfrak{g})$ is called the ith *Chern class*.

Notice that if M is n-dimensional, we have $\mathrm{c}_i(\mathcal{F}) = 0$ if $2i > n$. It turns out that if we diagonalize the curvature form $\mathcal{F}$, it is more easy to compute the different Chern classes. In most of the cases we shall take $G = \mathrm{SU}(n)$. In that case the generators of its Lie algebra $\mathfrak{su}(n)$ are chosen to be anti-Hermitian, so the Hermitian matrix $i\mathcal{F}/2\pi$ can be diagonalized by $g \in \mathrm{SU}(k)$ and there are two-forms $x_1,\ldots,x_k$ such that $g^{-1}(i\mathcal{F}/2\pi)g = A$, where $A = \mathrm{diag}(x_1,\ldots,x_k)$. By the property of determinants $\det(XY) = \det(X)\det(Y)$, we see that $\mathrm{c}(\mathcal{F})$ is invariant, so that we find $\mathrm{c}(I + i\mathcal{F}/2\pi) = \mathrm{c}(gg^{-1} + g(i\mathcal{F}/2\pi)g^{-1}) = \mathrm{c}(I + A)$. Now, we have

$$\begin{aligned}
\det(I + A) &= \det\Big(\mathrm{diag}(1 + x_1,\ldots,1 + x_k)\Big) \\
&= \prod_{i=1}^k (1 + x_i) \\
&= 1 + \sum_{i=1}^k x_i + \sum_{i<j} x_i x_j + \ldots + x_1 x_2 \ldots x_k \\
&= 1 + \mathrm{tr}A + \frac{1}{2}\Big((\mathrm{tr}A)^2 - \mathrm{tr}A^2\Big) + \ldots + \det A.
\end{aligned}$$

We see that each term is an invariant polynomal, so we find

$$\begin{aligned}
\mathrm{c}_0(\mathcal{F}) &= 1\\
\mathrm{c}_1(\mathcal{F}) &= \frac{i}{2\pi}\mathrm{tr}\mathcal{F}\\
\mathrm{c}_2(\mathcal{F}) &= \frac{1}{8\pi^2}\Big(\mathrm{tr}(\mathcal{F}\wedge\mathcal{F}) - \mathrm{tr}\mathcal{F}\wedge\mathrm{tr}\mathcal{F}\Big)\\
&\vdots\\
\mathrm{c}_k(\mathcal{F}) &= \left(\frac{i}{2\pi}\right)^k \det\mathcal{F}.
\end{aligned}$$

Remark. Since these are all invariant polynomials, we find that the *Chern numbers* $\mathrm{c}_i(E) := ([\mathrm{c}_i(\mathcal{F})], M) = \int_M \mathrm{c}_i(\mathcal{F})$ are independent of the connection. If it is clear which vector bundle we work with, we write c_i instead of $\mathrm{c}_i(E)$.

The number $\mathrm{n} = \frac{1}{8\pi^2}\int \mathrm{tr}(\mathcal{F}\wedge\mathcal{F})$ is called the *topological charge* for reasons shown below. We shall prove below that $\mathrm{n}\in\mathbb{Z}$ if $M = S^4$. Notice that since $\mathrm{tr}(A^2)$ is an invariant polynomial, we have that n is also independent of the choice of connection. Furthermore, notice that $\mathrm{c}_2 = \mathrm{n} + \frac{1}{2}\mathrm{c}_1$, so for bundles with zero first Chern number, we have $\mathrm{c}_2 = \mathrm{n}$.

Lemma 3.2.2. Let $E\oplus E'$ be the Whitney sum of two vector bundles E and E' over M. Let $\mathcal{F}_E$ and $\mathcal{F}_{E'}$ be field-strengths of E respectively E'. By (1.2.9) $\mathcal{F}_{E\oplus E'} = \mathcal{F}_E \oplus F_{E'}$ is a field-strength on $E\oplus E'$. Then we have

$$\mathrm{c}(\mathcal{F}_{E\oplus E'}) = \mathrm{c}(\mathcal{F}_E)\wedge\mathrm{c}(\mathcal{F}'_E).$$

Proof. Since we have

$$\mathcal{F}_{E\oplus E'} = \begin{pmatrix} \mathcal{F}_E & 0\\ 0 & \mathcal{F}_{E'}\end{pmatrix}$$

we find

$$\begin{aligned}
\mathrm{c}(\mathcal{F}_{E\oplus E'}) &= \det\left(I + \frac{i\mathcal{F}_{E\oplus E'}}{2\pi}\right)\\
&= \det\begin{pmatrix} I + \frac{i\mathcal{F}_E}{2\pi} & 0\\ 0 & I + \frac{i\mathcal{F}_{E'}}{2\pi}\end{pmatrix}\\
&= \det\left(I + \frac{i\mathcal{F}_E}{2\pi}\right)\wedge\det\left(I + \frac{i\mathcal{F}_{E'}}{2\pi}\right)\\
&= \mathrm{c}(\mathcal{F}_E)\wedge\mathrm{c}(\mathcal{F}_{E'}).
\end{aligned}$$

□

Corollary 3.2.3. Let E be a complex vector bundle over M with $\dim M = 4$ and $G = \mathrm{SU}(n)$. If $E = E_1\oplus\ldots\oplus E_n$ with E_i vector bundles over M with fibre $\mathbb{C}^2$. Then we have

$$\mathrm{c}_2(E) = \sum_{i=1}^n \mathrm{c}_2(E_i).$$

Proof. Let $\mathcal{F}_1,\ldots,F_n$ be the field-strengths of $E_1,\ldots,E_n$ and let $\mathcal{F}$ be the field-strength of E. Thus we have $\mathcal{F} = \mathcal{F}_1\oplus\ldots\oplus\mathcal{F}_n$. Since $\dim M = 4$, we have $\mathrm{c}_k(\mathcal{F}) = \mathrm{c}_k(\mathcal{F}_i) = 0$ for all

$i \in \{1, \ldots, n\}$ and $k > 2$. Furthermore, we have $\mathrm{c}_1(\mathcal{F}) = \mathrm{c}_1(\mathcal{F}_i) = 0$, since field-strengths on E and the E_i take values in $\mathfrak{su}(n)$, which is traceless. Then by previous lemma, we find

$$\begin{aligned} 1 + \mathrm{c}_2(\mathcal{F}) &= \mathrm{c}(\mathcal{F}) \\ &= c(\mathcal{F}_1) \wedge \ldots \wedge \mathrm{c}(\mathcal{F}_n) \\ &= \Big(1 + \mathrm{c}_2(\mathcal{F}_1)\Big) \wedge \ldots \wedge \Big(1 + \mathrm{c}_2(\mathcal{F}_n)\Big) \\ &= 1 + \Big(\mathrm{c}_2(\mathcal{F}_1) + \ldots + \mathrm{c}_2(\mathcal{F}_n)\Big), \end{aligned}$$

so we see that $\mathrm{c}_2(\mathcal{F}) = \mathrm{c}_2(\mathcal{F}_1) + \ldots + \mathrm{c}_2(\mathcal{F}_n)$, from which the statement follows. □

3.3 Instantons in Euclidean space

In this subsection we introduce the notion of an action, an object from which the equations of motion can be derived. These equations are important, since they describe the dynamics of the forces in gauge theories. We refer to A.4, where most of the notation used is introduced. Before we make the first definition, we need the following lemma.

Lemma 3.3.1. The object $\mathrm{tr}(\mathcal{F}_i \wedge *\mathcal{F}_i)$ is globally defined and gauge invariant.

Proof. By the cyclic property of the trace, $\mathrm{tr}(\mathcal{F}_i \wedge *\mathcal{F}_i)$ is ad-invariant. Then by Proposition 2.2.7 we see that $\mathrm{tr}(\mathcal{F}_i \wedge *\mathcal{F}_i) = \mathrm{tr}(\mathcal{F}_j \wedge *\mathcal{F}_j)$ on overlaps U_{ij}. Similarly, by the cyclic property of the trace and Theorem 2.4.26, it follows that $\mathrm{tr}(\mathcal{F}_i \wedge *\mathcal{F}_i)$ is gauge-invariant. □

Since $\mathrm{tr}(\mathcal{F}_i \wedge *\mathcal{F}_i)$ is globally defined, we shall drop the subscript i. Now we are able to define the action.

Definition 3.3.2. Let M be a four-dimensional Riemannian manifold. Then the *Yang-Mills action* is defined by

$$S_E[\mathcal{A}] = -\int_M \mathrm{tr}(\mathcal{F} \wedge *\mathcal{F}). \tag{3.3.1}$$

Notice that since $\mathrm{tr}(\mathcal{F} \wedge *\mathcal{F})$ is not obtained by an invariant symmetric polynomial, we cannot conclude that the action is independent of the connection. This is what we want, since we would like to put further constraints on the connection by the equations of motion we ontain by minimalizing the action. We note that by Definition A.4.3 and the following remark we have

$$S_E[\mathcal{A}] = \|\mathcal{F}\|^2 = \int_M |\mathcal{F}|^2 \mathrm{dVol}(g), \tag{3.3.2}$$

where we chose the normalization factor $\lambda = 1$. If we write $\mathcal{F} = \frac{1}{2}\mathcal{F}_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, we find

$$S_E[\mathcal{A}] = -\frac{1}{4}\int_M \mathrm{tr}(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu})\mathrm{dVol}(g),$$

which will be recognized by physicists as the usual action for a gauge theory. We recall that the Bianchi identity is given by $\mathcal{D}_{\mathcal{A}}\mathcal{F} = 0$. We can also state a similar identity for $*\mathcal{F}$.

Definition 3.3.3. The equation

$$\mathcal{D}_{\mathcal{A}} * \mathcal{F} = 0 \tag{3.3.3}$$

is called the *Yang-Mills equation*. Its solutions $\mathcal{A}$ are called *Yang-Mills connections*.

We emphasize that the nature of the Bianchi-identity is geometrical, while the nature of the Yang-Mills equation is dynamical. We write $\mathcal{F}[\mathcal{A}]$ when we want to express the $\mathcal{A}$-dependence of $\mathcal{F}$. Next proposition relates the Yang-Mills equation to the action.

Proposition 3.3.4. The solutions $\mathcal{A}$ of the Yang-Mills equations on a coordinate neighborhood U are precisely the minima of the Yang-Mills action.

Proof. Let $\alpha \in \Omega^1(U, \mathfrak{g})$, then we find that $\mathcal{A}$ is a minimum of $S_E[\mathcal{A}] = \|\mathcal{F}[\mathcal{A}]\|^2$ if

$$\frac{\mathrm{d}}{\mathrm{d}t} S_E[\mathcal{A} + t\alpha]\bigg|_{t=0} = 0 \tag{3.3.4}$$

for all $\alpha \in \Omega^1(U, \mathfrak{g})$. We start by expressing $\mathcal{F}[\mathcal{A} + \alpha]$ in terms of $\mathcal{F}[\mathcal{A}]$ and α. Since we have $\mathcal{F}[\mathcal{A}] = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$, we find

$$\begin{aligned}
\mathcal{F}[\mathcal{A} + t\alpha] &= \mathrm{d}(\mathcal{A} + t\alpha) + (\mathcal{A} + t\alpha) \wedge (\mathcal{A} + t\alpha) \\
&= \mathrm{d}\mathcal{A} + t\mathrm{d}\alpha + \mathcal{A} \wedge \mathcal{A} + t\alpha \wedge \mathcal{A} + t\mathcal{A} \wedge \alpha + t^2 \alpha \wedge \alpha \\
&= \mathcal{F}[\mathcal{A}] + t\mathrm{d}\alpha + t\alpha \wedge \mathcal{A} + t\mathcal{A} \wedge \alpha + t^2 \alpha \wedge \alpha \\
&= \mathcal{F}[\mathcal{A}] + t\mathrm{d}\alpha + t[\mathcal{A}, \alpha] + t^2 \alpha \wedge \alpha \\
&= \mathcal{F}[\mathcal{A}] + t\mathcal{D}_{\mathcal{A}}\alpha + t^2 \alpha \wedge \alpha,
\end{aligned}$$

where $[\mathcal{A}, \alpha] = \mathcal{A} \wedge \alpha + \alpha \wedge \mathcal{A}$ by (A.1.2). Then we find

$$\begin{aligned}
\|F[\mathcal{A} + t\alpha]\|^2 &= (\mathcal{F}[\mathcal{A}] + t\mathcal{D}_{\mathcal{A}}\alpha + t^2 \alpha \wedge \alpha, \mathcal{F}[\mathcal{A}] + t\mathcal{D}_{\mathcal{A}}\alpha + t^2 \alpha \wedge \alpha) \\
&= \|F[\mathcal{A}]\|^2 + 2t(\mathcal{F}[\mathcal{A}], \mathcal{D}_{\mathcal{A}}\alpha) + t^3(\alpha, \alpha \wedge \alpha) + t^4(\alpha \wedge \alpha, \alpha \wedge \alpha),
\end{aligned}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t} \|F[\mathcal{A} + t\alpha]\|^2\bigg|_{t=0} = 2(\mathcal{F}[\mathcal{A}], \mathcal{D}_{\mathcal{A}}\alpha). \tag{3.3.5}$$

By taking the adjoint of $\mathcal{D}_A$ we find by (3.3.4) that

$$(\mathcal{D}^*_{\mathcal{A}}\mathcal{F}[\mathcal{A}], \alpha) = 0 \tag{3.3.6}$$

for all $\alpha \in \Omega^1(U, \mathfrak{g})$ if and only if $\mathcal{D}^*_{\mathcal{A}}\mathcal{F}[\mathcal{A}] = 0$. By Theorem A.4.11 we see that this is precisely the case if $\mathcal{A}$ is a solution of the Yang-Mills equation $\mathcal{D}_{\mathcal{A}} * \mathcal{F}[\mathcal{A}] = 0$. ☐

Only for $G = \mathrm{U}(1)$ are the Yang-Mills equations relatively easy to solve. For more complicated groups, we obtain a non-linear second-order differential equation, for which it is nearly impossible to give all possible solutions.

Definition 3.3.5. The absolute minima of the Yang-Mills action are called *instantons*.

It follows immediately from the definition that instantons are solutions of the Yang-Mills equation. We shall see below that instantons are relatively easy to find. First consider a coordinate neighborhood U. Then by Proposition A.4.8 the operator $** : \Omega^2(U, \mathfrak{g}) \to \Omega^2(U, \mathfrak{g})$ equals the identity, since M is a four-dimensional manifold, so $*$ has eigenvalues ± 1. Hence we find that $\Omega^2(U, \mathfrak{g})$ can be split in an orthogonal decomposition of the eigenspaces of $*$:

$$\Omega^2(U, \mathfrak{g}) = \Omega^2_+(U, \mathfrak{g}) \oplus \Omega^2_-(U, \mathfrak{g}). \tag{3.3.7}$$

Definition 3.3.6. A $\mathfrak{g}$-valued two-form α is called *self-dual* if $\alpha \in \Omega^2_+(U, \mathfrak{g})$ and *anti-self-dual* if $\alpha \in \Omega^2_-(U, \mathfrak{g})$. Equivalently, we say that α is self-dual if $*\alpha = \alpha$ and anti-self-dual if $*\alpha = -\alpha$. A connection is called (anti-)self-dual if its gauge field-strength is (anti-)self-al, that is $\mathcal{F}$ satisfies the *(anti-)self-duality equation* (abbreviated as the (A)SD-equations)

$$\mathcal{F} = \pm * \mathcal{F}. \tag{3.3.8}$$

We denote the space of self-dual connections with $\mathscr{A}^+$ and the space of anti-self-dual connections with $\mathscr{A}^-$.

Let $\alpha \in \Omega^2(U,\mathfrak{g})$. Then we define $\alpha_\pm = \frac{1}{2}(\alpha \pm *\alpha)$. So we have $\alpha_\pm \in \Omega^2_\pm(U,\mathfrak{g})$ and $\alpha = \alpha_+ + \alpha_-$. Since $\Omega^2_+(U,\mathfrak{g})$ and $\Omega^2_-(U,\mathfrak{g})$ are orthogonal, we have

$$\alpha_+ \wedge \alpha_- = \alpha_+ \wedge * * \alpha_- = -\alpha_+ \wedge *\alpha_- = -\langle \alpha_+, \alpha_- \rangle \mathrm{dVol}(g) = 0. \tag{3.3.9}$$

Proposition 3.3.7. A connection $\mathcal{A}$ is an instanton if and only if its gauge field-strength $\mathcal{F}$ is either self-dual or anti-self-dual. In other words, $\mathcal{A}$ is an instanton if and only if $\mathcal{A} \in \mathscr{A}^\pm$.

Proof. Given the gauge field strength $\mathcal{F}$, we can split it into self-dual and anti-self-dual parts $\mathcal{F} = \mathcal{F}_+ + \mathcal{F}_-$. Since $\mathcal{F}_+ \perp \mathcal{F}_-$, we find

$$S_E[\mathcal{A}] = \int_M |\mathcal{F}_+|^2 \mathrm{dVol}(g) + \int_M |\mathcal{F}_-|^2 \mathrm{dVol}(g). \tag{3.3.10}$$

We shall proof that $8\pi^2\mathrm{n} = \int_M \mathrm{tr}(\mathcal{F} \wedge \mathcal{F})$, the topological charge, is an lower bound for $S_E[\mathcal{A}]$. We find

$$\begin{aligned}
8\pi^2\mathrm{n} &= \int_M \mathrm{tr}\Big((\mathcal{F}_+ + \mathcal{F}_-) \wedge (F_+ + \mathcal{F}_-)\Big) \\
&= \int_M \mathrm{tr}(F_+ \wedge \mathcal{F}_+) + \int_M \mathrm{tr}(\mathcal{F}_+ \wedge \mathcal{F}_-) + \int_M \mathrm{tr}(\mathcal{F}_- \wedge \mathcal{F}_+) + \int_M \mathrm{tr}(\mathcal{F}_- \wedge \mathcal{F}_-) \\
&= \int_M \mathrm{tr}(\mathcal{F}_+ \wedge * * \mathcal{F}_+) + \int_M \mathrm{tr}(\mathcal{F}_- \wedge * * \mathcal{F}_-) \\
&= \int_M \mathrm{tr}(\mathcal{F}_+ \wedge *\mathcal{F}_+) - \int_M \mathrm{tr}(\mathcal{F}_- \wedge *\mathcal{F}_-) \\
&= \int_M \langle \mathcal{F}_+, \mathcal{F}_+ \rangle \mathrm{dVol}(g) - \int_M \langle \mathcal{F}_-, \mathcal{F}_- \rangle \mathrm{dVol}(g) \\
&= -\int_M |\mathcal{F}_+|^2 \mathrm{dVol}(g) + \int_M |\mathcal{F}_-|^2 \mathrm{dVol}(g).
\end{aligned} \tag{3.3.11}$$

Hence we see that

$$S_E[\mathcal{A}] \geq 8\pi^2|\mathrm{n}|, \tag{3.3.12}$$

with equality if and only if $\mathcal{F}_+$ or $\mathcal{F}_-$ vanishes, that is if and only if $\mathcal{F} = \mp * \mathcal{F}$. □

We could also give the ASD-equations in coordinates. Then the equation $\mathcal{F}^+_{\mathcal{A}} = 0$ is equivalent to

$$\mathcal{F}_{12} + \mathcal{F}_{34} = 0; \tag{3.3.13}$$
$$\mathcal{F}_{14} + \mathcal{F}_{23} = 0; \tag{3.3.14}$$
$$\mathcal{F}_{13} + \mathcal{F}_{42} = 0. \tag{3.3.15}$$

or

$$[\mathcal{D}_1, \mathcal{D}_2] + [\mathcal{D}_3, \mathcal{D}_4] = 0; \tag{3.3.16}$$
$$[\mathcal{D}_1, \mathcal{D}_4] + [\mathcal{D}_2, \mathcal{D}_3] = 0; \tag{3.3.17}$$
$$[\mathcal{D}_1, \mathcal{D}_3] + [\mathcal{D}_4, \mathcal{D}_2] = 0. \tag{3.3.18}$$

by Proposition 2.3.13, where we used the abbreviation $\mathcal{D}_\mu$ instead of $(\mathcal{D}_\mathcal{A})_\mu$.

We remark that the (anti-)self-duality equation is of first order, and is therefore more easy to solve than the Yang-Mills equation. We have seen that solutions of the first equation are also solutions of the latter, but this also follows directly, since if $\mathcal{F} = \pm*\mathcal{F}$, we find $\mathcal{D}_\mathcal{A}*\mathcal{F} = \pm\mathcal{D}_\mathcal{A}\mathcal{F}$, which is always zero by the Bianchi identity.

Corollary 3.3.8. The action for an instanton equals $S_E[\mathcal{A}] = \int_M \mathrm{tr}(\mathcal{F} \wedge \mathcal{F}) = 8\pi^2\mathrm{n}$.

If $\mathcal{A}$ is an instanton with gauge field-strength $\mathcal{F}$, we often speak of the *instanton number* of $\mathcal{F}$ instead of the topological charge of $\mathcal{F}$. Notice that the matrices of $\mathfrak{su}(n)$ are traceless, so for $G = \mathrm{SU}(n)$, the first Chern number is zero and we find that $\mathrm{n} = \mathrm{c}_2$. Next proposition shows that instantons do not exist in electrodynamics, since its structure group U(1) is abelian.

Proposition 3.3.9. Let M be a four-dimensional base space of a principal bundle with abelian structure group G. Then if $\mathcal{F}$ is (anti-)self-dual, we have $\mathcal{F} = 0$.

Proof. Since G is an abelian group, we find that $\mathfrak{g}$ is abelian, so commutators vanish and we have $\mathcal{F} = \mathrm{d}\mathcal{A}$ by (2.2.9). Since we have $\mathrm{d}^\dagger = *\mathrm{d}*$ by Theorem A.4.9, we find

$$\begin{aligned} \|\mathcal{F}\|^2 &= (\mathcal{F},\mathcal{F}) = \pm(*\mathcal{F},\mathcal{F}) = \pm(*\mathrm{d}\mathcal{A},\mathrm{d}\mathcal{A}) \\ &= \pm(*\mathrm{d}**\mathcal{A},\mathrm{d}\mathcal{A}) = \pm(\mathrm{d}^\dagger*\mathcal{A},\mathrm{d}\mathcal{A}) = \pm(*\mathcal{A},\mathrm{d}^2\mathcal{A}) = 0. \end{aligned}$$

□

Proposition 3.3.10. Let P be a principal bundle. Then all instantons on P have the same instanton number n.

Proof. Since n is obtained from the invariant polynomial $\mathrm{n}(\mathcal{F}) = \mathrm{tr}(\mathcal{F}\wedge\mathcal{F})$, the statement follows directly from Theorem 3.1.8. □

Last proposition has a direct consequence that instantons cannot exist on trivial bundles.

Corollary 3.3.11. Let $P = M \times G$ be a trivial bundle with M four-dimensional. Then if $\mathcal{F}$ is (anti-)self-dual, we have $\mathcal{F} = 0$.

Proof. Since P is trivial, we can find a globally defined connection $\mathcal{A}$ with global defined gauge-field strength $\mathcal{F}$ such that $\mathcal{F} = 0$. Then we have $\mathrm{n} = \frac{1}{8\pi^2}\int_M \mathrm{tr}(\mathcal{F}\wedge\mathcal{F}) = 0$. Then by previous proposition we find for arbitrary (anti-)self-dual $\mathcal{F}$ that $\|\mathcal{F}\|^2 = S_E[\mathcal{A}] = 0$, whence $\mathcal{F} = 0$. □

3.4 The Minkowski case

Since we do not live in an Euclidean space $\mathbb{R}^4$, but in Minkowski space $\mathbb{R}^{1,3}$, it is interesting to investigate wether our concepts can be translated to latter space. In quantum field theory one usually switches between Minkowski space with coordinates x^0, x^1, x^2, x^3 and Euclidean space with coordinates x^1, x^2, x^3, x^4 by a Wick transformation $x^0 \mapsto ix^4$. So we can Wick transform our concepts in Euclidean space to Minkowski space by sending x^4 to $-ix^0$. Hence we find that $\mathrm{dVol}(g) = \mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3\wedge\mathrm{d}x^4$ in $\mathbb{R}^4$ transforms into $i\mathrm{d}x^0\wedge\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3$ in $\mathbb{R}^{1,3}$ (notice that moving $\mathrm{d}x^0$ to the left gives a change of sign). Let A be the gauge field and F the gauge-field strength in Minkowski space. Then the Yang-Mills action $S_M[A]$ in Minkowski space picks up an extra factor i compared with $S_E[\mathcal{A}]$ and becomes

$$\begin{aligned} S_M[A] &= -\frac{i}{4}\int_M \mathrm{tr}(F_{\mu\nu}F^{\mu\nu})\mathrm{dVol}(g) \\ &= -i\int_M \mathrm{tr}(F\wedge *F). \end{aligned}$$

By using the same technics as in the proof of Proposition 3.3.4, we see that the solutions of the Yang-Mills equation $\mathcal{D}_A * F = 0$ are precisely the minima of S_M. In contrary to the Euclidean case, we see by Proposition A.4.8 that $** : \Omega^2(U,\mathfrak{g}) \to \Omega^2(U,\mathfrak{g})$ equals $-\mathrm{Id}$. Thus we see that $*$ has eigenvalues $\pm i$, whence the absolute minima of the Euclidean action satisfy $*F = \pm iF$. Therefore, instantons do not exist in Minkowski space, since the gauge field-strength F and its dual $*F$ are real objects.

3.5 Classifying bundles over S^4

Since $\mathbb{R}^4$ is not compact, it is possible that the integral in the action S_E diverges. Therefore, we shall only consider connections with finite action, that is connections for which the integral in the action converges. We can accomplish this by restrict us to gauge field-strengths $\mathcal{F}$ which vanish at infinity. In other words, we demand that $\mathcal{F}|_x \to 0$ if $|x| \to \infty$, or equivalently that $\mathcal{A}|_x \to g^{-1}\mathrm{d}g$ if $|x| \to \infty$ for some function g with values in G.

Since S^4 minus a point and $\mathbb{R}^4$ are conformally equivalent (just like the two-sphere minus a point by stereographic projection is conformally equivalent to the plane), a connection for which the action is finite defines an action on S^4. Whence the Hodge star on two-forms is conformally invariant (Proposition A.4.7), we find that an instanton on $\mathbb{R}^4$ defines an instanton on S^4. Hence we shall consider $M = S^4$, the one-point compactification of $\mathbb{R}^4$, which is a much nicer space to work with. We shall prove that if G is a simple Lie group containing SU(2) as a subgroup, then two G-bundles over S^4 with the same instanton number are equivalent.

First we state a theorem which proof can be found in [6] and which will need in order to assure that all our analysis will not only be valid for G equal to SU(2), but also for Lie groups G containing SU(2) as a subgroup.

Theorem 3.5.1 (R. Bott)**.** Let G be a simple Lie group containing SU(2) as a subgroup. Then every map $S^3 \to G$ is homotopic to a map $S^3 \to \mathrm{SU}(2)$.

Convention. Unless otherwise stated, we shall from now on assume that G is a simple Lie group containing SU(2) as a subgroup.

We can cover S^4 by two charts U_N and U_S, being the northern respectively the southern hemisphere:

$$U_N = \left\{ x \in S^4 : |x| < \frac{1}{2} + \epsilon \right\}$$
$$U_S = \left\{ x \in S^4 : |x| > \frac{1}{2} - \epsilon \right\}$$

where $\epsilon > 0$ is a small number. Since all the topological information about the bundle lies in the transition maps, and we here only have one transition map $g_{NS} : U_{NS} \to G$, we can classify all principal bundles by classifying all maps $U_{NS} \to G$. Since $U_{NS} = U_N \cap U_S$ is homotopic to S^3, this means we can classify all principal bundles over S^4 by classifying all maps $g : S^3 \to G$. We will do this by assigning to these maps an integer, called the degree, which is introduced in next definition.

Definition 3.5.2. Let $f : S^n \to S^n$ be a continuous map and let $H_k(S^n)$ be the kth homology group of S^n. Then the *degree* of f is defined as the integer k such that $f_* : H_n(S^n) \to H_n(S^n)$ satisfies $f_*([x]) = k[x]$ for an arbitrary $[x] \in H_n(S^n)$.

Notice that $\deg(f)$ does not depend on the choice of $[x]$, for f_* is continuous with a discrete codomain, so it must be a constant function.

Theorem 3.5.3. Let $f, g : M \to N$ be homotopic to each other and let A be an abelian group. Then we have $f_* = g_* : H_k(M; A) \to H_k(N; A)$ and $f^* = g^* : H^k(N; A) \to H^k(M; A)$.

This result is proven in [7] as Theorem V.8.5. As a corollary, we find that $\deg(f) = \deg(g)$ if $f, g : S^n \to S^n$ are homotopic.

Proposition 3.5.4. We can classify all maps $g : S^3 \to G$ up to homotopy by assigning to each map g an integer $n(g)$.

Proof. By Bott's theorem, we see that every map $g : S^3 \to G$ is homotopic to a map $\bar{g} : S^3 \to \mathrm{SU}(2)$. Since latter group is homeomorphic as topological space to S^3, we can calculate the degree of $\bar{g}$. Then $n(g)$ is given by $\deg(\bar{g})$. Whence homotopy is an equivalence relation, it follows that $n(g)$ does not depend on the choice of a representative of the homotopy class of g. □

Since it might be hard to find this $\bar{g}$ homotopic to g, it is important that we find other ways to calculate $n(g)$.

Theorem 3.5.5. Let $g : S^3 \to G$ be a map. Then the integer $n(g)$ can be computed by the following integral

$$n(g) = \frac{1}{24\pi^2} \int_{S^3} \mathrm{tr}(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g). \tag{3.5.1}$$

Before we can proof this theorem, we formulate de Rham's Theorem, which is proven in [7] as Theorem V.9.1.

Theorem 3.5.6 (de Rham)**.** Let M be an n-manifold and let c, ω be representatives of a homology class in $H_k(M;\mathbb{R})$ respectively a de Rham cohomology class in $H^k_{\mathrm{dR}}(M)$. Then the pairing $\langle [c], [\omega]\rangle := \int_c \omega$ induces an isomorphism between the de Rham cohomology group $H^k_{\mathrm{dR}}(M)$ and the singular cohomology group $H^k(M;\mathbb{R})$.

Proof of Theorem 3.5.5. Using Bott's Theorem, let $\bar{g} : S^3 \to \mathrm{SU}(2)$ be the map homotopic to g. From the theorem we find that if $f : S^n \to S^n$ is a map and ω is a volume form on S^n, we have $\int_{f_*[c]} \omega = \int_{[c]} f^*\omega$, hence

$$\deg(f) \int_{S^n} \omega = \deg(f)\langle S^n, \omega\rangle = \langle f_* S^n, \omega\rangle = \langle S^n, f^*\omega\rangle = \int_{S^n} f^*\omega. \tag{3.5.2}$$

On SU(2) the form $\mathrm{tr}(\theta \wedge \theta \wedge \theta)$ is a volume form, and since S^3 is homeomorphic to SU(2), we find

$$n(g) = \deg(\bar{g}) \int_{\mathrm{SU}(2)} \mathrm{tr}(\theta \wedge \theta \wedge \theta) = \int_{S^3} \bar{g}^*\mathrm{tr}(\theta \wedge \theta \wedge \theta) = \int_{S^3} g^*\mathrm{tr}(\theta \wedge \theta \wedge \theta), \tag{3.5.3}$$

where last equality follows by Theorem 3.5.3. Without proof we state that $\int_{\mathrm{SU}(2)} \mathrm{tr}(\theta\wedge\theta\wedge\theta) = 24\pi^2$. By $g^*\theta = g^{-1}\mathrm{d}g$ we find

$$n(g) = \frac{1}{24\pi^2} \int_{S^3} \mathrm{tr}(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g). \tag{3.5.4}$$

□

We shall proceed by showing that $-n(g)$ is exactly the topological charge of a gauge-field strength $\mathcal{F}$.

Theorem 3.5.7. Let $\mathcal{F}$ be a gauge-field strength on $P(S^4, G)$. Then the topological charge n is an integer which equals $-n(g)$, where $g : U_{NS} \to G$ is the transition map and U_N and U_S are the two coordinate charts on S^4 defined as above.

Proof. First notice that U_N and U_S have opposite orientations, so that we have $\partial U_N = -\partial U_S = S^3$. Furthermore, we have $\mathcal{A}_N = g^{-1}\mathcal{A}_S g + g^{-1}\mathrm{d}g$ and $\mathcal{F}_N = g^{-1}\mathcal{F}_S g$. If K is the Chern-Simons form of $\mathrm{n}(\mathcal{F}) = \frac{1}{8\pi^2}\mathrm{tr}(\mathcal{F}\wedge\mathcal{F})$, which is calculated in Example 3.1.11, we find

$$\begin{aligned}
\mathrm{n} &= \int_{S^4} \mathrm{n}(\mathcal{F}) \\
&= \int_{U_N} \mathrm{n}(\mathcal{F}_N) + \int_{U_S} \mathrm{n}(F_S) \\
&= \int_{U_N} \mathrm{d}K(\mathcal{A}_N, 0) - \int_{-U_S} \mathrm{d}K(\mathcal{A}_S, 0) \\
&= \int_{\partial U_N} K(\mathcal{A}_N, 0) - \int_{-\partial U_S} K(\mathcal{A}_S, 0) \\
&= \frac{1}{8\pi^2}\int_{S^3} \mathrm{tr}\left(\mathcal{F}_N\wedge\mathcal{A}_N - \frac{1}{3}\mathcal{A}_N^3 - \mathcal{F}_S\wedge\mathcal{A}_S + \frac{1}{3}\mathcal{A}_S^3\right),
\end{aligned}$$

where we used the notation $\beta^n = \underbrace{\beta\wedge\beta\wedge\ldots\wedge\beta}_{n \text{ times}}$ for $\mathfrak{g}$-valued forms. By (2.2.11) we find

$$\begin{aligned}
\mathrm{tr}(\mathcal{F}_N\wedge\mathcal{A}_N) &= \mathrm{tr}\Big(g^{-1}\mathcal{F}_S g\wedge(g^{-1}\mathcal{A}_S g + g^{-1}\mathrm{d}g)\Big) \\
&= \mathrm{tr}(\mathcal{F}_S\wedge\mathcal{A}_S + g^{-1}\mathcal{F}_S\wedge\mathrm{d}g) \\
&= \mathrm{tr}(\mathcal{F}_S\wedge\mathcal{A}_S + g^{-1}\mathrm{d}\mathcal{A}_S\wedge\mathrm{d}g + g^{-1}\mathcal{A}_S^2\wedge\mathrm{d}g) \\
&= \mathrm{tr}\Big(\mathcal{F}_S\wedge\mathcal{A}_S + \mathrm{d}(g^{-1}\mathcal{A}\wedge\mathrm{d}g) - \mathrm{d}(g^{-1})\wedge\mathcal{A}_S\wedge\mathrm{d}g + g^{-1}\mathcal{A}_S^2\wedge\mathrm{d}g\Big) \\
&= \mathrm{tr}\Big(\mathcal{F}_S\wedge\mathcal{A}_S + \mathrm{d}(g^{-1}\mathcal{A}\wedge\mathrm{d}g) + g^{-1}\mathrm{d}gg^{-1}\wedge\mathcal{A}_S\wedge\mathrm{d}g + g^{-1}\mathcal{A}_S^2\wedge\mathrm{d}g\Big),
\end{aligned}$$

where we used (A.2.5) in the last equality and the fact that the trace is cyclic, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, in the second equality, which allowed us to eliminate some occurrences of g and g^{-1}. Furthermore, we have

$$\begin{aligned}
\frac{1}{3}\mathrm{tr}(\mathcal{A}_N^3) &= \frac{1}{3}\mathrm{tr}\Big((g^{-1}\mathcal{A}_S g + g^{-1}\mathrm{d}g)^3\Big) \\
&= \mathrm{tr}\left(\frac{1}{3}\mathcal{A}_S^3 + g^{-1}\mathcal{A}_S^2\wedge\mathrm{d}g + g^{-1}\mathrm{d}g\wedge g^{-1}\mathcal{A}_S\wedge\mathrm{d}g + \frac{1}{3}(g^{-1}\mathrm{d}g)^3\right),
\end{aligned}$$

where the terms with a factor 3 are obtained by using (A.1.7). Then we find

$$\begin{aligned}
\mathrm{n} &= -\frac{1}{8\pi^2}\int_{S^3}\mathrm{tr}\left(\mathrm{d}(g^{-1}\mathcal{A}_S\wedge\mathrm{d}g) + (g^{-1}\mathrm{d}g)^3\right) \\
&= -\frac{1}{8\pi^2}\int_{S^3}\mathrm{tr}(g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g\wedge g^{-1}\mathrm{d}g) \\
&= -n(g),
\end{aligned}$$

where we used that S^3 is a manifold without boundary in the second equality. □

Corollary 3.5.8. Let $P(S^4, G)$ and $P'(S^4, G)$ be principal bundles and let $\mathcal{A}$ be an n-instanton on P and $\mathcal{A}'$ an n'-instanton on P' such that $n \neq n'$. Then P and P' are inequivalent principal bundles.

3.6 BPST instantons

We shall now investigate solutions of the self-dual equations of topological charge -1 in the Euclidean space $\mathbb{R}^4$ with structure group SU(2). The solution relies heavily on the properties of the Pauli matrices σ_i. Furthermore, we shall use quaternions in order to describe the solutions. This is not necessary, a description of the -1-instanton solutions without using quaternions can be found [34] for example, but the use of quaternions will be helpful in the ADHM construction, which we will discuss below.

Lorentz generators

We start by defining the $\sigma_{\mu\nu}$-matrices, which are 2×2 representatives of the Lorentz generators in Euclidian space.

Definition 3.6.1. $\sigma_{\mu\nu}$ and $\bar{\sigma}_{\mu\nu}$ are defined as follows:

$$\sigma_{\mu\nu} = \frac{1}{2}(\sigma_\mu\sigma_\nu^\dagger - \sigma_\nu\sigma_\mu^\dagger) \tag{3.6.1}$$
$$\bar{\sigma}_{\mu\nu} = \frac{1}{2}(\sigma_\mu^\dagger\sigma_\nu - \sigma_\nu^\dagger\sigma_\mu). \tag{3.6.2}$$

where σ_i are the *Pauli matrices* and $\sigma_4 = iI$. Pauli matrices are Hermitian; $\sigma_i^\dagger = \sigma_i$, while σ_4 is anti-Hermitian; $\sigma_4^\dagger = -\sigma_4$. Furthermore, the σ-matrices are unitary, $\sigma_\mu^\dagger = \sigma_\mu^{-1}$, and

$$\frac{1}{2}(\sigma_\mu\sigma_\nu^\dagger + \sigma_\nu\sigma_\mu^\dagger) = g_{\mu\nu}I_2, \tag{3.6.3}$$

which can be derived from $\{\sigma_i, \sigma_j\} = 2g_{ij}I_2$ and the anti-Hermiticity of σ_4.

Proposition 3.6.2. We $\sigma_{\mu\nu}$ is anti-self-dual, while $\bar{\sigma}_{\mu\nu}$ is self-dual.

$$\frac{1}{2}\epsilon_{\mu\nu\rho\tau}\sigma_{\rho\tau} = -\sigma_{\mu\nu} \tag{3.6.4}$$
$$\frac{1}{2}\epsilon_{\mu\nu\rho\tau}\bar{\sigma}_{\rho\tau} = \bar{\sigma}_{\mu\nu}. \tag{3.6.5}$$

Notice that $\sigma_{\mu\nu}$ is self-dual if we replace σ_4 by σ_0 and use the oriented basis $\{x^0, x^1, x^2, x^3\}$ instead of $\{x^1, x^2, x^3, x^4\}$. Similarly $\bar{\sigma}_{\mu\nu}$ is anti-self-dual in that case. We shall use the basis $\{x^1, x^2, x^3, x^4\}$ from now on. Since $-i\sigma_1, -i\sigma_2, -i\sigma_3$ form a basis of $\mathfrak{su}(2)$, it is convenient to assign a special symbol to these elements. Hence we define

$$\tau_\mu = -i\sigma_\mu.$$

Notice that this means that $\tau_4 = I_2$. We give are now ready to introduce quaternions.

Quaternions

Definition 3.6.3. The algebra of *quaternions* is defined as

$$\mathbb{H} = \{ix^1 + jx^2 + kx^3 + x^4 : x^1, x^2, x^3, x^4 \in \mathbb{R}\},$$

where $i^2 = j^2 = k^2 = ijk = -1$. The set $\{\pm 1, \pm i, \pm j, \pm k\}$ form a group, called the *quaternion group*. Notice that $\mathbb{H}$ is a 4-dimensional vector space over $\mathbb{R}$. If $x = x^0 + ix^1 + jx^2 + kx^3 \in \mathbb{H}$, we define the *quaternionic conjugate* $\bar{x}$ of x by

$$\bar{x} = -ix^1 - jx^2 - kx^3 + x^4.$$

We define the *quaternionic imaginary part* of x by

$$\text{Im } x = x - x^4 = ix^1 + jx^2 + kx^3.$$

We define the *absolute value* of a quarternion x by $|x| = \sqrt{x\bar{x}}$.

Remarks.

1. Contrary to the definition for complex numbers we have Im $x \notin \mathbb{R}$.

2. We can easily calculate the imaginary part by the identity Im $x = \frac{1}{2}(x - \bar{x})$.
3. Conjugation for quaternions is an anti-involution: $\overline{xy} = \bar{y}\bar{x}$.

By setting $x^2 = x^3 = 0$ we can identify $\mathbb{C}$ as a subalgebra of $\mathbb{H}$. Moreover, we have a group isomorphism induced by $(i, j, k, 1) \mapsto (-i\sigma_1, -i\sigma_2, -i\sigma_3, I_2) = (\tau_1, \tau_2, \tau_3, \tau_4)$, which allows us to write

$$x = x^\mu \tau_\mu.$$

In this notation we have $\bar{x} = x^\mu \tau_\mu^\dagger$. If we define complex coordinates z_1, z_2 by

$$z_1 = x_2 + ix_1$$
$$z_2 = x_4 + ix_3$$

we have

$$x = \begin{pmatrix} \bar{z}_2 & -z_1 \\ \bar{z}_1 & z_2 \end{pmatrix}, \qquad \bar{x} = \begin{pmatrix} z_2 & z_1 \\ -\bar{z}_1 & \bar{z}_2 \end{pmatrix} \tag{3.6.6}$$

Remark. We see that $1 \in \mathbb{H}$ corresponds with $x = x^\mu \tau_\mu$ with $x_1 = x_2 = x_3 = 0$ and $x_4 = 1$, and with $x = (z_1, z_2) \in \mathbb{C}^2$ with $z_1 = 0$ and $z_2 = 1$. Especially the identification with the element in $\mathbb{C}^2$ seems to be a little bit odd, but it turns out to be convenient.

Since $\{\tau_1, \tau_2, \tau_3\}$ is the set of generators of $\mathfrak{su}(2)$, we can identify latter Lie algebra with pure imaginary quaternions with basis $\{i, j, k\}$. Hence we can write SU(2)-gauge potential as

$$\mathcal{A}|_x = \mathcal{A}_\mu(x)\mathrm{d}x^\mu,$$

with $\mathcal{A}_\mu$ taking values in Im $\mathbb{H}$.

It is convenient to rewrite the action as

$$S_E[\mathcal{A}] = \frac{1}{2}\int_M |\mathcal{F}|^2 \mathrm{dVol}(g), \tag{3.6.7}$$

where the factor $\frac{1}{2}$ arises by choosing the normalization factor λ (defined in the remark below Definition A.4.3) equal to $\frac{1}{2}$, so that we have $|a\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu| = 1$ with $a = i, j, k$. This means that in case of $\mathcal{F}$ being (anti-)self-dual, we have

$$\mathrm{n} = \frac{1}{4\pi^2}\int_M |\mathcal{F}|^2 \mathrm{dVol}(g). \tag{3.6.8}$$

Definition 3.6.4. We define the *quaternionic differential* by

$$\mathrm{d}x = i\mathrm{d}x^1 + j\mathrm{d}x^2 + k\mathrm{d}x^k + \mathrm{d}x^4.$$

Its conjugate is defined by

$$\mathrm{d}\bar{x} = -i\mathrm{d}x^1 - j\mathrm{d}x^2 - k\mathrm{d}x^k + \mathrm{d}x^4.$$

We easily see that we can represent all SU(2)-gauge potential $\mathcal{A}|_x$ by

$$\mathcal{A}|_x = \mathrm{Im}(f(x)\mathrm{d}x) = \frac{1}{2}\left(f(x)\mathrm{d}x - \mathrm{d}\bar{x}\overline{f(x)}\right), \tag{3.6.9}$$

where $f:\mathbb{H}\to\mathbb{H}$ is a function. Sometimes it is more convenient to represent $\mathcal{A}|_x$ by

$$\mathcal{A}|_x = \mathrm{Im}(\tilde{f}(x)\mathrm{d}\bar{x}) \tag{3.6.10}$$

with $\tilde{f}:\mathbb{H}\to\mathbb{H}$. Both representations are equivalent, since $\tilde{f}=-\bar{f}$ gives the same gauge potential. Now let us denote $f(x)\mathrm{d}x$ as $a_\mu \mathrm{d}x^\mu$ with $a_\mu\in\mathbb{H}$. Thus for all μ we have $a_\mu = ia_\mu^1 + ja_\mu^2 + ka_\mu^3 + a_\mu^4$, and so Im $a_\mu = a_\mu - a_\mu^0$. Then we find

$$\begin{aligned}
\mathrm{Im}(f\mathrm{d}x)\wedge\mathrm{Im}(f\mathrm{d}x) &= \mathrm{Im}(a_\mu)\mathrm{Im}(a_\nu)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= (a_\mu - a_\mu^4)(a_\nu - a_\nu^4)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= (a_\mu a_\nu - a_\mu a_\nu^4 - a_\mu^4 a_\nu + a_\mu^4 a_\nu^4)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= a_\mu a_\nu \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu,
\end{aligned}$$

since $a_\mu^4\in\mathbb{R}$ commutes with all other object, so $-a_\mu a_\nu^4 - a_\mu^4 a_\nu + a_\mu^4 a_\nu^4$ is antisymmetric in μ and ν, so vanishes under contraction with $\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$. On the other hand, we have

$$\begin{aligned}
\mathrm{Im}(f\mathrm{d}x\wedge f\mathrm{d}x) &= \mathrm{Im}(a_\mu a_\nu)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= \Big(a_\mu a_\nu - (a_\mu a_\nu)^4\Big)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= \Big(a_\mu a_\nu + a_\mu^1 a_\nu^1 + a_\mu^2 a_\nu^2 + a_\mu^3 a_\nu^3 - a_\mu^4 a_\nu^4\Big)\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\\
&= a_\mu a_\nu \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu,
\end{aligned}$$

since $a_\mu^1 a_\nu^1 + a_\mu^2 a_\nu^2 + a_\mu^3 a_\nu^3 - a_\mu^4 a_\nu^4$ is clearly symmetric in μ and ν. So we see that $\mathrm{Im}(f\mathrm{d}x)\wedge\mathrm{Im}(f\mathrm{d}x) = \mathrm{Im}(f\mathrm{d}x\wedge f\mathrm{d}x)$, whence by (2.2.11), the corresponding gauge-field strength $\mathcal{F}$ of $\mathcal{A}$ is given by

$$\mathcal{F}|_x = \mathrm{Im}\Big(\mathrm{d}f(x)\wedge\mathrm{d}x + f(x)\mathrm{d}x\wedge f(x)\mathrm{d}x\Big). \tag{3.6.11}$$

Proposition 3.6.5. The volume form $\mathrm{d}x\wedge\mathrm{d}\bar{x}$ is anti-self-dual, while the form $\mathrm{d}\bar{x}\wedge\mathrm{d}x$ is self-dual. Both volume forms are purely imaginary.

Proof. We have $\mathrm{d}x\wedge\mathrm{d}\bar{x} = \tau_\mu\tau_\nu^\dagger \mathrm{d}x^\mu\wedge\mathrm{d}x^\nu = -\sigma_\mu\sigma_\nu^\dagger\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$. Since $\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$ is antisymmetric in μ and ν, we may replace $\sigma_\mu\sigma_\nu^\dagger$ by $\frac{1}{2}(\sigma_\mu\sigma_\nu^\dagger - \sigma_\nu\sigma_\mu^\dagger) = \sigma_{\mu\nu}$ which is anti-self-dual. We could also give a proof by direct calculation, which yields an expression which is clearly purely imnaginary:

$$\begin{aligned}
\mathrm{d}x\wedge\mathrm{d}\bar{x} &= (i\mathrm{d}x^1 + j\mathrm{d}x^2 + k\mathrm{d}x^3 + \mathrm{d}x^4)\wedge(-i\mathrm{d}x^1 - j\mathrm{d}x^2 - k\mathrm{d}x^3 + \mathrm{d}x^4)\\
&= 2(\mathrm{d}x^1\wedge\mathrm{d}x^4 - \mathrm{d}x^2\wedge\mathrm{d}x^3)i + 2(\mathrm{d}x^2\wedge\mathrm{d}x^4 - \mathrm{d}x^1\wedge\mathrm{d}x^3)j + 2(\mathrm{d}x^3\wedge\mathrm{d}x^4 - \mathrm{d}x^1\wedge\mathrm{d}x^2)k.
\end{aligned}$$

Using (A.4.8), we see that every term is anti-self-dual, so the whole form is anti-self-dual. Similarly, we find

$$\begin{aligned}
\mathrm{d}\bar{x}\wedge\mathrm{d}x &= (-i\mathrm{d}x^1 - j\mathrm{d}x^2 - k\mathrm{d}x^3 + \mathrm{d}x^4)\wedge(i\mathrm{d}x^1 + j\mathrm{d}x^2 + k\mathrm{d}x^3 + \mathrm{d}x^4)\\
&= -2(\mathrm{d}x^1\wedge\mathrm{d}x^4 + \mathrm{d}x^2\wedge\mathrm{d}x^3)i - 2(\mathrm{d}x^2\wedge\mathrm{d}x^4 + \mathrm{d}x^1\wedge\mathrm{d}x^3)j - 2(\mathrm{d}x^3\wedge\mathrm{d}x^4 + \mathrm{d}x^1\wedge\mathrm{d}x^2)k,
\end{aligned}$$

which is clearly self-dual. □

The (−1)-instanton

Proposition 3.6.6. The gauge field

$$\mathcal{A}|_x = \mathrm{Im}\left(\frac{(x-b)\mathrm{d}\bar{x}}{\rho^2 + |x-b|^2}\right) \tag{3.6.12}$$

with $\rho^2\in\mathbb{R}$ and $b\in\mathbb{H}$ is anti-self-dual.

Proof. First we calculate the corresponding gauge field-strength:

$$\begin{aligned}
\mathcal{F}|_x &= \mathrm{Im}\left(\frac{\mathrm{d}(x-b)\wedge \mathrm{d}\bar{x}}{\rho^2+|x-b|^2} + (x-b)\mathrm{d}\left(\frac{1}{\rho^2+|x-b|^2}\right)\wedge \mathrm{d}\bar{x} + \frac{(x-b)\mathrm{d}\bar{x}\wedge(x-b)\mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2}\right)\\
&= \mathrm{Im}\left(\frac{\mathrm{d}x\wedge \mathrm{d}\bar{x}}{\rho^2+|x-b|^2} - \frac{(x-b)\mathrm{d}\bar{x}\wedge(x-b)\mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2} - \frac{(x-b)(\bar{x}-\bar{b})\mathrm{d}x\wedge \mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2} + \frac{(x-b)\mathrm{d}\bar{x}\wedge(x-b)\mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2}\right)\\
&= \mathrm{Im}\left((\rho^2+|x-b|^2)\frac{\mathrm{d}x\wedge \mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2} - \frac{|x-b|^2\mathrm{d}x\wedge \mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2}\right)\\
&= \mathrm{Im}\left(\frac{\rho^2\mathrm{d}x\wedge \mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2}\right)\\
&= \frac{\rho^2\mathrm{d}x\wedge \mathrm{d}\bar{x}}{(\rho^2+|x-b|^2)^2},
\end{aligned}$$

where the last step follows from Proposition 3.6.5. By the same proposition, we see that $\mathcal{A}$ is anti-self-dual. ☐

By interchanging x and $\bar{x}$ in (3.6.12), we see that the gauge potential

$$\mathcal{A} = \mathrm{Im}\left(\frac{(\bar{x}-\bar{b})\mathrm{d}x}{\rho^2+|x-b|^2}\right) \tag{3.6.13}$$

is self-dual, since its corresponding gauge field-strength is given by

$$\mathcal{F} = \frac{\rho^2\mathrm{d}\bar{x}\wedge \mathrm{d}x}{(\rho^2+|x-b|^2)^2}. \tag{3.6.14}$$

Proposition 3.6.7. The instanton defined in (3.6.13) has instanton number $\mathrm{n} = 1$.

Proof. We use the abbreviation $\mathrm{d}x^{\mu\nu}$ for $\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$, and the notation $|\alpha|_g^2 = g(\alpha,\alpha)$ for ordinary k-forms α. Then using (3.6.8) and the expression for $\mathrm{d}\bar{x}\wedge\mathrm{d}x$ we found in Proposition 3.6.5, we find

$$\begin{aligned}
|\mathrm{n}| &= \frac{1}{4\pi^2}\int_{\mathbb{R}^4}|\mathcal{F}|^2\mathrm{d}^4x\\
&= \frac{1}{4\pi^2}\int_{\mathbb{R}^4}\frac{4\rho^4}{(\rho^2+|x-b|^2)^2}\Big|(\mathrm{d}x^{14}+\mathrm{d}x^{23})i + (\mathrm{d}x^{24}+\mathrm{d}x^{13})j + (\mathrm{d}x^{34}+\mathrm{d}x^{12})k\Big|^2\mathrm{d}^4x\\
&= \frac{1}{4\pi^2}\int_{\mathbb{R}^4}\frac{4\rho^4}{(\rho^2+|x-b|^2)^2}\Big(|\mathrm{d}x^{14}+\mathrm{d}x^{23})|_g^2 + |\mathrm{d}x^{24}+\mathrm{d}x^{13})|_g^2 + |\mathrm{d}x^{34}+\mathrm{d}x^{12}|_g^2\Big)\mathrm{d}^4x\\
&= \frac{1}{4\pi^2}\int_{\mathbb{R}^4}\frac{24\rho^4}{(\rho^2+|x-b|^2)^2}\mathrm{d}^4x,
\end{aligned}$$

where we used that $\mathrm{d}x^{14}$ and $\mathrm{d}x^{23}$ are normalized basis vectors, which are perpendicular to each other, so that we have $|\mathrm{d}x^{14}+\mathrm{d}x^{23}|_g^2 = |\mathrm{d}x^{14}|_g^2 + |\mathrm{d}x^{23}|_g^2 = 2$. Similarly the squared length of the other self-dual 2-forms also equals 2. Now, by translation invariance of the integral, we can replace $|x-b|$ by $|x$, while using spherical polar coordinates gives

$$\begin{aligned}
|\mathrm{n}| &= \frac{6}{\pi^2}\mathrm{Vol}(S^3)\int_0^\infty\frac{r^3\mathrm{d}r}{(1+r^2)^4}\\
&= \frac{1}{2\pi^2}\mathrm{Vol}(S^3)\\
&= 1,
\end{aligned}$$

where we used that the volume of the unit sphere in $\mathbb{R}^4$ is $2\pi^2$. ☐

3.7 The moduli space

Convention. In this section we assume that M is a connected 4-manifold and G is a compact connected matrix group with semisimple Lie algebra $\mathfrak{g}$.

We recall that we denote the space of connections with $\mathscr{A}$, the space of irreducible connections with $\mathscr{A}_o$ and the space of self-dual connections with $\mathscr{A}^+$. We denote the space of irreducible self-dual connections $\mathscr{A}_o \cap \mathscr{A}^+$ with $\mathscr{A}_o^+$. Furthermore, we have $\tilde{\mathcal{G}} = \mathcal{G}/Z(\mathcal{G})$. For $\omega \in \mathscr{A}^+$, we can identify $T_\omega\mathscr{A}$ with $\Omega^1(M, \mathrm{ad}\ P)$, since $\mathscr{A}$ is an affine space over $\Omega^1(M, \mathrm{ad}\ P)$.

We define the moduli space to be the space of all self-dual connections up to gauge transformation. That is

Definition 3.7.1. Let $\mathscr{M} = \mathscr{A}^+/\tilde{\mathcal{G}}$. Then $\mathscr{M}$ is called the *moduli space* of the bundle $P(M, G)$.

The purpose of this paragraph is to determine the dimension of the moduli space. We will do this by calculation the dimension of its tangent space $T_{\mathcal{A}}\mathscr{M}$ with $\mathcal{A} \in \mathscr{A}^+$, so that we can conclude that $\mathscr{M}$ must have the same dimension. This last step can be made only if $\mathscr{M}$ is a finite dimensional manifold, a fact which we will not proof, but later on we will give the required conditions instead.

Lemma 3.7.2. The Lie algebra $\mathfrak{G} = T_{\mathrm{Id}}\mathcal{G}$ of $\mathcal{G}$ is equal to $\Gamma^\infty(M, \mathrm{ad}\ P) = \Omega^0(M, \mathrm{ad}\ P)$.

Proof. By the map H defined in Theorem 2.4.15, we see that $\mathcal{G}$ can be identified with $\Gamma^\infty(M, \mathrm{Ad}\, P)$. Since the fibres of Ad P are equal to G, while the fibres of ad P are equal to $\mathfrak{g}$, we shall use the notations $\pi_{\mathrm{Ad}}^{-1}(x) = G_x$ and $\pi_{\mathrm{ad}}^{-1}(x) = \mathfrak{g}_x$. Then we have Ad $P = \bigcup_{x\in M} G_x$ and ad $P = \bigcup_{x\in M} \mathfrak{g}_x$. We have by Theorem 1.5.6 that $([p, e])$ is the identity of $\pi_{\mathrm{Ad}}^{-1}(x)$ for all $p \in \pi^{-1}(x)$, so we find

$$\mathfrak{g}_x = T_e(G_x) = T_{[(p,e)]}(\pi_{\mathrm{Ad}}^{-1}(x)) = V_{[(p,e)]}\mathrm{Ad}\ P,$$

where we used Proposition 2.1.2 in the last equality. By the second remark below Lemma 2.4.14, we have that $[(s_i(x), e)]$ is the identity element $e(x)$ of $\Gamma^\infty(M, \mathrm{Ad}\ P)$ with $s_i : U_i \to P$ the canonical section induced by the trivialization ψ_i. Since we have by definition of a section that $s_i(x) \in \pi^{-1}(x)$, we find that $V_{[(p,e)]}\mathrm{Ad}\ P = V_{e(x)}\mathrm{Ad}\ P$.

Now, let $s \in \Gamma^\infty(M, \mathrm{Ad}\ P)$ be a section and let I me a small interval around 0 such that $\gamma : I \to \Gamma^\infty(M, \mathrm{Ad}\, P)$ is a curve with $\gamma(0) = s$. In other words, we have $\gamma'(0) \in T_s\Gamma^\infty(M, \mathrm{Ad}\, P)$. Then, for all $x \in M$, $\gamma_x(t) = \gamma(t)(x)$ is a curve in $\pi_{\mathrm{Ad}}^{-1}(x)$. Hence $\gamma_x'(0) \in T_{\gamma_x(0)}(\pi_{\mathrm{Ad}}^{-1}(x)) = T_{s(x)}(\pi^{-1}(x)) = V_{s(x)}\mathrm{Ad}\ P$ (again using Proposition 2.1.2).

Now if we take $s(x) = e(x)$, we find $\gamma_x'(0) \in \mathfrak{g}_x$ and so $\gamma'(0) : M \to \mathrm{ad}\ P = \bigcup_{x\in M} \mathfrak{g}_x$ is a section. Thus $\gamma'(0) \in \Gamma^\infty(M, \mathrm{ad}\ P)$ if and only if $\gamma'(0) \in T_{\mathrm{Id}}\mathcal{G}$, which proves the lemma. □

Definition 3.7.3. We define the exponential map $\exp : \mathfrak{G} \to \mathcal{G}$ fibrewise, that is if $\Theta \in \mathfrak{G}$ is described by local functions $\{\theta_i : U_i \to \mathfrak{g}\}$ (not to be confused with the Maurer-Cartan form), we define $\exp(t\Theta) \in \Gamma^\infty(M, \mathrm{Ad}\ P)$ to be the function described by the family of local functions $\{\exp(t\theta_i) : U_i \to G\}$, where exp denotes the matrix exponent. We denote the element $H^{-1}\Big(\exp(t\Theta)\Big) \in \mathcal{G}$ with Φ_t, where H is the isomorphism from Theorem 2.4.15.

Remarks.

1. Since we have $\exp(\mathrm{ad}_C X) = \mathrm{Ad}_C \exp(X)$ for $n \times n$-matrices X and C with the latter invertible (see for instance Proposition 2.3 of [18] for a proof), we see that $\exp(t\theta_i(x)) = \mathrm{Ad}_{g_{ij}(x)} \exp(t\theta_j(x))$ if and only if $\theta_i(x) = \mathrm{ad}_{g_{ij}(x)}\theta_j(x)$, so Φ_t is well-defined.

2. Using the homomorphism $H_i : \mathfrak{G} \to C^\infty(U_i, G)$ defined in Proposition 2.4.12, we see by Theorem 2.4.15 that $H_i(\Phi_t) = \exp(t\theta_i)$.

Proposition 3.7.4. Let $\mathcal{A} = \{\mathcal{A}_i\} \in \mathscr{A}$. Then $\mathrm{im}\{\mathcal{D}_\mathcal{A} : \Omega^0(M, \mathrm{ad}\ P) \to \Omega^1(M, \mathrm{ad}\ P)\}$ equals the tangent space $T_\mathcal{A}(\mathcal{G}\cdot\mathcal{A})$ of the orbit of $\mathcal{A}$ under the action of the gauge group.

Proof. Let $\Theta \in \mathfrak{G} = \Gamma^\infty(M, \mathrm{ad}\ P)$ and let $\Phi_t = \exp(t\Theta)$. Then by Theorem 2.4.26 the gauge transform of $\mathcal{A}_i$ is given by

$$\Phi_t(\mathcal{A}_i) = \exp(-t\theta_i)\mathcal{A}_i\exp(t\theta_i) + \exp(-t\theta_i)\mathrm{d}\exp(t\theta_i) \in \Omega^1(U_i, \mathfrak{g}). \tag{3.7.1}$$

We have by definition that $\Phi_t(\mathcal{A}_i)$ is a curve in $\mathcal{G}\cdot\mathcal{A}$, while $\Phi_0(\mathcal{A}_i) = \mathcal{A}_i$. So we see that that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mathcal{A}_i)\Big|_{t=0} \in T_\omega(\mathcal{G}\cdot\mathcal{A})$. Since an explicit calculation gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mathcal{A}_i) =& -\exp(-t\theta_i)\theta_i\mathcal{A}_i\exp(t\theta_i) + \exp(-t\theta_i)\mathcal{A}_i\theta_i\exp(t\theta_i)\\ &-\theta_i\exp(-t\theta_i)\mathrm{d}\exp(t\theta_i) + \exp(-t\theta_i)\mathrm{d}\Big(\theta_i\exp(t\theta_i)\Big)\\ =&\exp(-t\theta_i)\Big(\mathcal{A}_i\theta_i - \theta_i\mathcal{A}_i + \mathrm{d}\theta\Big)\exp(t\theta_i)\\ =&\Phi_{-t}\mathcal{D}_{\mathcal{A}_i}\theta_i\Phi_t,\end{aligned} \tag{3.7.2}$$

where we used that $\exp(t\theta_i)$ commutes with θ_i, since both are 0-forms and e^A always commutes with A for all $n\times n$-matrices A. Furthermore, in the second equality we used Proposition A.2.2 in order to expand $\mathrm{d}\Big(\theta_i\exp(t\theta_i)\Big)$. If we evaluate this derivative in $t = 0$, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mathcal{A}_i)\Bigg|_{t=0} = \mathcal{D}_\mathcal{A}\theta_i, \tag{3.7.3}$$

we find that $\mathcal{D}_\mathcal{A}\Theta \in T_\mathcal{A}(\mathcal{G}\cdot\mathcal{A})$. ☐

By Proposition 2.5.17, we have $\mathcal{G}\cdot\mathcal{A} \simeq \tilde{\mathcal{G}}\cdot\mathcal{A}$ for irreducible connections $\mathcal{A}$, hence we have the following corollary:

Corollary 3.7.5. Let $\mathcal{A} \in \mathscr{A}_o$. Then $\mathrm{im}\{\mathcal{D}_\mathcal{A} : \Omega^0(M, \mathrm{ad}\ P) \to \Omega^1(M, \mathrm{ad}\ P)\}$ equals $\mathcal{T}_\mathcal{A}(\tilde{\mathcal{G}}\cdot\mathcal{A})$.

Definition 3.7.6. We define $\mathcal{D}^-_\mathcal{A} : \Omega^1(M, \mathrm{ad}\ P) \to \Omega^2_-(M, \mathrm{ad}\ P)$, the space of anti-self-dual forms, by $\mathcal{D}^-_\mathcal{A}\tau = (\mathcal{D}_\mathcal{A}\tau)^-$, where $\tau \in \Omega^1(M, \mathrm{ad}\ P)$.

Lemma 3.7.7. Let $\mathcal{A}$ be a self-dual connection. Then we have $\mathcal{D}^-_\mathcal{A}\circ\mathcal{D}_\mathcal{A} = 0$.

Proof. By definition of $\mathcal{D}^-_\mathcal{A}$ and (2.3.8) we find $\mathcal{D}^-_\mathcal{A}\circ\mathcal{D}_\mathcal{A} = \mathcal{F}^-_\mathcal{A} = 0$, since $\mathcal{A}$ is self-dual. ☐

Proposition 3.7.8. Let $\mathcal{A} = \{\mathcal{A}_i\} \in \mathscr{A}^+$. Then a $\tau \in \Omega^1(M, \mathrm{ad}\ P)$ is tangent to $\mathscr{A}^+$ if and only if $\mathcal{D}^-_\mathcal{A}\tau = 0$.

Proof. Define the curve $\mathcal{A}_t = \mathcal{A} + t\tau \in \mathscr{A}$, then by (3.1.6), we have $\mathcal{F}_t = \mathcal{F} + t\mathcal{D}_\mathcal{A}\tau + t^2\tau\wedge\tau$ and we see that $\mathcal{F}_t$ is self-dual up to first order if and only if $\mathcal{D}_\mathcal{A}\tau$ is self-dual, that is if $\mathcal{D}^-_\mathcal{A}\tau = 0$. So we see that $\mathcal{A}_t \in \mathscr{A}^+$ if and only $\mathcal{D}^-_\mathcal{A}\tau = 0$ and since $\mathcal{A}_0 = \mathcal{A}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}_t\Big|_{t=0} = \tau$, so we see that $\tau \in T_\mathcal{A}\mathscr{A}^+$ if and only if $\mathcal{D}^-_\mathcal{A}\tau = 0$. ☐

We can summarize all our results in next theorem.

Theorem 3.7.9. Let $\mathcal{A} \in \mathscr{A}^+$. Then the following diagram given by

$$0 \longrightarrow \Omega^0(M, \mathrm{ad}\ P) \xrightarrow{\mathcal{D}_\mathcal{A}} \Omega^1(M, \mathrm{ad}\ P) \xrightarrow{\mathcal{D}^-_\mathcal{A}} \Omega^2_-(M, \mathrm{ad}\ P) \longrightarrow 0.$$

is a complex, called the *deformation complex*, with $\mathrm{im}\{\mathcal{D}_\mathcal{A} : \Omega^0(M, \mathrm{ad}\ P) \to \Omega^1(M, \mathrm{ad}\ P)\} = T_\mathcal{A}(\mathcal{G}\cdot\mathcal{A})$ and $\ker\{\mathcal{D}^-_\mathcal{A} : \Omega^1(M, \mathrm{ad}\ P) \to \Omega^2(M, \mathrm{ad}\ P)\} = T_\mathcal{A}\mathscr{A}^+$.

Theorem 3.7.10. Let $M = S^4$ and $G = \mathrm{SU}(2)$. Then all self-dual connections with nonzero instanton number must be irreducible.

Proof (sketch). Let ω be a self-dual connection. Then if ω is reducible, the bundle over M splits (Theorem 3.1 of [16]) and there must be a circle S^1-subbundle over S^4. Since S^1-bundles are classified by $H_3(S^1)$, which is zero, we find that these bundles must be trivial. So ω must have zero instanton number. ☐

We see that for S^4 we have $\mathscr{A}^+ = \mathscr{A}_o^+$. Hence we have following corollary:

Corollary 3.7.11. Let $G = \mathrm{SU}(2)$. Then the tangent space of the moduli space $T_{\mathcal{A}}\mathscr{M}$ is isomorphic to the first twisted cohomology group $H^1_{\mathcal{A}}$.

Proof. Since $\mathscr{M} = \mathscr{A}_o^+/\tilde{\mathcal{G}} \simeq \mathscr{A}_o^+/\tilde{\mathcal{G}}\cdot\mathcal{A}$, we have $T_{\mathcal{A}}\mathscr{M} \simeq T_{\mathcal{A}}\mathscr{A}_o^+/T_{\mathcal{A}}(\tilde{\mathcal{G}}\cdot\mathcal{A}) = \ker\mathcal{D}^-_{\mathcal{A}}/\mathrm{im}\,\mathcal{D}_{\mathcal{A}}$, which is the first cohomology group $H^1_{\mathcal{A}}$. ☐

Lemma 3.7.12. Let G be a connected and compact matrix group with semisimple Lie algebra $\mathfrak{g}$ and let $\mathcal{A}$ be an irreducible connection. Then the zeroth twisted cohomology group $H^0_{\mathcal{A}}$ vanishes.

Proof. We have $H^0_{\mathcal{A}} = \ker\mathcal{D}_{\mathcal{A}}$, which is nonzero if and only if there is a $\Theta \in \Omega^0(M, \mathrm{ad}\ P)$ such that $\mathcal{D}_{\mathcal{A}}\Theta = 0$. Let $\Theta \in \ker\mathcal{D}_{\mathcal{A}}$, then Θ defines a gauge transformation $\Phi_t = \exp(t\Theta)$, which acts on $\mathcal{A}$ by (3.7.1). We see by (3.7.2) that $\Phi_t(\mathcal{A})$ must be independent of t, so we are allowed to choose substitute $t = 0$ in the right-hand side of (3.7.1), whence we find $\Phi_t(\mathcal{A}) = \mathcal{A}$, and therefore $\Phi_t \in \Gamma_{\mathcal{A}} = \{\Phi \in \mathcal{G} : \Phi(\mathcal{A}) = \mathcal{A}\}$, the isotropy group of $\mathcal{A}$, for all t. Since $\Theta = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t\Big|_{t=0}$ and $\Phi_0 = \mathrm{Id}$, we see that $\Theta \in \mathrm{Lie}(\Gamma_{\mathcal{A}})$, so we have proven that $\ker\mathcal{D}_{\mathcal{A}} \subset \mathrm{Lie}(\Gamma_{\mathcal{A}})$. By Proposition 2.5.16 we find that $\ker\mathcal{D}_{\mathcal{A}} = \mathrm{Lie}(Z(G))$. From Paragraph 3.1 of [14], we find that $\mathrm{Lie}(Z(G)) = Z(\mathfrak{g}) := \{X \in \mathfrak{g} : [X,Y] = 0\ \forall Y \in \mathfrak{g}\}$, which is an abelian ideal of $\mathfrak{g}$. Since $\mathfrak{g}$ is semisimple, it follows that $Z(\mathfrak{g})$ and thus $\ker\mathcal{D}_{\mathcal{A}}$ must be zero. ☐

Lemma 3.7.13. On $M = S^4$ we have $H^2_{\mathcal{A}} = 0$.

Proof (sketch). Since $H^2_{\mathcal{A}} = \Omega^2_-(M, \mathrm{ad}\,P)/\mathrm{im}\,\mathcal{D}^-_{\mathcal{A}}$, it is isomorphic to the subspace of $\Omega^2_-(M, \mathrm{ad}\,P)$ orthogonal to the image of $\mathcal{D}^-_{\mathcal{A}}$. But this is exactly the kernel of $(\mathcal{D}^-_{\mathcal{A}})^*$. Since we have $\ker(\mathcal{D}^-_{\mathcal{A}})^* \subset \ker\mathcal{D}^-_{\mathcal{A}}(\mathcal{D}^-_{\mathcal{A}})^*$, it is sufficient to show that the kernel of $\mathcal{D}^-_{\mathcal{A}}(\mathcal{D}^-_{\mathcal{A}})^*$ is zero. This can be accomplished by showing that $\mathcal{D}^-_{\mathcal{A}}(\mathcal{D}^-_{\mathcal{A}})^*$ is a positive-definite operator. For the details we refer to Appendix B of [4]. ☐

Now we are able to state following theorem. We refer to [2], [13] or [16] for a proof.

Theorem 3.7.14. Let M be a manifold with $H^2_{\mathcal{A}} = 0$. Then the moduli space $\mathscr{M}$ is a finite dimensional smooth manifold.

Finally, we are able to state the following theorem:

Theorem 3.7.15. Let P be a SU(2)-bundle over S^4. Then the dimension of the moduli space $\mathscr{M}$ equals $8|\mathrm{n}| - 3$, where n is the instanton number.

Proof (sketch). Instead of the operators $\mathcal{D}_A$ and $\mathcal{D}^-_{\mathcal{A}}$, we consider the operator $\mathcal{D}^-_{\mathcal{A}} \oplus \mathcal{D}^*_A$: $\Omega^1(M, \mathrm{ad}\ P) \to \Omega^2_-(M, \mathrm{ad}\ P) \oplus \Omega^0(M, \mathrm{ad}\ P)$, which turns out to be an elliptic operator. Then we have

$$\begin{aligned}
\mathrm{Index}(\mathcal{D}^-_{\mathcal{A}} \oplus \mathcal{D}^*_A) &= \dim\ker(\mathcal{D}^-_{\mathcal{A}} \oplus \mathcal{D}^*_A) - \dim\ker(\mathcal{D}^-_{\mathcal{A}} \oplus \mathcal{D}^*_A)^* \\
&= \dim\ker\mathcal{D}^-_{\mathcal{A}} + \dim\ker\mathcal{D}^*_A - \dim\ker(\mathcal{D}^-_{\mathcal{A}})^* - \dim\ker\mathcal{D}_{\mathcal{A}} \\
&= -\dim H^2_{\mathcal{A}} + \dim H^1_{\mathcal{A}} - \dim H^0_{\mathcal{A}} \\
&= \dim H^1_{\mathcal{A}},
\end{aligned}$$

where the third equality follows from the fact that $H^0_{\mathcal{A}} = \ker \mathcal{D}_{\mathcal{A}}$, $H^1_{\mathcal{A}} = \ker \mathcal{D}^-_{\mathcal{A}}/\mathrm{im}\ \mathcal{D}_{\mathcal{A}} = \ker \mathcal{D}^-_{\mathcal{A}}/\ker(\mathcal{D}^-_{\mathcal{A}})^*$ and $H^2_{\mathcal{A}} = \ker(\mathcal{D}^-_{\mathcal{A}})^*$. Last equality follows from previous lemmas in this section. Using the Atiyah-Singer Index Theorem, we find that the index of $\mathcal{D}^-_{\mathcal{A}} \oplus \mathcal{D}^*_{\mathcal{A}}$ equals the topological index, which is $8|\mathrm{n}| - 3$ for $G = \mathrm{SU}(2)$ (see Table 8.1 of [2]). $\square$

4 The ADHM construction

In section 3.6 we found a 5-parameter solution with topological charge n $= -1$ for the anti-self-duality equation. In previous section we found that the total number of (anti)-self-dual solutions with topological charge n $= n$ must be $8|n| - 3$, so we see that for n $= -1$ we found all solutions. In this section we will discuss the ADHM construction, which helps us finding all $|8|n - 3$ solutions for arbitrary n.

4.1 Holomorphic vector bundles

Definition 4.1.1. Let E be a vector bundle over a complex manifold M. If the fibre V is a complex vector space and the structure group G equals $\mathrm{GL}(n, \mathbb{C})$, we call E a *complex vector bundle*. If $\pi : E \to M$ is holomorphic and E has a *holomorphic structure* $\mathcal{E}$, i.e. one can find biholomorphic trivializations $\psi_i : U_i \times V \to \pi^{-1}(U_i)$ for which the transition functions $g_{ij} : U_i \cap U_j \to \mathrm{GL}(n, \mathbb{C})$ are holomorphic, we call E a *holomorphic vector bundle*.

Let E be a vector bundle with holomorphic structure $\mathcal{E}$. In Appendix B, we introduced complex and Hermitian manifolds, and the operators $\partial : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M)$, $\bar{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$. Since the ordinary covariant derivative can be seen an extension of the exterior derivative, it is interesting to see how we can extent ∂ and $\bar{\partial}$ to operators acting on $\Omega^{p,q}(M, E)$, the space E-valued (p, q)-forms. Next proposition proves the existence of an operator $\partial_{\mathcal{E}}$ which is such an extension and which is compatible with the holomorphic structure $\mathcal{E}$.

Proposition 4.1.2. Let E be a holomorphic vector bundle with holomorphic structure $\mathcal{E}$. We have linear operators $\partial_E : \Omega^{0,q}(M, E) \to \Omega^{0,q+1}(M, E)$ and $\bar{\partial}_E : \Omega^{p,0}(M, E) \to \Omega^{p+1,0}(M, E)$ uniquely determined by the properties

1. If f is a complex-valued function and s a vector-valued section, $\partial_{\mathcal{E}}$ and $\bar{\partial}_{\mathcal{E}}$ satisfy the Leibnitz rule:
$$\partial_{\mathcal{E}}(fs) = (\partial f)s + f(\partial_{\mathcal{E}} s)$$
$$\bar{\partial}_{\mathcal{E}}(fs) = (\bar{\partial} f)s + f(\bar{\partial}_{\mathcal{E}} s).$$

2. $\partial_{\mathcal{E}} s$ and $\bar{\partial}_{\mathcal{E}} s$ vanish on an open subset $U \subset M$ if and only if s is anti-holomorphic respectively holomorphic over U.

Proof. We only give the proof for $\bar{\partial}_{\mathcal{E}}$, which we construct as follows. The second property implies that $\bar{\partial}_{\mathcal{E}}$ is local, which allows us to work in a local holomorphic trivialization U_i, in which sections s of E are represented by vector-valued functions $s_i \in \Omega^{0,q}(U_i, V)$ by Corollary 2.3.4. So we define $\bar{\partial}_{\mathcal{E}}$ by letting the ordinary $\bar{\partial}$ operator act on the separate components. We have to check that this is independent of the trivialization. By Proposition 1.5.10, we have that two representations s_i and s_j of a section s are related by a map g_{ij} into $\mathrm{GL}(n, \mathbb{C})$ such that $s_i = g_{ij} s_j$. Since E is a holomorphic vector bundle, g_{ij} is also holomorphic, so we have $\bar{\partial}(g_{ij}) = 0$, whence
$$\bar{\partial}(s_i) = \bar{\partial}(g_{ij} s_j) = \bar{\partial}(g_{ij}) s_j + g_{ij} \bar{\partial}(s_j) = g_{ij} \bar{\partial}(s_j).$$
So we see that $\bar{\partial}(s_i) = g_{ij} \bar{\partial}(s_j)$, hence $\bar{\partial}_{\mathcal{E}}$ is globally well-defined. The proof for $\partial_{\mathcal{E}}$ goes in a similar way. $\square$

Let E be a holomorphic vector bundle and $\mathcal{A}$ a connection on its associated principle bundle. By the projection operators defined in Definition B.2.5, we can decompose the operator $\mathrm{d}_{\mathcal{A}} : \Omega^0(M, E) \to \Omega^1(M, E)$ into holomorphic and anti-holomorphic parts: $\mathrm{d}_{\mathcal{A}} = \partial_{\mathcal{A}} \oplus \bar{\partial}_{\mathcal{A}}$, with

$\partial_{\mathcal{A}} : \Omega^0(M,E) \to \Omega^{1,0}(M,E)$ and $\bar{\partial}_{\mathcal{A}} : \Omega^0(M,E) \to \Omega^{0,1}(M,E)$. One could ask what the conditions are that $\bar{\partial}_{\mathcal{A}}$ is compatible with a holomorphic structure on E, that is that $\bar{\partial}_{\mathcal{A}} = \bar{\partial}_{\mathcal{E}}$.

We start by defining a so called *partial connection* on a complex vector bundle E over a complex manifold M with fibre V, which is a an operator $\bar{\partial}_\alpha : \Omega^0(M,E) \to \Omega^{0,1}(M,E)$ which satisfies the Leibniz rule given in Proposition 4.1.2. Analogous to (2.3.10), we have that $\bar{\partial}_\alpha$ can be expressed locally on a patch U_i as

$$(\bar{\partial}_\alpha)_i = \bar{\partial} + \alpha_i, \tag{4.1.1}$$

where $\alpha_i \in \Omega^{0,1}(U_i, V)$. Since partial connections satisfy only the first condition in Proposition 4.1.2, one could ask if it is possible to find another characterization of the second. This can indeed be done by looking to $\bar{\partial}_\alpha^2$.

Theorem 4.1.3. Let $\bar{\partial}_\alpha$ be a partial connection on a complex vector bundle E over a complex manifold M. Then $\bar{\partial}_\alpha$ is compatible with some holomorphic structure on E if and only if $\bar{\partial}_\alpha^2 = 0$.

The proof of this theorem can be found in section 2.2.2 of [13]. Note that if E is a vector bundle with holomorphic structure $\mathcal{E}$, we clearly have $\bar{\partial}_{\mathcal{E}}^2 = 0$, which is consistent with the theorem.

Corollary 4.1.4. Let E be a vector bundle with holomorphic structure $\mathcal{E}$ and let $\bar{\partial}_{\mathcal{A}}$ be the partial connection corresponding to the connection $\mathcal{A}$. Then $\bar{\partial}_{\mathcal{A}}$ is compatible with $\mathcal{E}$ if and only if $\mathcal{F}_{\mathcal{A}}^{0,2} = 0$, where $\mathcal{F}_{\mathcal{A}} = \mathcal{F}_{\mathcal{A}}^{2,0} + \mathcal{F}_{\mathcal{A}}^{1,1} + \mathcal{F}_{\mathcal{A}}^{0,2}$ is decomposed into different (p,q)-forms.

Proof. This follows directly from the fact that $\bar{\partial}_{\mathcal{A}}^2 = \mathcal{F}_{\mathcal{A}}^{0,2}$. □

Furthermore, we can specify the conditions for a partial connection to be induced from a connection. Before doing so, we introduce the concept of a Hermitian metric on a vector bundle.

Definition 4.1.5. Let E be a complex vector bundle over M with fibre V. Then a map $h_x : \pi^{-1}(x) \times \pi^{-1}(x) \to \mathbb{C}$ which depends smoothly on $x \in M$ is called a *Hermitian inner product* on the fibres if it satisfies

1. $h_x(u, av + bw) = ah_x(u,v) + bh_x(u,w)$ for all $u, v, w \in \pi^{-1}(x)$ and $a, b \in \mathbb{C}$;
2. $h_x(u,v) = \overline{h_x(v,u)}$ for all $u, v \in \pi^{-1}(x)$;
3. $h_x(u,v) \geq 0$ for all $u, v \in \pi^{-1}(x)$ and $h_x(u,u) = 0$ if and only if $u = 0$.

Definition 4.1.6. A complex vector bundle with a Hermitian inner product on the fibres is called a *Hermitian vector bundle*. A connection on the associated fibre bundle P of a Hermitian vector bundle E with Hermitian inner product $(\cdot,\cdot)$ is called *Hermitian* if it is compatible with the Hermitian inner product:

$$\mathrm{d}(\alpha_1, \alpha_2) = (\mathrm{d}_{\mathcal{A}}\alpha_1, \alpha_2) + (\alpha_1, \mathrm{d}_{\mathcal{A}}\alpha_2),$$

where $\alpha_1, \alpha_2 \in \Omega^1(M,E)$.

Lemma 4.1.7. Let $\mathcal{A}$ be a Hermitian connection. Then both $\mathcal{A}$ and $\mathcal{F}_{\mathcal{A}}$ are skew-Hermitian.

By the lemma we can prove that all partial connections come from a Hermitian connection.

Proposition 4.1.8. Let E be a Hermitian vector bundle over M, then for all partial connections $\bar{\partial}_\alpha$ on E, there is a unique Hermitian connection $\mathcal{A}$ such that $\bar{\partial}_{\mathcal{A}} = \bar{\partial}_\alpha$.

Proof. We work in a local unitary trivialization, so that we have $\bar{\partial}_\alpha = \bar{\partial}+\alpha$ with α a (0,1)-form. Since previous lemma requires that we must have $\mathcal{A} = -\mathcal{A}^*$, we have to choose $\mathcal{A} = \alpha - \alpha^*$. Hence we see that the (0,1)-component $\bar{\partial}_{\mathcal{A}}$ of $\mathrm{d}_{\mathcal{A}} = \mathrm{d} + \mathcal{A}$ is exactly $\bar{\partial}_\alpha$. □

The fact that also $\mathcal{F}_{\mathcal{A}}$ is skew-Hermitian leads to next proposition.

Proposition 4.1.9. Let $\mathcal{A}$ be a Hermitian connection on a Hermitian vector bundle over M. Then $\mathcal{A}$ is compatible with a holomorphic structure if and only if it has a field strength $\mathcal{F}_{\mathcal{A}}$ of type (1,1). In his case the connection is uniquely determined by the metric and the holomorphic structure.

Proof. Since the field strength $\mathcal{F}$ can be split into $\mathcal{F}^{2,0}_{\mathcal{A}}+\mathcal{F}^{1,1}_{\mathcal{A}}+\mathcal{F}^{0,2}_{\mathcal{A}}$, and Hermitian connections have skew-Hermitian field strengths, we must have $\mathcal{F}^{0,2}_{\mathcal{A}} = -(\mathcal{F}^{2,0}_{\mathcal{A}})^*$. By Corollary 4.1.4 we find that $\mathcal{A}$ is compatible with a holomorphic structure if and only if $\mathcal{F}^{0,2} = 0$, so also $\mathcal{F}^{2,0}_{\mathcal{A}}$ vanishes in that case, whence $\mathcal{F}_{\mathcal{A}} = \mathcal{F}^{1,1}_{\mathcal{A}}$. □

In Appendix B we introduced the Kähler form ω (not to be confused with the connection one-form), which is has bidegree (1,1). Hence if E is a Hermitian vector bundle over M we can decompose $\Omega^{1,1}(M,E) = \Omega^{1,1}_0 \oplus \Omega^0(M,E)\cdot\omega$, where $\Omega^{1,1}_0(M,E)$ is the space of forms pointwise orthogonal ω.

Lemma 4.1.10. The complexified self-dual forms over a complex manifold M of complex dimension 2 are given by

$$\Omega^+(M,E)^{\mathbb{C}} = \Omega^{2,0}(M,E) \oplus \Omega^0(M,E)\cdot\omega \oplus \Omega^{0,2}(M,E)$$

and the complexified anti-self-dual forms are

$$\Omega^-(M,E)^{\mathbb{C}} = \Omega^{1,1}_0(M,E).$$

Proof. Since $\dim_{\mathbb{C}} M = 2$, we use the model space $V = \mathbb{C}^2$ and endow M with complex coordinates $z^1 = ix^1 + x^2$ and $z^2 = ix^3 + x^4$. Then $\Omega^{2,0}(M,E)$ is spanned by

$$\mathrm{d}z^1 \wedge \mathrm{d}z^2 = -\mathrm{d}x^1 \wedge \mathrm{d}x^3 + \mathrm{d}x^2 \wedge \mathrm{d}x^4 + i(\mathrm{d}x^2 \wedge \mathrm{d}x^3 + \mathrm{d}x^1 \wedge \mathrm{d}x^4).$$

Similarly $\Omega^{0,2}(M,E)$ is spanned by

$$\mathrm{d}\bar{z}^1 \wedge \mathrm{d}\bar{z}^2 = -\mathrm{d}x^1 \wedge \mathrm{d}x^3 + \mathrm{d}x^2 \wedge \mathrm{d}x^4 - i(\mathrm{d}x^2 \wedge \mathrm{d}x^3 + \mathrm{d}x^1 \wedge \mathrm{d}x^4),$$

while we have on $\mathbb{C}^2$ that $g_{1\bar{2}} = g_{2\bar{1}} = 0$ and $g_{1\bar{1}} = g_{2\bar{2}} = 1$ so that by (B.3.1) we find

$$\omega = i\mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^1 + i\mathrm{d}z^2 \wedge \mathrm{d}\bar{z}^2 = -2(\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^3 \wedge \mathrm{d}x^4).$$

Now by (A.4.8), we have

$$\begin{aligned} *\mathrm{d}x^1 \wedge \mathrm{d}x^3 &= -\mathrm{d}x^2 \wedge \mathrm{d}x^4; \\ *\mathrm{d}x^1 \wedge \mathrm{d}x^4 &= \mathrm{d}x^2 \wedge \mathrm{d}x^3; \\ *\mathrm{d}x^1 \wedge \mathrm{d}x^2 &= \mathrm{d}x^3 \wedge \mathrm{d}x^4. \end{aligned}$$

Since $*^2 = 1$, we can move the star operator to the right-hand sides, so we see that $\Omega^{2,0}(M,E)$, $\Omega^{0,2}(M,E)$ and $\Omega^0(M,E)\cdot\omega$ span all self-dual forms. Since latter three spaces together with $\Omega^{1,1}_0(M,E)$ form a decomposition of $\Omega^2(M,E)^{\mathbb{C}}$, while $\Omega^+(M,E) \oplus \Omega^-(M,E)$ is another decomposition, we find that $\Omega^-(M,E)$ must be equal to $\Omega^{1,1}_0(M,E)$. □

We are now able to describe anti-self-dual connections in terms of holomorphic vector bundles.

Theorem 4.1.11. Let $\mathcal{A}$ be an anti-self-dual connection on a complex vector bundle E over the Hermitian manifold M. Then the operator $\bar{\partial}_{\mathcal{A}}$ defines a holomorphic structure on E. If $\mathcal{E}$ is a holomorphic structure on E and $\mathcal{A}$ is a compatible unitary connection, then $\mathcal{A}$ is anti-self-dual if and only if the component of the field strength $\mathcal{F}_{\mathcal{A}}$ along the Kähler form ω is zero: $(\mathcal{F}_{\mathcal{A}}, \omega) = 0$.

Proof. Assume that $\mathcal{A}$ is anti-self dual, so by previous lemma $\mathcal{F}_{\mathcal{A}} \in \Omega_0^{1,1}(M, E)$. Thus $\mathcal{F}_{\mathcal{A}}^{0,2} = 0$ and by Proposition 4.1.9 it follows that $\mathcal{A}$ is compatible with some holomorphic structure on E. For the second statement, let $\mathcal{A}$ be a unitary connection compatible with a holomorphic structure $\mathcal{E}$ on E. Then again by Proposition 4.1.9 we have that $\mathcal{F}_{\mathcal{A}} \in \Omega^{1,1}(M, E) = \Omega_0^{1,1}(M, E) \oplus \Omega^0(M, E) \cdot \omega$. Then $\mathcal{F}_{\mathcal{A}}$ is anti-self-dual if and only if $\mathcal{F}_{\mathcal{A}} \in \Omega_0^{1,1}(M, E)$, which is exactly the case if and only if $(\mathcal{F}_{\mathcal{A}}, \omega) = 0$. $\square$

Remark. If we restrict ourselves to $M = \mathbb{C}^2$, we find by

$$\mathcal{F}_{\mathcal{A}} = \frac{1}{2}(\mathcal{F}_{12}\mathrm{d}x^{12} + \mathcal{F}_{13}\mathrm{d}x^{13} + \mathcal{F}_{14}\mathrm{d}x^{14} + \mathcal{F}_{23}\mathrm{d}x^{23} + \mathcal{F}_{24}\mathrm{d}x^{24} + \mathcal{F}_{34}\mathrm{d}x^{34}),$$

with $\mathrm{d}x^{\mu\nu} = \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, and $\omega = -2(\mathrm{d}x^{12} + \mathrm{d}x^{34})$, which we found above, that the condition $(\mathcal{F}_{\mathcal{A}}, \omega) = 0$ is equivalent with the ASD equation (3.3.13). Furthermore, we easily see from the fact that $\Omega^{0,2}(M, E)$ is spanned by $-\mathrm{d}x^1 \wedge \mathrm{d}x^3 + \mathrm{d}x^2 \wedge \mathrm{d}x^4 - i(\mathrm{d}x^2 \wedge \mathrm{d}x^3 + \mathrm{d}x^1 \wedge \mathrm{d}x^4)$ that $\mathcal{F}_{\mathcal{A}}^{0,2} = 0$ is equivalent with the other two ASD equations (3.3.14) and (3.3.15). So latter two equations are equivalent with the condition that $\mathrm{d}_{\mathcal{A}}$ is compatible with the holomorphic structure.

4.2 Connections and projections

Let M be a smooth manifold and K and L complex vector spaces. Then a smooth map $R : M \to \mathrm{Hom}(K, L)$ is actually a family of linear maps R_x, and so R is a bundle map

$$R : \underline{K} \to \underline{L},$$

where $\underline{V}$ denotes the trivial vector bundle over M with fibre V. Assume that M is a complex manifold and that we have complex vector spaces K_0, K_1 and K_2 and holomorphic bundle maps

$$\underline{K}_0 \xrightarrow{\alpha} \underline{K}_1 \xrightarrow{\beta} \underline{K}_2,$$

such that $\beta \circ \alpha = 0$. In other words, we have a family of complexes

$$K_0 \xrightarrow{\alpha_x} K_1 \xrightarrow{\beta_x} K_2,$$

where α_x and β_x vary holomorphic with x. If α is injective and β surjective, we can define a vector bundle E with fibres $E_x = \ker \beta_x / \mathrm{im}\ \alpha_x$.

Proposition 4.2.1. Let E be the vector bundle defined as above. Then we can find a holomorphic structure $\mathcal{E}$ on E.

Proof. We can accomplish the proof by showing that all local sections $s : U \subset M \to E$ lift to a smooth section $s' : U \to \ker \beta$.

$$\begin{tikzcd} & \ker\beta \arrow[d] \\ U \arrow[ur, "s'"] \arrow[r, "s"'] & E \end{tikzcd} = \ker\beta/\mathrm{im}\ \alpha$$

Fix an $x_0 \in M$ and choose a $k_1 \in \ker\beta_{x_0}$. Let $P : K_2 \to K_1$ be a right inverse for β_{x_0} (which exists, since β is assumed to be surjective), then we shall prove that it is possible to find a holomorphic section of $\ker\beta$ of the form $s'(x) = k_1 + j(x)$, where $j(x_0) = 0$. Let $\eta : M \to K_2$ be given by $\eta_x = \beta_x - \beta_{x_0}$, so that

$$\beta_x = \beta_{x_0} + \eta_x. \tag{4.2.1}$$

Then $s'(x) \in \ker\beta_x$ if

$$(1 + P\eta_x)j_x = -P\eta_x(k_1), \tag{4.2.2}$$

since if we let β_{x_0} act on both sides, we obtain $\beta_{x_0}j_x + \eta_x j_x = -\eta_x(k_1)$. Then since $k_1 \in \ker\beta_{x_0}$, we can subtract $\beta_{x_0}k_1$ from the right-hand side, whence we have $\beta_x j_x = -\beta_x k_1$ using (4.2.1). But this is exactly $\beta_x s'(x) = 0$. If we choose U small enough, then $P\eta_x$ is small, whence we can invert $1 + P\eta_x$. Hence we can find a unique solution j_x to (4.2.2), which varies holomorphically with x, since η_x depends holomorphically on x. Using this method, we can find a set of local holomorphic sections of E, forming a basis for the fibres near x. $\square$

If we assume that the vector space K_i have Hermitian metrics, then the fibres E_x can be identified with the orthogonal complement of im $\alpha_x \subset \ker\beta_x$, that is $E_x = (\text{im } \alpha_x)^\perp \cap \ker\beta_x$. Now, we introduce the bundle map $R : \underline{K_1} \to \underline{K_2 \times K_0}$ given by

$$R_x = (\beta_x, \alpha_x^\dagger) = \begin{pmatrix} \beta_x \\ \alpha_x^\dagger \end{pmatrix}, \tag{4.2.3}$$

where R_x acts on $k \in K_1$ by $R_x k = (\beta_x k, \alpha_x^\dagger k)$. We clearly have $\ker R_x = \ker\alpha_x^\dagger \cap \ker\beta_x = (\text{im } \alpha_x)^\perp \cap \ker\beta_x = E_x$.

Proposition 4.2.2. The orthogonal projections P_α and P_β of K_1 on $\ker\alpha_x^\dagger = (\text{im } \alpha)^\perp$ respectively $\ker\beta_x$ are given by

$$\begin{aligned} P_x^\alpha &= 1 - \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger; \\ P_x^\beta &= 1 - \beta_x^\dagger(\beta_x\beta_x^\dagger)^{-1}\beta_x. \end{aligned}$$

The orthogonal projection $P_x : K_1 \to E_x$ is given by

$$P_x = 1 - \beta_x^\dagger(\beta_x\beta_x^\dagger)^{-1}\beta_x + \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger$$

Proof. It is clear that im $P_x^\alpha = \ker\alpha_x^\dagger$, since $\alpha_x^\dagger P_x^\alpha = 0$. Furthermore, P_x^α is clearly Hermitian, and since

$$\begin{aligned} (P_x^\alpha)^2 &= (1 - \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger)^2 \\ &= 1 - 2\alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger + \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger\alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger \\ &= 1 - 2\alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger + \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger \\ &= P_x^\alpha. \end{aligned}$$

In a similar way we find that P_x^β is an orthogonal projection. Since $\beta_x\alpha_x = 0$, we have

$$\begin{aligned} P_x^\beta P_x^\alpha &= (1 - \beta_x^\dagger(\beta_x\beta_x^\dagger)^{-1}\beta_x)(1 - \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger) \\ &= 1 - \beta_x^\dagger(\beta_x\beta_x^\dagger)^{-1}\beta_x - \alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger + \beta_x^\dagger(\beta_x\beta_x^\dagger)^{-1}\beta_x\alpha_x(\alpha_x^\dagger\alpha_x)^{-1}\alpha_x^\dagger \\ &= P_x, \end{aligned}$$

and similarly we find $P_x = P_x^\alpha P_x^\beta$, since $\alpha_x^\dagger\beta_x^\dagger = (\beta_x\alpha_x)^\dagger = 0$. So we found that P_x^α and P_x^β commute, whence $P_x^2 = P_x^\alpha P_x^\beta P_x^\alpha P_x^\beta = P_x^\alpha P_x^\alpha P_x^\beta P_x^\beta = P_x^\alpha P_x^\beta = P_x$, so P_x is an idempotent. Furthermore, P_x is clearly Hermitian, so it is an orthogonal projection. From the way we can express $P_x = P_x^\alpha P_x^\beta$ it is clear that P_x maps K_1 onto $E_x = \ker\alpha_x^\dagger \cap \ker\beta_x$. $\square$

Now we have found a bundle projection $P : \underline{K}_1 \to E$ and given the flat product connection d on $\underline{K}_1$, we can define an induced connection $\mathcal{A}$ on E by defining the covariant derivative $\mathrm{d}_{\mathcal{A}}$ by

$$\mathrm{d}_{\mathcal{A}} f = P\mathrm{d}f, \tag{4.2.4}$$

where $f : U \subset M \to E$ a section.

Lemma 4.2.3. The unitary connection $\mathcal{A}$ is compatible with the holomorphic structure $\mathcal{E}$.

Proof. We first remark that α depends holomorphically on x, so $\bar{\partial}\alpha = 0$. Furthermore, the k are holomorphic coordinates so we find

$$\bar{\partial}P^\alpha(k) = -\alpha\bar{\partial}(\alpha^\dagger\alpha)^{-1}\alpha^{-1}k,$$

thus we have $\bar{\partial}P^\alpha \in \mathrm{im}\ \alpha$, and so $P^\alpha\bar{\partial}P^\alpha = 0$.

Now, let s' be a holomorphic section of $\ker\beta$. Then we can associate the section s'' of E by the projection $s'' = P^\alpha(s') \in (\mathrm{im}\ \alpha)^\perp$, so that we find

$$P^\alpha(\bar{\partial}s'') = P^\alpha\Big(\bar{\partial}(P^\alpha(s'))\Big) = P^\alpha\Big((\bar{\partial}P^\alpha)s' + \bar{\partial}s'\Big) = P^\alpha\bar{\partial}P^\alpha s' = 0.$$

Since $P^\beta P^\alpha\bar{\partial} = \bar{\partial}_{\mathcal{A}}$ on E, we see that $\bar{\partial}_{\mathcal{A}}$ annihilates holomorphic sections on E, whence $\mathrm{d}_{\mathcal{A}}$ is compatible with the holomorphic structure. ☐

4.3 The ADHM construction

Before we introduce the ADHM construction, we give a motivation for the equations we shall introduce. Recall that the ASD-equation is equivalent to equations (3.3.16)-(3.3.18). If we restrict us to the Euclidean space $M = \mathbb{R}^4$ and assume we have a Hermitian metric, we can introduce complex coordinates $z_1 = x_2 + ix_1$, $z_2 = x_4 + ix_3$, but we could also express $\mathcal{D}_{\mathcal{A}}$ in these complex coordinates. So we introduce the operators D_1, D_2 defined by

$$\begin{aligned} D_1 &= \frac{1}{2}(\mathcal{D}_2 - i\mathcal{D}_1) \\ D_2 &= \frac{1}{2}(\mathcal{D}_4 - i\mathcal{D}_3), \end{aligned}$$

where we once again used the notation $\mathcal{D}_\mu$ instead of $(\mathcal{D}_{\mathcal{A}})_\mu$. Since the metric on $M \simeq \mathbb{C}^2$ is Hermitian, D_1, D_2 are skew-Hermitian if we demand that the connection is Hermitian. Thus we have $D_1^\dagger = \frac{1}{2}(-\mathcal{D}_2 - i\mathcal{D}_1)$ and $D_2^\dagger = \frac{1}{2}(-\mathcal{D}_4 - i\mathcal{D}_3)$. Then we find

$$\begin{aligned} 4[D_1, D_2] &= [\mathcal{D}_2 - i\mathcal{D}_1, \mathcal{D}_4 - i\mathcal{D}_3] \\ &= -[\mathcal{D}_1, \mathcal{D}_3] + [\mathcal{D}_2, \mathcal{D}_4] - i\Big([\mathcal{D}_1, \mathcal{D}_4] + [\mathcal{D}_2, \mathcal{D}_3]\Big), \end{aligned}$$

so we see that $[D_1, D_2] = 0$ is equivalent to equations (3.3.17) and (3.3.18). Furthermore, we have

$$\begin{aligned} 4[D_1, D_1^\dagger] &= [\mathcal{D}_2 - i\mathcal{D}_1, -\mathcal{D}_2 - i\mathcal{D}_1] \\ &= i[\mathcal{D}_1, \mathcal{D}_2] - i[\mathcal{D}_2, \mathcal{D}_1] \\ &= 2i[\mathcal{D}_1\, \mathcal{D}_2]. \end{aligned}$$

Similarly, we find $4[D_2, D_2^\dagger] = 2i[\mathcal{D}_3, \mathcal{D}_4]$, hence we see that (3.3.16) is equivalent to $[D_1, D_1^\dagger] + [D_2, D_2^\dagger] = 0$. Summarizing, the ASD-equations are equivalent to

$$\begin{aligned} [D_1, D_2] &= 0 \\ [D_1, D_1^\dagger] + [D_2, D_2^\dagger] &= 0. \end{aligned}$$

By the remark below Theorem 4.1.11, we see that the second equation is equivalent with the condition $(\mathcal{F}_{\mathcal{A}}, \omega) = 0$, while the first equation is equivalent with the condition that the connection is compatible with the holomorphic structure. This is not surprising, since $[D_1, D_2] = 0$ corresponds with the fact that the $(2,0)$-part of $\mathcal{F}_{\mathcal{A}}$ vanishes, thus $\mathcal{F}_{\mathcal{A}}^{2,0} = 0$, which is exactly the condition that the connection is compatible.

Now, the idea of the ADHM construction is that we take the Fourier transforms B_i, which are matrices, of D_i and solve the equations above in the Fourier-transformed space. Furthermore, we have to add source terms depending on $k \in \mathbb{Z}$ in order to get instantons with instanton number k. So we get equations of the form

$$\begin{aligned} [B_1, B_2] &= S_1 \\ [B_1, B_1^\dagger] + [B_2, B_2^\dagger] &= S_2. \end{aligned}$$

We remark that if this method works, then something remarkable has been achieved. In that case, we started with the ASD equation, which is a non-linear partial differential equation, and ended up with a set of quadratic equations whose solutions are relatively easy to find and are in 1-1 correspondence with solutions of the ASD equations.

We are now ready to give a precise formulation of the ADHM construction. First we define a certain sysem of data indexed by an integer k.

Definition 4.3.1. Let U be a four-dimensional space with a complex structure, so that we have the coordinates (z_1, z_2) on U. We shall refer to the following spaces and maps as an *ADHM system*:

1. Complex vector spaces V and W of dimension k and n respectively.
2. Complex $k \times k$ matrices B_1, B_2, a complex $k \times n$ matrix I and a complex $n \times k$ matrix J.

A system of *ADHM data* is a system $(U, V, W, B_1, B_2, I, J)$ that satisfies the following two conditions:

1. B_1, B_2, I, J satisfy the *ADHM equations*
$$[B_1, B_2] + IJ = 0 \tag{4.3.1}$$
$$[B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J = 0. \tag{4.3.2}$$
2. For all $(x, y) \neq (0,0) \in U^2$ with $x = (z_1, z_2)$, $y = (w_1, w_2)$, the map $\alpha_{(x,y)} : V \to W \oplus V \otimes U$
$$\alpha_{(x,y)} = \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \tag{4.3.3}$$
is injective, while $\beta_{(x,y)} : W \oplus V \otimes U \to V$ given by
$$\beta_{(x,y)} = \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \end{pmatrix}, \tag{4.3.4}$$
is surjective. Note that we implicitly multiplied z_1 and z_2 by the $k \times k$ identity matrix. This condition is called the *non-degeneracy condition* and implies that the $2k \times (n + 2k)$ matrix
$$R_{(x,y)} = \begin{pmatrix} \beta_{(x,y)} \\ \alpha^\dagger_{(x,y)} \end{pmatrix} : W \oplus V \otimes U \to V \times V$$
is surjective.

Lemma 4.3.2. Let $g \in \mathrm{U(k)}$, $h \in \mathrm{SU}(n)$ act on (B_1, B_2, I, J) by

$$(B_1, B_2, I, J) \mapsto (gB_1g^{-1}, gB_2g^{-1}, gIh^{-1}, hJg^{-1}).$$

Then if B_1, B_2, I, J satisfy the ADHM equation, so do the transformed matrices under these transformations.

Proof. Trivial. □

The ADHM complex

At first sight it might not be not clear how one could construct instantons from ADHM data. However, it turns out that α and β define a complex, and so we can define by the methods of section 4.2 a vector bundle and a connection which turns out to be anti-self-dual.

Proposition 4.3.3. Let $\alpha_{(x,y)}$ and $\beta_{(x,y)}$ be defined by (4.3.3) respectively (4.3.4). Then the diagram

$$V \xrightarrow{\alpha_{(x,y)}} V \otimes U \oplus W \xrightarrow{\beta_{(x,y)}} V,$$

is a complex for all $(x, y) \neq (0, 0)$ if and only if (B_1, B_2, I, J) satisfying the ADHM equations,.

Proof. The diagram is a complex if $\beta\alpha = 0$. So let us calculate $\beta\alpha$. This is

$$\begin{aligned}
\beta\alpha &= \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \end{pmatrix} \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \\
&= (w_2 I + w_1 J^\dagger)(w_2 J - w_1 I^\dagger) + (w_2 B_2 - w_1 B_1^\dagger + z_2)(-w_2 B_1 - w_1 B_2^\dagger - z_1) \\
&\quad + (w_2 B_1 + w_1 B_2^\dagger + z_1)(w_2 B_2 - w_1 B_1^\dagger + z_2) \\
&= (w_2^2 IJ + w_1 w_2 J^+ J - w_1 w_2 II\dagger - w_1^2 J^\dagger I^\dagger) + (-w_2^2 B_2 B_1 - w_1 w_2 B_2 B_2^\dagger - w_2 z_1 B_2 \\
&\quad + w_1 w_2 B_1^\dagger B_1 + w_1^2 B_1^\dagger B_2^\dagger + w_1 z_1 B_1^\dagger - z_2 w_2 B_1 - z_2 w_1 B_2^\dagger - z_1 z_2) + (w_2^2 B_1 B_2 - w_1 w_2 B_1 B_1^\dagger \\
&\quad + z_2 w_2 B_1 + w_1 w_2 B_2^\dagger B_2 - w_1^2 B_2^\dagger B_1^\dagger + z_2 w_1 B_2^\dagger + z_1 w_2 B_2 - z_1 w_1 B_1^\dagger + z_1 z_2) \\
&= w_2^2 (IJ + B_1 B_2 - B_2 B_1) + w_1^2 (-J^\dagger I^\dagger - B_2^\dagger B_1^\dagger + B_1^\dagger B_2^\dagger) \\
&\quad - w_1 w_2 (II^\dagger - J^\dagger J + B_1 B_1^\dagger - B_1^\dagger B_1 + B_2 B_2^\dagger - B_2^\dagger B_2) \\
&= w_2^2([B_1, B_2] + IJ) - w_1^2([B_1, B_2] + IJ)^\dagger - w_1 w_2([B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J).
\end{aligned}$$

So we see that $\beta\alpha = 0$ if and only if the ADHM equations (4.3.1) and (4.3.2) hold for (B_1, B_2, I, J). □

Convention. If we choose $y = 1$ (which is $(w_1, w_2) = (0, 1)$), we write x instead of $(x, 1)$. So $\beta_{(x,1)}$ is denoted by β_x, etcetera.

If U, V, W have Hermitian product, we can assign a connection $\mathcal{A}$ to the ADHM data U, V, W by constructing the fibre bundle E with fibres $E_x = \ker R_x$. The connection $\mathcal{A}$ is then given by (4.2.4). We can now state the following important theorem:

Theorem 4.3.4. We have a one-to-one correspondence between the equivalence classes of ADHM data under the action of the groups $\mathrm{U}(k)$ and $\mathrm{SU}(n)$, and gauge equivalent classes of anti-self-dual $\mathrm{SU}(n)$-connections $\mathcal{A}$ with instanton number $-k$.

A complete proof of this theorem is given in Chapter 3 of [13]. We shall only prove one direction, namely that the ADHM data define an SU(n)-instanton with instanton number $-k$.

We start by showing that the instanton number of the bundle E equals $-k$. First we introduce the adjoint of $R_{(x,y)}$, which we denote by $\Delta_{(x,y)}$, since it is a Dirac-like operator. We refer to section 3.1.1 of [13] for the mathematical details of the Dirac operator. So $\Delta_{(x,y)}$ is an $(n+2k)\times 2k$ operator given by

$$\Delta_{(x,y)} = \begin{pmatrix} \beta^\dagger_{(x,y)} & \alpha_{(x,y)} \end{pmatrix}.$$

Then we have

$$\Delta_{(x,y)} = \begin{pmatrix} \bar{w}_2 I^\dagger + \bar{w}_1 J & w_2 J - w_1 I^\dagger \\ \bar{w}_2 B_2^\dagger - \bar{w}_1 B_1 + \bar{z}_2 & -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ \bar{w}_2 B_1^\dagger + \bar{w}_1 B_2 + \bar{z}_1 & w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix}.$$

The matrix Δ_x could be seen as a Fourier transform of an ordinary Dirac operator. If we define

$$a = \begin{pmatrix} I^\dagger & J \\ B_2^\dagger & -B_1 \\ B_1^\dagger & B_2 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 & 0 \\ I_k & 0 \\ 0 & I_k \end{pmatrix} \tag{4.3.5}$$

and let $x = (z_1, z_2)$, $y = (w_1, w_2)$ be quaternions defined by (3.6.6), we have

$$\Delta_{(x,y)} = ay + bx \tag{4.3.6}$$

if we implicitly identify x with $x \otimes I_k$ and similar for y. So for instance we have

$$bx = b(x \otimes I_k) = \begin{pmatrix} 0 & 0 \\ I_k & 0 \\ 0 & I_k \end{pmatrix} \begin{pmatrix} \bar{z}_2 I_k & -z_1 I_k \\ \bar{z}_1 I_k & z_2 I_k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \bar{z}_2 & -z_1 \\ \bar{z}_1 & z_2 \end{pmatrix},$$

where $z_i, \bar{z}_i$ are implicitly multiplied with I_k in last matrix.

Lemma 4.3.5. We have $E_{(x,y)} = E_{(xq,yq)}$, where q is a non-zero quaternion defined by

$$q = \begin{pmatrix} \bar{q}_2 & -q_1 \\ \bar{q}_1 & q_2 \end{pmatrix}$$

with $q_1, q_2 \in \mathbb{C}$.

Proof. Recall that we have $E_{(x,y)} = \ker R_{(x,y)}$. Let us examine what $R_{(xq,yq)}$ looks like.

$$\begin{aligned} R_{(xq,yq)} &= \Delta^\dagger_{(xq,yq)} \\ &= (ayq + bxq)^\dagger \\ &= \bar{q}(ay + bx)^\dagger \\ &= \bar{q} R_{(x,y)}, \end{aligned}$$

from which it is clear that that $\ker R_{(xq,yq)} = \ker R_{(x,y)} = E_{(x,y)}$. ☐

We see that we can consider (x, y) homogeneous coordinates, which makes E a bundle over $\mathbb{P}_1(\mathbb{H}) = S^4$. We are now able to prove that $\mathcal{A}$ on E must have instanton number $-k$

Proposition 4.3.6. The bundle E over S^4 has topological charge $-k$.

Proof. We have $\underline{K}_1 = \underline{V \otimes U \oplus W} = \underline{\mathbb{C}}^{n+2k}$. Then we have $E = \ker R$, while the rows of $R_{(x,y)}$ span by the non-degeneracy condition a subspace of $\mathbb{C}^{n+2k}$ of complex dimension $2k$, which is complementary to $E_{(x,y)}$. Notice that this subspace only depends on the ratio xy^{-1}, that is on the point of S^4. Hence we see that the complement of E in $\underline{\mathbb{C}}^{n+2k}$ is given by k quaternionic line bundles. Since the one-dimensional quaternionic subspace $L_{x,y} = \{(xq, yq) : q \in \mathbb{H}\} \subset \mathbb{H}^2$ can be associated to $v_{x,y} = (x, y) \in \mathbb{H}^2$, and $L_{x,y}$ corresponds to a point in $\mathbb{P}_1(\mathbb{H})$, we see that each line bundle can be identified with Σ, which is the bundle over $\mathbb{P}_1(\mathbb{H})$ defined by

$$\Sigma = \{(L, v) \in \mathbb{P}_1(\mathbb{H}) \times \mathbb{H}^2 : v \in L\}$$
$$\pi_\Sigma : \Sigma \to \mathbb{P}_1(\mathbb{H}), \quad \big(L, v\big) \mapsto L.$$

Note that $\pi^{-1}(L) = \{(L', v) : \pi_\Sigma((L', v)) = L\} = \big\{(L', v) : v \in L = L'\big\} \simeq \{v \in L\} = \mathbb{H}$, so Σ is indeed a line bundle, which is called the quaternionic *tautological line bundle*. Hence $\underline{\mathbb{C}}^{n+2k} = E \oplus \Sigma^{\oplus k}$ and since the left-hand side is trivial, we have

$$0 = \mathrm{c}_2(\underline{\mathbb{C}}^{n+2k}) = \mathrm{c}_2(E) + k\mathrm{c}_2(\Sigma) \tag{4.3.7}$$

by Corollary 3.2.3. So we can calculate $\mathrm{c}_2(E)$ if we know what $\mathrm{c}_2(\Sigma)$ is. Now, just like the first Chern number of the complex tautological bundle is 1 (see for instance page 309 of [23]), the second Chern number of the quaternionic tautological bundle is 1 (that in the quaternionic case it is the second Chern class which is 1, and not the first Chern class, has to do with the fact that the dimension of $\mathbb{H}$ is twice the dimension of $\mathbb{C}$). $\square$

If we choose $y = 1$, which is $(y_1, y_2) = (0, 1)$, we obtain a subbundle over $\mathbb{R}^4$. As mentioned before, we shall use the notation x instead of $(x, 1)$. Then we have

$$\Delta_x^\dagger = a^\dagger + \bar{x}b^\dagger = \begin{pmatrix} I & B_2 + z_2 & B_1 + z_1 \\ J^\dagger & -B_1^\dagger - \bar{z}_1 & B_2^\dagger + \bar{z}_2 \end{pmatrix}, \tag{4.3.8}$$

with

$$\bar{x}b^\dagger = (\bar{x} \otimes I_k)b^\dagger$$

Note that we have

$$\Delta_x^\dagger = \begin{pmatrix} \beta_x \\ \alpha_x^\dagger \end{pmatrix},$$

so Δ_x is of maximum rank by the non-degeneracy condition for α_x and β_x. Furthermore, we remark that be can write

$$\Delta_x = a + bx = a + b\tau_\mu x^\mu. \tag{4.3.9}$$

We shall see that we can obtain $\mathrm{SU}(n)$ instantons with instanton number $-k$ by looking at solutions Ψ of the equations

$$\Delta_x^\dagger \Psi = 0 \tag{4.3.10}$$

The reason why we use the notation $\Delta_x^\dagger \Psi = 0$ instead of $R_x \Psi = 0$ is to emphasize that the Ψ can be seen as solutions of the Dirac equation. We can give an orthonormal basis of the nullspace of $\Delta_x^\dagger$ and construct a matrix M whose columns are exactly the basis vectors of the null space. Hence we have

$$\Delta_x^\dagger M = 0 \tag{4.3.11}$$
$$M^\dagger M = 1. \tag{4.3.12}$$

The second equation is the orthonormalization condition. A simple calculation gives

$$\Delta_x^\dagger \Delta_x = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

with

$$\begin{aligned}
&p = \beta_x \beta_x^\dagger = II^\dagger + B_1 B_1^\dagger + B_2 B_2^\dagger + B_1 \bar{z}_1 + B_2 \bar{z}_2 + B_1^\dagger z_1 + B_2^\dagger z_2 + |z_1|^2 + |z_2|^2 \\
&q = IJ + [B_1, B_2] \\
&r = J^\dagger I^\dagger + [B_2^\dagger, B_1^\dagger] \\
&s = \alpha_x^\dagger \alpha_x = J^\dagger J + B_1^\dagger B_1 + B_2^\dagger B_2 + B_1 \bar{z}_1 + B_2 \bar{z}_2 + B_1^\dagger z_1 + B_2^\dagger z_2 + |z_1|^2 + |z_2|^2.
\end{aligned}$$

The first ADHM equation (4.3.1) implies that $q = 0$. Since $r = q^\dagger$, we also find that $r = 0$. Clearly p and s are Hermitian $k \times k$ matrices. By adding the second ADHM equation (4.3.2) to s, we see that $s = p$. It is conventional to write $p = s = f^{-1}$, so that we have

$$\Delta_x^\dagger \Delta_x = \begin{pmatrix} f^{-1} & 0 \\ 0 & f^{-1} \end{pmatrix}.$$

Note that since Δ_x is of maximum rank, f must be of maximum rank. Furthermore, if we identify f with

$$I_2 \otimes f = \begin{pmatrix} f & 0 \\ 0 & f \end{pmatrix},$$

we have $\Delta_x^\dagger \Delta_x = f^{-1}$. If we use a similar identification, we can define the projection operator Q by

$$Q_x = \Delta_x f \Delta_x^\dagger. \tag{4.3.13}$$

This is indeed an orthogonal projection, since it is clearly Hermitian and

$$Q_x^2 = \Delta_x f \Delta_x^\dagger \Delta_x f \Delta_x^\dagger = \Delta_x f f^{-1} f \Delta_x^\dagger = Q_x.$$

Lemma 4.3.7. We have $P + Q = 1$.

Proof. By direct calculation. Since $f^{-1} = \alpha_x^\dagger \alpha_x = \beta_x \beta_x^\dagger$, we have

$$\begin{aligned}
Q_x &= \Delta_x f \Delta^\dagger \\
&= \begin{pmatrix} \beta_x^\dagger & \alpha_x \end{pmatrix} \begin{pmatrix} f & 0 \\ 0 & f \end{pmatrix} \begin{pmatrix} \beta_x \\ \alpha_x^\dagger \end{pmatrix} \\
&= \beta_x^\dagger f \beta_x + \alpha_x f \alpha_x^\dagger \\
&= \beta_x^\dagger (\beta_x \beta_x^\dagger)^{-1} \beta_x + \alpha_x (\alpha_x^\dagger \alpha_x)^{-1} \alpha_x^\dagger \\
&= 1 - P_x.
\end{aligned}$$

□

Now we are able to describe P_x in terms of M.

Lemma 4.3.8. We have $P_x = MM^\dagger$.

Proof. First we remark that $P'_x = MM^\dagger$ is indeed an orthogonal projection, since it is clearly Hermitian and is an idempotent by the orthonormalization condition (4.3.12).

We shall prove that $P = P'$. Since $\Delta_x^\dagger M = 0$ (thus $M^\dagger \Delta_x = 0$), we find that

$$P'_x Q_x = Q_x P'_x = 0,$$

so P'_x and Q_x are orthonormal. This means that the operator $1 - Q_x - P'_x$ is a projection operator. Note that projection operators only have 0 and 1 as eigenvalue, so projection operators are *positive*. Whence the operator is clearly Hermitian, since P'_x and Q_x are, we find that

$$1 - Q_x - P'_x = VV^\dagger$$

for some operator V (for an advanced proof that this V exists, see Theorem VIII.3.5 of [11]). Now, we have $Q_x VV^\dagger = 0$, or equivalently $\Delta_x f \Delta_x^\dagger VV^\dagger = 0$, and since f and Δ_x are of maximum rank, we find $\Delta_x^\dagger VV^\dagger = 0$. With other words, V must consist of other vectors annihilated by $\Delta_x^\dagger$, linear independent of the vectors which are the building stones of M. Since M consists by definition of the complete basis of the nullspace of $\Delta_x^\dagger$, we must have $V = 0$. So we have $1 - Q_x - P'_x = 0$ and since $P_x + Q_x = 1$, it follows that $P_x = P'_x$. □

We have arrived at the point where we can express $\mathcal{A}$ in terms of M.

Proposition 4.3.9. We have

$$\mathcal{A} = M^\dagger \mathrm{d}M. \tag{4.3.14}$$

Notice that M is not an element of $\mathrm{SU}(n)$, since it is not an $n \times n$ matrix, unless $k = n$, so $\mathcal{A}$ is not pure gauge.

Proof. Let s be a section. Then we find

$$M\mathrm{d}s + M\mathcal{A}s = \mathrm{d}_{\mathcal{A}}(Ms) = P\mathrm{d}(Ms) = MM^\dagger \mathrm{d}(Ms) = M\Big(\mathrm{d}s + (M^\dagger \mathrm{d}M)s\Big),$$

whence $\mathcal{A} = M^\dagger \mathrm{d}M$. □

Now we are able to prove that the field-strength of a connection defined by (4.3.14) is anti-self-dual. We introduce the following notation:

Notation. $A_{[\mu}B_{\nu]}$ is defined to be the object $A_\mu B_\nu - A_\nu B_\mu$.

Theorem 4.3.10. The connection defined by (4.3.14) is $\mathfrak{su}(n)$-valued, is anti-self-dual and has instanton number $\mathrm{n} = -k$.

Proof. Using the notation we introduced above, we find when substituting (4.3.14) into (2.2.12)

$$\begin{aligned}
\mathcal{F}_{\mu\nu} &= \partial_{[\mu}\mathcal{A}_{\nu]} + \mathcal{A}_{[\mu}\mathcal{A}_{\nu]} \\
&= \partial_{[\mu}(M^\dagger \partial_{\nu]} M) + (M^\dagger \partial_{[\mu} M)(M^\dagger \partial_{\nu]} M) \\
&= (\partial_{[\mu} M^\dagger)(\partial_{\nu]} M) - (\partial_{[\mu} M^\dagger) M (M^\dagger \partial_{\nu]} M) && \text{[By } (A.2.5)] \\
&= (\partial_{[\mu} M^\dagger)(1 - MM^\dagger)(\partial_{\nu]} M) \\
&= (\partial_{[\mu} M^\dagger) Q (\partial_{\nu]} M) && \text{[By Lemma 4.3.8]} \\
&= (\partial_{[\mu} M^\dagger)\Delta_x f \Delta_x^\dagger (\partial_{\nu]} M)
\end{aligned}$$

Now, by (4.3.11) we have $0 = \partial_\nu(\Delta_x^\dagger M) = (\partial_\nu\Delta_x^\dagger)M + \Delta_x^\dagger\partial_\nu M$, so we have $\Delta_x^\dagger\partial_\nu M = -(\partial_\nu\Delta_x^\dagger)M$. Similarly, we have $(\partial_\mu M^\dagger)\Delta_x = -M^\dagger\partial_\nu\Delta_x$, whence

$$\begin{aligned}\mathcal{F}_{\mu\nu} &= M^\dagger(\partial_{[\mu}\Delta_x)f(\partial_{\nu]}\Delta_x^\dagger)M\\ &= M^\dagger b\tau_{[\mu}f\tau_{\nu]}^\dagger b^\dagger M.\end{aligned}$$

We have identified x with $x\otimes I_k$, whence τ_μ must be identified with $\tau_\mu\otimes I_k$. Then by $(A\otimes B)(C\otimes D) = AC\otimes BD$ we find

$$\tau_{[\mu}f\tau_{\nu]}^\dagger = (\tau_{[\mu}\otimes I_k)(I_2\otimes f)(\tau_{\nu]}^\dagger\otimes I_k) = \tau_{[\mu}\tau_{\nu]}^\dagger\otimes f = -2\sigma_{\mu\nu}\otimes f,$$

where we used that $\tau_\mu = -i\sigma_\mu$ and (3.6.1). So by (3.6.4), we find that $\mathcal{F}_{\mu\nu}$ is anti-self-dual.

Differentiating (4.3.12) we find

$$\mathcal{A}^\dagger = (\mathrm{d}M^\dagger)M = -M^\dagger\mathrm{d}M = -\mathcal{A},$$

so $\mathcal{A}$ is anti-Hermitian, hence it is $\mathfrak{u}(n)$-valued. We can calculate its trace and since $\mathcal{A}$ is anti-Hermitian, we find

$$\mathrm{Im}\ \mathrm{tr}(\mathcal{A}) = \frac{1}{2}\Big(\mathrm{tr}(\mathcal{A}) - \overline{\mathrm{tr}(\mathcal{A})}\Big) = \frac{1}{2}\Big(\mathrm{tr}(\mathcal{A}) - \mathrm{tr}(\mathcal{A}^\dagger)\Big) = \mathrm{tr}(\mathcal{A}),$$

so $\mathrm{tr}(\mathcal{A})$ is pure imaginary. Hence we find that $\mathcal{A} = \mathcal{A}^1 + \mathcal{A}^2$ with $\mathcal{A}_2 = \frac{1}{n}\mathrm{tr}(\mathcal{A})I_n$ $\mathfrak{u}(1)$-valued, where we identified $\mathfrak{u}(1)$ as the subalgebra $\{irI_n : r\in\mathbb{R}\}$ of $\mathfrak{u}(n)$, and $\mathcal{A}^1$ $\mathfrak{su}(n)$-valued, since we have indeed $\mathrm{tr}(\mathcal{A}^1) = \mathrm{tr}(\mathcal{A}) - \frac{1}{n}\mathrm{tr}(\mathcal{A})\mathrm{tr}(I_n) = 0$. By (2.2.11) the field-strength of $\mathcal{A}$ is given by

$$\begin{aligned}\mathcal{F} &= \mathrm{d}(\mathcal{A}^1+\mathcal{A}^2) + (\mathcal{A}^1+\mathcal{A}^2)\wedge(\mathcal{A}^1+\mathcal{A}^2)\\ &= \mathrm{d}\mathcal{A}^1 + \mathrm{d}\mathcal{A}^2 + \mathcal{A}^1\wedge\mathcal{A}^1 + \mathcal{A}^2\wedge\mathcal{A}^2 + \mathcal{A}^1\wedge\mathcal{A}^2 + \mathcal{A}^2\wedge\mathcal{A}^1\end{aligned}$$

The latter two terms are equal to $(\mathcal{A}^1_\mu\mathcal{A}^2_\nu + \mathcal{A}^2_\mu\mathcal{A}^1_\nu)\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$. Notice that $\mathcal{A}^1_\mu\mathcal{A}^2_\nu + \mathcal{A}^2_\mu\mathcal{A}^1_\nu$ is symmetric in μ and ν since $\mathcal{A}^1_\mu$ commutes with $\mathcal{A}^2_\nu$ for the latter is a diagonal matrix. So $\mathcal{A}^1\wedge\mathcal{A}^1 + \mathcal{A}_2\wedge\mathcal{A}^2 + \mathcal{A}^1\wedge\mathcal{A}^2 = 0$ and we find that $\mathcal{F} = \mathcal{F}^1 + \mathcal{F}^2$, where $\mathcal{F}^i$ comes from $\mathcal{A}^i$ by (2.2.11). Since the Hodge star is linear, we find that the ADS-equation $*\mathcal{F} = -\mathcal{F}$ is equivalent with $*\mathcal{F}^1 + *\mathcal{F}^2 = -\mathcal{F}^1 - \mathcal{F}^2$. Now, $\mathcal{F}^1$ is traceless, while $\mathcal{F}^2$ only has entries on the diagonal, so we obtain two decoupled ASD equations $*\mathcal{F}^1 = -\mathcal{F}^1$ and $*\mathcal{F}^2 = -\mathcal{F}^2$. So $\mathcal{F}^2$ is the field-strength of an ASD U(1)-gauge potential $\mathcal{A}^2$, but by Proposition 3.3.9, we see that $\mathcal{F}^2$ must be zero. By Corollary 2.6.6, $\mathcal{A}^2$ is pure gauge, which disappears after an appropriate U(1)-gauge transformation. Thus we can assume that $\mathcal{A}$ is $\mathfrak{su}(n)$-valued.

Recall that by Proposition A.4.7 the anti-self-duality equations are conformal invariant. So the instanton $\mathcal{A}$ can seen as an instanton defined on the bundle E with fibre $E_{(x,y)}$ over $S^4-\{0\}$. Then by Proposition 4.3.6 it follows that the instanton number of $\mathcal{A}$ is $-k$.

□

The (-1)-instanton revisited

Finally we show that the solution of the ADHM equation for $n=2$ and $k=1$ equals (3.6.12).

Since the B_i matrices are 1×1, all commutators are zero, so (4.3.1) and (4.3.2) become

$$\begin{aligned}IJ &= 0;\\ II^\dagger - J^\dagger J &= 0,\end{aligned}$$

and we have a free choice for B_1 and B_2. We shall choose $B_1 = -b_1 \in \mathbb{C}$, $B_2 = -b_2 \in \mathbb{C}$. Now, if we represent I and $J^\dagger$ by the complex 2-vectors $i = (i_1, i_2)$ and $j = (j_1, j_2)$, the ADHM equations are equivalent to

$$\begin{aligned} i \cdot j &= 0; \\ |i| &= |j|. \end{aligned}$$

We shall choose $i = (\rho, 0)$ and $j = (0, \rho)$ with $\rho \in \mathbb{R}$, whence $\Delta_x^\dagger$ becomes

$$\Delta_x^\dagger = \begin{pmatrix} \rho & 0 & z_2 - b_2 & z_1 - b_1 \\ 0 & \rho & -\bar{z}_1 + \bar{b}_1 & \bar{z}_2 - \bar{b}_2 \end{pmatrix} = \begin{pmatrix} \rho & \bar{x} - \bar{b} \end{pmatrix},$$

with x defined as in (3.6.6). We define $b \in \mathbb{H}$ in a similar way. Then we have two vectors spanning the null space of $\Delta_x^\dagger$:

$$v_1 = \lambda \begin{pmatrix} -z_2 + b_2 \\ \bar{z}_1 - \bar{b}_1 \\ \rho \\ 0 \end{pmatrix}, \qquad v_2 = \lambda \begin{pmatrix} -z_1 + b_1 \\ -\bar{z}_2 + \bar{b}_2 \\ 0 \\ \rho \end{pmatrix},$$

with λ a scalar. Hence we find

$$M = \lambda \begin{pmatrix} -z_2 + b_2 & -z_1 + b_1 \\ \bar{z}_1 - \bar{b}_1 & -\bar{z}_2 + \bar{b}_2 \\ \rho & 0 \\ 0 & \rho \end{pmatrix} = \lambda \begin{pmatrix} -\bar{x} + \bar{b} \\ \rho \end{pmatrix}.$$

Then by $x^\dagger = (\tau_\mu x^\mu)^\dagger = \tau_\mu^\dagger x^\mu = \bar{x}$, we find

$$M^\dagger = \lambda \begin{pmatrix} -x + b & \rho \end{pmatrix},$$

so that $M^\dagger M = \lambda^2(|x - b|^2 + \rho^2)$. By (4.3.12) we see that this implies that $\lambda = \frac{1}{\sqrt{\rho^2 + |x-b|^2}}$, hence we have

$$M = \frac{1}{\sqrt{\rho^2 + |x - b|^2}} \begin{pmatrix} -\bar{x} + \bar{b} \\ \rho \end{pmatrix}.$$

Simply differentiating gives

$$\partial_\mu M = \begin{pmatrix} \frac{-\tau_\mu^\dagger}{\sqrt{\rho^2 + |x-b|^2}} + \frac{(\bar{x} - \bar{b})(x_\mu - b_\mu)}{(\rho^2 + |x-b|^2)^{3/2}} \\ \frac{-\rho(x_\mu - b_\mu)}{(\rho^2 + |x-b|^2)^{3/2}} \end{pmatrix},$$

hence

$$\begin{aligned} \mathcal{A}_\mu &= M^\dagger \partial_\mu M \\ &= \frac{(x - b)\tau_\mu^\dagger}{\rho^2 + |x - b|^2} - \frac{|x - b|^2 (x_\mu - b_\mu)}{(\rho^2 + |x - b|^2)^2} - \frac{\rho^2 (x_\mu - b_\mu)}{(\rho^2 + |x - b|^2)^2} \\ &= \frac{(x - b)\tau_\mu^\dagger - (x_\mu - b_\mu)}{\rho^2 + |x - b|^2}. \end{aligned}$$

Now, if $y = x - b$ we have

$$\begin{aligned}
y\tau_1^\dagger - y_1 &= (iy^1 + jy^2 + ky^3 + y^4)(-i) - y_1 = \mathrm{im}(-yi)\\
y\tau_2^\dagger - y_2 &= (iy^1 + jy^2 + ky^3 + y^4)(-j) - y_2 = \mathrm{im}(-yj)\\
y\tau_2^\dagger - y_2 &= (iy^1 + jy^2 + ky^3 + y^4)(-k) - y_3 = \mathrm{im}(-yk)\\
y\tau_4^\dagger - y_4 &= (iy^1 - jy^2 + ky^3 + y^4) - y_4 = \mathrm{im}(y),
\end{aligned}$$

hence we see

$$\begin{aligned}
\mathcal{A} &= \frac{y\tau_\mu^\dagger - y_\mu}{\rho^2 + |y|^2}\mathrm{d}x^\mu\\
&= \mathrm{Im}\left(\frac{-yi\mathrm{d}x^1 - yj\mathrm{d}x^2 - yk\mathrm{d}x^3 + y\mathrm{d}x^4}{\rho^2 + |y|^2}\right)\\
&= \mathrm{Im}\left(\frac{y\mathrm{d}\bar{x}}{\rho^2 + |y|^2}\right)\\
&= \mathrm{Im}\left(\frac{(x-b)\mathrm{d}\bar{x}}{\rho^2 + |x-b|^2}\right),
\end{aligned}$$

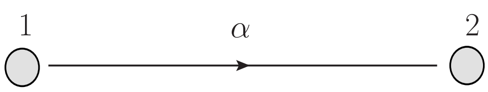
which is exactly the instanton defined in (3.6.12).

4.4 Quivers

In this section we shall introduce the concept of a *quiver*, which plays an important role in recent research, since it has connections with many concepts in mathematics and physics as Dynkin diagrams in the theory of Lie algebras, the study of Feynman diagrams and with D-branes in string theory. The link of quivers with instantons is given by the fact that one of the first interesting quivers was inspired by the ADHM construction.

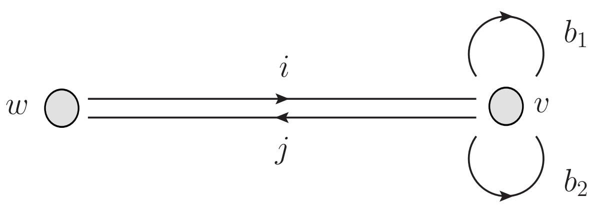
Definition 4.4.1. A *quiver* Q is a directed graph, that is, a quadruple $\{Q_V, Q_A, s, t\}$, where Q_V is a collection of *vertices*, Q_A a collection of *arrows*, $s, t : Q_A \to Q_V$ are functions which assign to each arrow a *starting vertex* respectively a *terminating vertex*. If both V and A are finite, the quiver is called a *finite quiver*.

Example 4.4.2. Next picture shows the quiver Q with $Q_V = \{1, 2\}$, $Q_A = \{\alpha\}$, $s(\alpha) = 1$, $t(\alpha) = 2$.



Next example shows that quivers may have multiple arrows between the same two vertices and arrows from one to the same vertex. The quiver is called the ADHM quiver, which is derived from the ADHM construction.

Example 4.4.3. The *ADHM quiver* Q_{ADHM} is described by $Q_V = \{v, w\}$, $Q_A = \{b_1, b_2, i, j\}$, $s(b_1) = s(b_2) = s(j) = t(b_1) = t(b_2) = t(i) = v$, $t(j) = s(i) = w$. The graph is given by



Definition 4.4.4. A *representation of a quiver* Q over a field $\mathbb{F}$ is an assignment of a vector space $V(x)$ over $\mathbb{F}$ to each vertex $x \in Q_V$ and an $\mathbb{F}$-linear map $V(\alpha) : V(x) \to V(y)$ to each arrow $\alpha \in Q_A$ where $x = s(\alpha)$ and $y = t(\alpha)$. We call the representation *finite* if $\dim_{\mathbb{F}} V = \sum_{x \in Q_V} \dim_{\mathbb{F}} V(x)$ is finite.

Examples 4.4.5.

1. For any quiver the *zero representation* can be defined by assigning the zero space to each vertex and the zero map to each arrow.
2. Let Q be the quiver from Example 4.4.2. Then a representation of Q is given by $V(1) = \mathbb{F}^n$, $V(2) = \mathbb{F}^m$ and $V(\alpha) = M$, with M any $m \times n$-matrix with entries in $\mathbb{F}$.
3. For Q_{ADHM} a representation V_{ADHM} is given by an ADHM system. Thus $V(v) = V$, $V(w) = W$ with V and W complex vector spaces of dimension k and n respectively, $V(b_1) = B_1$, $V(b_2) = B_2$, $V(i) = I$, $V(j) = J$, where B_1, B_2 are complex $k \times k$ matrices, I is a complex $k \times n$ matrix and J is a complex $n \times k$ matrix.

Definition 4.4.6. Let V and W be representations over $\mathbb{F}$ of a quiver Q. Then a *morphism* from V to W is a family $\{\varphi(x)\}_{x \in Q_V}$ of $\mathbb{F}$-linear maps $\varphi(x) : V(x) \to W(x)$ such that for all arrows α we have $W(\alpha)\varphi(x) = \varphi(y)V(\alpha)$, where $x = s(\alpha)$ and $y = t(\alpha)$. In other words, the diagram

$$\begin{tikzcd} V(x) \arrow[r, "\phi(x)"] \arrow[d, "V(\alpha)"'] & W(x) \arrow[d, "W(\alpha)"] \\ V(y) \arrow[r, "\phi(y)"] & W(y) \end{tikzcd}$$

commutes. If $\varphi(x)$ is bijective for all $x \in Q_V$, we call $\{\varphi(x)\}_{x \in Q_V}$ an *isomorphism* of representations.

Given two representations of a quiver Q, we can define their direct sum.

Definition 4.4.7. Let V and W be representations over $\mathbb{F}$ of a quiver Q. Then the *direct sum* $V \oplus W$ over V and W is defined by setting $(V \oplus W)(x) = V(x) \oplus W(x)$ for all $n \in Q_V$ and $(V \oplus W)(\alpha) = V(\alpha) \oplus W(\alpha)$ for all $\alpha \in Q_A$.

By direct sums one could decompose representations of quivers into simpler representations. This leads to the following definition:

Definition 4.4.8. Let V be a representation over $\mathbb{F}$ of a quiver Q. Then V is called an *irreducible* representation of Q if there are no representation V_1 and V_2 over $\mathbb{F}$ of Q such that V is isomorphic to $V_1 \oplus V_2$.

Every finite dimensional representation V of a quiver Q can be decomposed into a finite sum of irreducible representations $V_1, \ldots, V_n$, which can be shown by induction. Next theorem tells that this decomposition is unique. A proof can be found in [3] as Theorem 1.8.

Theorem 4.4.9 (Krull-Remak-Schmidt)**.** Let $V_1, \ldots, V_n, W_1, \ldots, W_m$ be irreducible representations of a quiver Q such that $V_1 \oplus \ldots \oplus V_n \simeq W_1 \oplus \ldots \oplus W_m$. Then $n = m$ and there is a permutation π of $(1, \ldots, n)$ such that $V_i \simeq W_{\pi(i)}$ for all $1 \leq i \leq n$.

Another way of studying a quiver Q is by its path algebra.

Definition 4.4.10. Let Q be a quiver. A *path* p is a sequence $\alpha_1\alpha_2\ldots\alpha_n$ of arrows $\alpha_i \in Q_A$ such that $s(\alpha_i) = t(\alpha_{i+1})$ for $i = 1, \ldots, n-1$. We extent s and t to the space of all paths by defining $t(p) = t(\alpha_1)$ and $s(p) = s(\alpha_n)$. If $q = \beta_1\beta_2\ldots\beta_m$ is another path such that $t(q) = s(p)$, we define the concatenation pq of p and q by

$$pq = \alpha_1\alpha_2\ldots\alpha_n\beta_1\beta_2\ldots\beta_m.$$

If $x \in Q_V$, we define the trivial path e_x by the conditions $s(e_x) = t(e_x) = x$ and $e_{t(p)}p = pe_{s(x)} = p$ for all paths p in Q.

Definition 4.4.11. Let Q a quiver and $\mathbb{F}$ a field. Then we define the path algebra $\mathbb{F}Q$ as the vector space spanned by all paths in Q with multiplication defined by

$$p \cdot q = \begin{cases} pq, & s(p) = t(q); \\ 0, & s(p) \neq t(q). \end{cases}$$

Equivalently we could say that the path algebra $\mathbb{F}Q$ is the algebra generated by all $\alpha \in Q_A$ and e_x with $x \in Q_V$ such that

$$\begin{array}{ll} \alpha\beta = 0 & [s(\alpha) \neq t(\beta), \alpha, \beta \in Q_A] \\ \alpha e_x = 0 & [s(\alpha) \neq x, \alpha \in Q_A, x \in Q_V] \\ \alpha e_{s(\alpha)}x = \alpha & [\alpha \in Q_A] \\ e_x\alpha = 0 & [t(\alpha) \neq x, \alpha \in Q_A, x \in Q_V] \\ e_{t(\alpha)}\alpha = \alpha & [\alpha \in Q_A] \\ e_xe_y = 0 & [x, y \in Q_V, x \neq y] \\ e_x^2 = e_x & [x \in Q_V]. \end{array}$$

Examples 4.4.12.

1. Let Q be the quiver from Example 4.4.2. Then $\mathbb{F}Q$ is the algebra generated by $\{e_1, e_2, \alpha\}$.

2. Let Q be the quiver with one vertex 1 and one arrow α with $s(\alpha) = t(\alpha) = 1$. Then $\mathbb{F}Q$ is the algebra spanned by the paths $\{e_1, \alpha, \alpha^2, \alpha^3, \ldots\}$, which is isomoprhic to the polynomial ring $K[\alpha]$, where e_1 is the 1-element in the ring.

Just like a representation of a Lie algebra is equivalent with the notion of a $\mathfrak{g}$-module, it turns out that a representation of a quiver Q is equivalent with an $\mathbb{F}Q$-module.

Proposition 4.4.13. The category $\mathrm{Rep}_{\mathbb{F}}(Q)$ of finite-dimensional representations over $\mathbb{F}$ of a finite quiver Q and the category $\mathbb{F}Q-$ mod of finite-dimensional $\mathbb{F}Q$-modules are equivalent.

Proof. Let V be a finite-dimensional representation of a quiver Q. Then we define the left $\mathbb{F}Q$-module $\bar{V} = \bigoplus_{x\in Q_V} V(x)$ with multiplication $\mathbb{F}Q \times \bar{V} \to \bar{V}$ defined as follows. Let $v = (v_x)_{x\in Q_V}$ be a vector in $\bar{V}$. Then if p is a path $\alpha_1\alpha_2\ldots\alpha_n$ with $s(p) = y$ and $t(p) = z$ we define $pv = (pv_x)_{x\in Q_V}$ such that $(pv)_z = V(\alpha_1)V(\alpha_2)\ldots V(\alpha_n)(v_y)$, while all other components of pv are defined to be zero.

Conversely given a finite-dimensional left $\mathbb{F}Q$-module M, we define $V(x) = e_xM = \{e_xm : m \in M\}$. Hence we have $M = \bigoplus_{x\in Q_V} V(i)$ and we can define a representation by setting $V(\alpha) : V(x) \to V(y)$ by $e_xm \mapsto \alpha m$ for all arrows $\alpha : x \to y$. $\square$

As we have seen in Example 4.4.12.2, it is possible to obtain an infinite-dimensional path algebra. One could resolve this by imposing extra conditions.

Definition 4.4.14. Let Q be a quiver and $\mathbb{F}$ a field. Then an *admissible relation* is an element r of $\mathbb{F}Q$ of the form $r = \sum_{i=1}^{n} c_i p_i$ where $c_i \in \mathbb{F}$ for all i and p_i are paths such that $s(p_1) = s(p_2) = \ldots = s(p_n)$ and $t(p_1) = t(p_2) = \ldots = t(p_n)$. An *admissible ideal* is a two-sided ideal generated by admissible relations. $\mathbb{F}Q/I$ with I an admissible ideal is called an *algebra of a quiver with relations.*

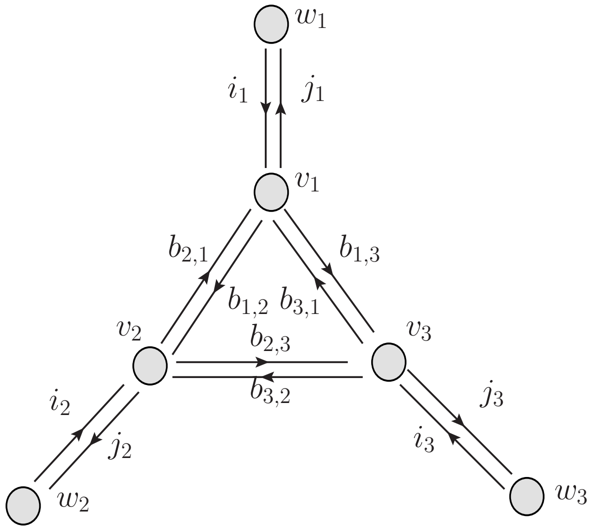
Next proposition might not come as a surprise.

Proposition 4.4.15. Let $\mathbb{F}Q/I$ be an algebra of a quiver with relations. Then modules of $\mathbb{F}Q/I$ correspond to representations of Q for which the equations $r = 0$ for all $r \in I$ hold.

So we find that representations of the ADHM quiver Q_{ADHM} over $\mathbb{C}$ with the relation $r = b_1 b_2 - b_2 b_1 + ij$ are exactly the representation for which the first ADHM equation (4.3.1) holds. There are two ways to make sure the second ADHM equation also holds. The most elegant solution is restricting ourself to holomorphic representations. Then a solution of the first ADHM equation is also a solution of the second ADHM equation if and only if the corresponding holomorphic vector bundle is stable. If we now the moduli space $\mathscr{A}^-/\tilde{\mathcal{G}}^{\mathbb{C}}$ instead of $\mathscr{M} = \mathscr{A}^-/\tilde{\mathcal{G}}$, where $\tilde{\mathcal{G}}^{\mathbb{C}}$ is the complexification of $\tilde{\mathcal{G}}$, this means that the second ADHM equation holds if and only if the connection is irreducible. If we recall Theorem 3.7.10, we see that the case of SU(2) all anti-self-dual connections are irreducible, so we see for SU(2) we can drop the second ADHM equation as a condition if we restrict ourself to holomorphic vector bundles and use the complexification of the gauge group instead of the ordinary gauge group. For more details we refer to chapter 6 of [13].

Another solution is extending the quiver in the following way. We consider the opposite quiver $Q_{\mathrm{ADHM}}^{\mathrm{op}}$, which is exactly the same quiver as Q_{ADHM}, but only with reversed arrows. If we denote the reversed arrow of α by α^*, we can define a new quiver Q with set of arrows $Q_A = (Q_{\mathrm{ADHM}})_A \cup (Q_{\mathrm{ADHM}}^{\mathrm{op}})_A$. We demand that the representations respect Hermitian structures: $V(\alpha^*) = V(\alpha)^*$, and introduce another relation: $r' = b_1 b_1^* - b_1^* b_1 + b_2 b_2^* - b_2^* b_2 + ii^* - j^* j$. The representations of this quiver are exactly the solutions of both ADHM equations. Hence we see that it is possible to study the ADHM contruction by the properties of the ADHM quiver and its representations.

It turns out that quivers offer a way to generalize the ADHM construction on other spaces than four-dimensional Euclidean space, such as ALE spaces. These are spaces obtained by considering $\mathbb{C}^2/\Gamma$, where $\Gamma \subset \mathrm{SU}(2)$ is a discrete subgroup. Then an ALE space (which is an abbreviation of Asymptotically Locally Euclidean space) is a hyperkähler manifold which is diffeomorphic to a so called minimal resolution of $\mathbb{C}^2/\Gamma$. It turns out that the discrete subgroups Γ of SU(2) can be classified by the A-, D- and E-series of Dynkin diagrams. Since Dynkin diagrams can be seen as quivers where the direction of the arrows has been forgotten, we see that ALE spaces can be classified by quivers. An interesting side-issue is that ALE spaces are examples of graviational instantons, which are the equivalents of Yang-Mills instantons in general relativity. Also on ALE spaces one could define gauge theories and try to examine what its Yang-Mills instantons are. This can be done by adjusting the ADHM quiver, which means that depending on the Dynkin diagram corresponding to the ALE space we have to add extra vertices and arrows. For instance the quiver for the ALE space corresponding to the Dynkin diagram A_2 (or equivalently $\Gamma = \mathbb{Z}_3$) is given by



Hence we find that quivers offer the opportunity to generalize the ADHM construction not only to S^4 and $\mathbb{R}^4$, but to a wide range of four-manifolds. ALE spaces form an important subject in string theory, especially in the theory of D-branes, which illustrates that quivers play an important role in modern research, not only in helping to define new spaces, but also in generalizing the ADHM construction on those spaces.

5 Conclusion

In this thesis we have studied space of instanton solutions with topological charge n in two ways. In the first one we used the moduli space over S^4, from which we could derive the number of solutions by calculating its dimension. The moduli space itself is a very fruitful subject for study, and the moduli space over S^4 is only one example, since one could easily define instantons and moduli spaces on other spaces than S^4. In this thesis we only studied the dimension of the moduli space, but latter has many more interesting properties. For instance, one could introduce metrics on the moduli space and study its diameter and volume [15]. Furthermore, one could ask how this metric should be related to the metric of the underlying space. The moduli spaces can be used to define new geometrical invariants. For example Donaldson used cobordism theory to show the existence of exotic differential structures on $\mathbb{R}^4$ by studying the moduli spaces of the boundaries of cobordisms [13].

The second method of studying instantons we used was the ADHM construction. This construction was introduced by Atiyah, Drinfeld, Hitchin and Manin and translated the problem of finding instanton solutions into the problem of solving a linear equation. We showed that every solution of the ADHM construction corresponds with an instanton, but we did not show the converse. The completeness of the ADHM construction, i.e. that all anti-self-dual solutions can be found by the ADHM construction, can be proved in several ways. For instance, one could count all solutions and conclude that the number of solutions equals the dimension of the moduli space, which is done in [4]. Another method is by the use of Penrose twistor spaces [1]. A modern proof without the explicit use of twistor spaces is given in [13].

Finally, we gave a short introduction to quivers and discussed the ADHM quiver, whose representations in some cases turn out to be solutions of the ADHM equations. We briefly discussed that some adjusted ADHM quivers correspond with the ADHM construction on ALE spaces. One could ask which other adjusted ADHM quivers are still relevant. For instance, what happens if one changes the relations between the paths in the ADHM quiver? Are the representations of that quiver still physical relevant?

Furthermore, quivers are interesting because of their link with algebras. For instance there is a strong link between quivers and the study of affine Lie algebras, since generalized Coxeter diagrams can be seen as quivers without orientation. Other examples of modern research fields in which quivers play an important role are the fields of Hall algebras and quantum groups. So it is reasonable to conclude that we can expect many interesting results in the research of quivers in the (nearby) future.

A Differential Geometry

In this section we assume that M is an m-dimensional orientable manifold with either a Riemannian or a Lorentzian metric g. Furthermore, $\mathfrak{g}$ is a Lie algebra with generators $\{T_i\}$.

A.1 Lie algebra-valued p-forms

We introduced the space $\Omega^r(U,F)$ of r-forms on a subset U of M with values in a manifold F in Definition 1.3.1. If F is an vector space with basis $\{T_i\}$, we can find a basis of $\Omega^r(U,F)$ by elements of the form $\eta^i T_j$, where $\{\eta^i\}$ is a basis of $\Omega^r(U)$. This means that every $\alpha\in\Omega^r(X,F)$ can be written as $\alpha=\alpha^i T_i$, where α^i are ordinary r-forms. From now on we shall assume that E and F are vector space with an action $E\times F\to F$ of E on F, for instance if $E=\mathrm{End}(F)$. Notice that if F is an algebra (for instance if F is a subspace of $\mathrm{End}(V)$ for some vector space V), it also acts on itself by left multiplication. We assume that $\{V_i\}$ is a basis of E and $\{T_i\}$ a basis of F.

Definition A.1.1. We extend the definition of the wedge product of ordinary forms to $\wedge:\Omega^p(U,E)\times\Omega^q(U,F)\to\Omega^{p+q}(U,F)$ by

$$\alpha\wedge\beta=\frac{1}{p!q!}\sum_{P\in S_{p+q}}\mathrm{sign}(P)\alpha(X_{P(1)},\ldots,X_{P(p)})\beta(X_{P(p+1)},\ldots,X_{P(p+q)}),\tag{A.1.1}$$

where $\alpha\in\Omega^p(U,E)$, $\beta\in\Omega^q(U,F)$, S_n the permutation group of order n and $X_1,\ldots,X_{p+q}$ vector fields on U. If $E=F$ is an algebra, we can also define the commutator $[\alpha,\beta]$ by

$$[\alpha,\beta]=\alpha\wedge\beta-(-1)^{pq}\beta\wedge\alpha.\tag{A.1.2}$$

Remarks.

1. If we write $\alpha=A\otimes\eta$ and $\beta=B\otimes\xi$ with $A\in E$ and $B\in F$, and η,ξ ordinary forms, we have

$$\alpha\wedge\beta=AB\eta\wedge\xi,\tag{A.1.3}$$

where $\wedge$ is the usual wedge product between ordinary forms.

2. If $E=F$ is an algebra, we have

$$[\alpha,\beta]=[A,B]\eta\wedge\xi,\tag{A.1.4}$$

which follows from $[\alpha,\beta]=AB\eta\wedge\xi-(-1)^{pq}BA\xi\wedge\eta$. We can also choose $\alpha^i\in\Omega^p(U)$ and $\beta^i\in\Omega^q(U)$ such that $\alpha=V_i\alpha^i$ and $\beta=T_i\beta^i$. Then we have

$$\alpha\wedge\beta=V_iT_j\alpha^i\wedge\beta^j.\tag{A.1.5}$$

3. If $E=F$ is an algebra, and $\alpha=T_i\alpha^i$ and $\beta=T_i\beta^i$, then the commutator $[\alpha,\beta]$ satisfies

$$[\alpha,\beta]=[T_i,T_j]\alpha^i\wedge\beta^j,\tag{A.1.6}$$

where $[T_i,T_j]$ is the usual matrix commutator of T_i and T_j.

4. Let $E=F$ be a space of $n\times n$-matrices. Then for $\alpha_i\in\Omega^p(U,F)$, we have

$$\mathrm{tr}(\alpha_1\wedge\alpha_2\wedge\ldots\wedge\alpha_k)=(-1)^{kp}\mathrm{tr}(\alpha_2\wedge\ldots\wedge\alpha_k\wedge\alpha_1),\tag{A.1.7}$$

since $\mathrm{tr}(\alpha_1\wedge\alpha_k)=\mathrm{tr}(T_{i_1}T_{i_2}\ldots T_{i_k})\alpha_1^{i_1}\wedge\alpha_2^{i_2}\wedge\ldots\wedge\alpha_k^{i_k}$, whence the identity follows from the fact that the trace is cyclic and from $\alpha\wedge\beta=(-1)^pr\beta\wedge\alpha$ for ordinary forms $\alpha\in\Omega^p(U)$ and $\beta\in\Omega^r(U)$.

5. It follows immediately from the definition that both $[\cdot,\cdot]$ and $\wedge$ are bilinear.

Proposition A.1.2. Let $\alpha \in \Omega^p(U,F)$, $\beta \in \Omega^q(U,F)$ and $\gamma \in \Omega^r(U,F)$. Then we have the following identities:

$$[\alpha,\beta] = -(-1)^{pq}[\beta,\alpha] \tag{A.1.8}$$

$$[\alpha,\alpha] = \begin{cases} 2\alpha\wedge\alpha, & p \text{ is odd;} \\ 0, & p \text{ is even.} \end{cases} \tag{A.1.9}$$

Proof. The first identity follows from

$$\begin{aligned} [\alpha,\beta] &= [T_i,T_j]\alpha^i\wedge\beta^j \\ &= T_iT_j\alpha^i\wedge\beta^j - T_jT_i\alpha^i\wedge\beta^j \\ &= T_iT_j\alpha^i\wedge\beta^j - (-1)^{pq}T_jT_i\beta^j\wedge\alpha^i \\ &= [T_i,T_j]\alpha^i\wedge\beta^j. \end{aligned}$$

The second and third identities follow directly from the first. □

Remark. If $p=1$, the second identity implies

$$[\alpha,\alpha](X,Y) = 2[\alpha(X),\alpha(Y)], \tag{A.1.10}$$

where X and Y are vector fields, since

$$\begin{aligned} [\alpha,\alpha](X,Y) &= 2\alpha\wedge\alpha(X,Y) \\ &= 2T_iT_j\alpha^i\wedge\alpha^j(X,Y) \\ &= 2T_iT_j\Big(\alpha^i(X)\alpha^j(Y) - \alpha^i(Y)\alpha^j(X)\Big) \\ &= 2\Big(\alpha(X)\alpha(Y) - \alpha(Y)\alpha(X)\Big) \\ &= 2[\alpha(X),\alpha(Y)]. \end{aligned}$$

A.2 The Exterior Derivative

Definition A.2.1. Let $U \subset M$. We define the linear map $\mathrm{d} : \Omega^k(U) \to \Omega^{k+1}(U)$ by the following properties:

1. If $f \in \Omega^0(U)$, then $\mathrm{d}f$ is defined by $\mathrm{d}f(X) = Xf$, where X is a vector field on U.
2. if $\alpha \in \Omega^p(U)$ and $\beta \in \Omega^q(U)$, then $\mathrm{d}(\alpha\wedge\beta) = \mathrm{d}\alpha\wedge\beta + (-1)^p\alpha\wedge\mathrm{d}\beta$.
3. $\mathrm{d}\circ\mathrm{d} = 0$.

We define $\mathrm{d} : \Omega^k(U,F) \to \Omega^{k+1}(U,F)$ by $\mathrm{d}\alpha = A\otimes\mathrm{d}\eta$ for $\alpha \in \Omega^k(U,E)$, where $\eta \in \Omega^k(U)$ and $A \in E$ such that $\alpha = A\otimes\eta$. Notice that this implies that if we write $\alpha = \alpha^iT_i$ for $\alpha_i \in \Omega^k(U)$, we have $\mathrm{d}\alpha = \mathrm{d}\alpha^iT_i$.

Proposition A.2.2. Let $\alpha \in \Omega^p(U,E)$ and $\beta \in \Omega^q(U,F)$. Then

$$\mathrm{d}(\alpha\wedge\beta) = \mathrm{d}\alpha\wedge\beta + (-1)^p\alpha\wedge\mathrm{d}\beta \tag{A.2.1}$$

and

$$\mathrm{d}[\alpha,\beta] = [\mathrm{d}\alpha,\beta] + (-1)^p[\alpha,\mathrm{d}\beta], \tag{A.2.2}$$

if $E = F = \mathrm{End}(V)$ for some vector space V.

Proof. Write $\alpha = \alpha^i V_i$ and $\beta = \beta^i T_i$ with $\alpha^i \in \Omega^p(U)$ and $\beta^i \in \Omega^q(U)$. Then we have

$$\begin{aligned} \mathrm{d}(\alpha \wedge \beta) &= V_i T_j \mathrm{d}(\alpha^i \wedge \beta^j) && \text{[By } (A.1.5)] \\ &= V_i T_j (\mathrm{d}\alpha^i \wedge \beta^j + (-1)^p \alpha^i \wedge \mathrm{d}\beta^j) && \text{[By Definition } A.2.1] \\ &= \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta && \text{[By } (A.1.5)]. \end{aligned}$$

Using the first identity, we find

$$\begin{aligned} \mathrm{d}[\alpha, \beta] &= \mathrm{d}(\alpha \wedge \beta - (-1)^{pq} \beta \wedge \alpha) && \text{[By } (A.1.2)] \\ &= \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta - (-1)^{pq} \mathrm{d}\beta \wedge \alpha - (-1)^{p^2 q} \beta \wedge \mathrm{d}\alpha \\ &= \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta - (-1)^{p^2 q} \mathrm{d}\beta \wedge \alpha - (-1)^{pq} \beta \wedge \mathrm{d}\alpha && \left[(-1)^p = (-1)^{p^2}\right] \\ &= [\mathrm{d}\alpha, \beta] + (-1)^p [\alpha, \mathrm{d}\beta] && \text{[By } (A.1.2)]. \end{aligned}$$

□

Proposition A.2.3. Let $\omega \in \Omega^1(U, F)$. Then for any two vector fields X and Y we have

$$\mathrm{d}\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]). \tag{A.2.3}$$

Proof. First assume $\omega \in \Omega^1(U)$. Let v be a smooth function. Since $\Omega^1(U)$ is one-dimensional, it is spanned by $\mathrm{d}v$, so there is a smooth function u such that $\omega = u\mathrm{d}v$. Then the left-hand side of (A.2.3) equals

$$\begin{aligned} \mathrm{d}\omega(X, Y) &= \mathrm{d}(u\mathrm{d}v)(X, Y) = \mathrm{d}u \wedge \mathrm{d}v(X, Y) + u\mathrm{d} \circ \mathrm{d}(X, Y) \\ &= \mathrm{d}u \wedge \mathrm{d}v(X, Y) \\ &= \mathrm{d}u(X)\mathrm{d}u(Y) - \mathrm{d}v(X)\mathrm{d}u(Y). \end{aligned}$$

The right-hand side is

$$\begin{aligned} X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]) &= X(u\mathrm{d}v(Y)) - Y(u\mathrm{d}v(X)) - u\mathrm{d}v([X, Y]) \\ &= X(uYv) - Y(uXv) - u[X, Y]v \\ &= (XuYu - uXYv) - (YuXv + uYXv) - (u(XYv - YXv), \end{aligned}$$

which is equal to the left-hand side after canceling the $uXYv$ and $uYXv$ terms. Now let $\omega \in \Omega^1(U, F)$. Since we can write $\mathrm{d}\omega = \mathrm{d}\omega^i T_i$, $X(\omega(Y)) = X(\omega^i(Y))T_i$ and $\omega([X, Y]) = \omega^i([X, Y])T_i$, and ω^i satisfies (A.2.3), it follows that ω satisfies (A.2.3). □

Remark. A more general formule can be given for k-forms with $k \geq 1$. For all $\alpha \in \Omega^k(U, F)$ and any smooth vector fields $X, \ldots, X_{k+1}$ on U we have

$$\begin{aligned} \mathrm{d}\alpha(X_1, \ldots, X_{k+1}) = &\sum_{i=1}^{k+1} (-1)^{i-1} X_i\Big(\alpha(X_1, \ldots, \widehat{X_i}, \ldots, X_{k+1})\Big) \\ &+ \sum_{i<j} (-1)^{i+j} \alpha\Big([X_i, X_j], X_1, \ldots, \widehat{X_i}, \ldots, \widehat{X_j} \ldots, X_{k+1}\Big), \end{aligned} \tag{A.2.4}$$

where that hats indicate omitted arguments. This is proved in [27] as Proposition 12.19.

The following proposition will be helpful.

Proposition A.2.4. Let $g : M \to G$ be some smooth function from M to a Lie group G. Then we have the following identity

$$g\mathrm{d}(g^{-1}) = -(\mathrm{d}g)g^{-1}. \tag{A.2.5}$$

Proof. Since the differential of a constant function vanishes, we have

$$0 = \mathrm{d}e = \mathrm{d}(gg^{-1}) = (\mathrm{d}g)g^{-1} + g\mathrm{d}(g^{-1}).$$

□

A.3 Integration

Definition A.3.1. If U is a chart with coordinates x^μ, we define the *invariant volume element*

$$\mathrm{dVol}(g) = \sqrt{|g|}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \ldots \wedge \mathrm{d}x^m, \tag{A.3.1}$$

where $|g| = \det g_{\mu\nu}$, the determinant of the metric.

Proposition A.3.2. $\mathrm{dVol}(g)$ is invariant under coordinate transformations.

Proof. If y^ρ are coordinates on V with $U \cap V \neq \emptyset$, we have by $\mathrm{d}y^\rho = \frac{\partial y^\rho}{\partial x^\mu}\mathrm{d}x^\mu$

$$\begin{aligned}
\mathrm{dVol}(g) &= \sqrt{\left|\det\left(\frac{\partial x^\mu}{\partial y^\rho}\frac{\partial x^\nu}{\partial y^\sigma}\right)\right|}\mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^m \\
&= \left|\det\left(\frac{\partial x^\mu}{\partial y^\rho}\right)\right|\sqrt{|g|}\det\left(\frac{\partial y^\sigma}{\partial x^\nu}\right)\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m \\
&= \pm\sqrt{|g|}\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m,
\end{aligned} \tag{A.3.2}$$

where we have a positive sign if x^μ and y^ρ define the same orientation, since in that case we have $\det\left(\frac{\partial x^\mu}{\partial y^\rho}\right) > 0$. ☐

Definition A.3.3. If $f : M \to \mathbb{R}$ is continuous, we define the integral of f over M by

$$\int_M f\mathrm{dVol}(g) = \int_M f\sqrt{|g|}\mathrm{d}x^1\mathrm{d}x^2 \ldots \mathrm{d}x^m. \tag{A.3.3}$$

The integral is finite if f has a compact support, which is always the case if M is compact.

A.4 The Hodge star

In this subsection we assume that $U \subset M$ is a coordinate neighborhood and F is a vector space, which acts on itself, and let $\{T_i\}$ be a basis for F. We shall introduce the Hodge star operation, which allows us to define a natural isomorphism between $\Omega^r(U, F)$ and $\Omega^{m-r}(U, F)$. First we recall the definition of a metric tensor.

Definition A.4.1. Let E be a vector bundle over M. Then a *metric tensor* is a map $g : E \times E \to \mathbb{R}$ such that the restriction to the fibres $g_p : E_p \times E_p \to \mathbb{R}$ is a non-degenerate bilinear form.

If $E = TM$, g is a tensor of signature $(0, 2)$ and we write $g = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$, so that we have $g(V, W) = g_{\mu\nu}V^\mu W^\nu$, where $V|_p = V^\mu\partial_\mu|_p, W = W^\mu\partial_\mu|_p \in T_pM$. The metric defined a natural isomorphism between the tangent bundle TM and the cotangent bundle T^*M by $V|_p \mapsto g_p(V|_p, \cdot)$, for $V|_x \in T_xM$. In coordinates this is $V^\mu\partial_\mu \mapsto V_\mu\mathrm{d}x^\mu$, where $V_\nu = g_{\mu\nu}V^\mu$. We notate this isomorphism with $I : TM \to T^*M$. We shall also denote the induced metric on T^*M by $g : T^*M \times T^*M \to \mathbb{R}$, which is given by $g(\omega, \eta) = g(I^{-1}(\omega), I^{-1}(\xi))$, with $\omega, \xi \in T^*M$. In coordinates this is $g(\omega, \xi) = g^{\mu\nu}\omega_\mu\xi_\nu$, with ω_μ, ξ_μ such that $\omega = \omega_\mu\mathrm{d}x^\mu$, $\xi = \xi_\mu\mathrm{d}x^\mu$, and $g^{\mu\nu}$ the inverse of $g_{\mu\nu}$. As we already have done before, we shall use g to raise and lower indices: $V_\mu = g_{\mu\nu}V^\nu$, $\omega^\mu = g^{\mu\nu}\omega_\nu$. Before we extent g to $\Lambda^r(T^*M)$, we introduce the following notation.

Definition A.4.2. $I = (\mu_1, \mu_2, \ldots, \mu_r)$ with $1 \leq \mu_1 < \mu_2 < \ldots < \mu_r \leq m$ is called a *multi-index of length* $|I| = r$. We denote with $\bar{I}$ the multi-index of length $m - r$ such that $I \cup \bar{I} = \{1, 2, \ldots, m\}$.

If we use the notation $\mathrm{d}x^I = \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$, we have $\{\mathrm{d}x^I|_p : I$ is a multi-index of length $r\}$ as a basis for $\Lambda^r(T_p^*M)$. Now, if $J = (\nu_1, \nu_2, \ldots, \nu_r)$ is also an multi-index, we can extent g to the basis elements of $\Lambda^r(T^*M)$ by

$$g(\mathrm{d}x^I, \mathrm{d}x^J) = g^{\mu_1\nu_1} \ldots g^{\mu_r\nu_r}. \tag{A.4.1}$$

and to general elements of $\Lambda^r(T^*M)$ by bilinear extension. That is, for $\omega = \omega_{\mu_1\ldots\mu_r}\mathrm{d}x^{\mu_1} \wedge \ldots x^{\mu_r}$ and $\xi = \xi_{\nu_1\ldots\nu_r}\mathrm{d}x^{\nu_1} \wedge \ldots x^{\nu_r}$ in $\Omega^r(U)$, we have

$$g(\omega, \xi) = g^{\mu_1\nu_1} \ldots g^{\mu_r\nu_r}\omega_{\mu_1\mu_2\ldots\mu_r}\xi_{\nu_1\nu_2\ldots\nu_r}, \tag{A.4.2}$$

or if we use the metric g to higher or lower indices

$$g(\omega, \xi) = \omega_{\mu_1\mu_2\ldots\mu_r}\xi^{\mu_1\nu_2\ldots\mu_r}. \tag{A.4.3}$$

Since $\Omega^r(U)$ is defined as the space of sections of $\Lambda^r(T^*U)$, g acts also on r-forms. We can use the metric to define an inner product on $\Omega^r(U, F)$.

Definition A.4.3. Let $\alpha, \beta \in \Omega^r(U, \mathfrak{g})$ and let $\mu : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ be an inner product. Then we define

$$\begin{aligned} \langle \alpha, \beta \rangle &= \mu(T_a, T_b)g(\alpha^a, \beta^b) \\ (\alpha, \beta) &= \int \langle \alpha, \beta \rangle \mathrm{dVol}(g). \end{aligned}$$

Furthermore, we shall write $|\alpha| = \langle \alpha, \alpha \rangle$ and $\|\alpha\| = (\alpha, \alpha)$.

Remark. If $\mathfrak{g} \subset \mathfrak{u}(n)$ then by Theorem 1.4.19 the map $\mu(A, B) = -\lambda \operatorname{tr}(AB)$ with $\lambda > 0$ is an inner product on $\mathfrak{g}$, so that we can take $\langle \alpha, \beta \rangle = -\lambda \operatorname{tr}(T_aT_b)g(\alpha^a, \beta^b)$. The factor λ is called the *normalization factor*. For $\mathfrak{g} = \mathfrak{su}(2)$, we shall we have the generators $\tau_i := -i\sigma_i$ with σ_i the Pauli matrices. We usually take $\lambda = 1$, so that we have $\mu(\tau_i, \tau_j) = 2\delta_{ij}$, but sometimes it is more convenient to have $\mu(\tau_i, \tau_j) = \delta_{ij}$, for which we have to take $\lambda = \frac{1}{2}$.

Now let $p : \Omega^{m-r}(U, F) \to \mathrm{Lin}\Big(\Omega^k(U, F), \Omega^m(U, F)\Big)$ be given by $p(\alpha)(\beta) = \beta \wedge \alpha$. Clearly p is an isomorphism. Notice that $\Omega^m(U)$ can be spanned by $\mathrm{dVol}(g)$, so we can define another isomorphism $m : \Omega^r(U, F) \to \mathrm{Lin}\Big(\Omega^r(U, F), \Omega^m(U, F)\Big)$ by $\alpha \mapsto \langle \alpha, \cdot \rangle \mathrm{dVol}(g)$. Now we are able to define the Hodge star operator.

Definition A.4.4. The *Hodge star operator* is the operator $* : \Omega^r(U, F) \to \Omega^{m-r}(U, F)$ defined by $* = p^{-1} \circ m$.

Since p and m are isomorphisms, $*$ is also an isomorphism.

Proposition A.4.5. Let $O : \Omega^r(U, F) \to \Omega^{m-r}(U, F)$ be a linear operator satisfying

$$\beta \wedge O(\alpha) = \langle \alpha, \beta \rangle \mathrm{dVol}(g), \tag{A.4.4}$$

for $\alpha \in \Omega^r(U)$ and $\beta \in \Omega^{m-r}(U)$. Then $O = *$.

Proof. First we show that $*$ satisfies (A.4.4):

$$\begin{aligned} \beta \wedge *\alpha &= p(*\alpha)(\beta) \\ &= p \circ p^{-1} \circ m(\alpha)(\beta) \\ &= m(\alpha)(\beta) \\ &= \langle \alpha, \beta \rangle \mathrm{dVol}(g). \end{aligned}$$

Then we have $\beta \wedge \Big(* \alpha - O(\alpha)\Big) = \beta \wedge *\alpha - \beta \wedge O(\alpha) = 0$, from which we easily see that $O = *$. ☐

Corollary A.4.6. For $\mathfrak{g} \subset \mathfrak{u}(n)$ the Hodge star allows us to write the inner product $(\cdot,\cdot)$ on $\Omega^r(U,\mathfrak{g})$ in the following way:

$$(\alpha,\beta) = -\int \mathrm{tr}(\alpha \wedge *\beta). \tag{A.4.5}$$

In order to give a local description of the Hodge star, let x^μ be the coordinates on U such that at $p \in U$ we have

$$g(\partial_\mu, \partial_\nu) = \begin{cases} \epsilon_\mu, & \mu = \nu; \\ 0, & \mu \neq \nu. \end{cases} \tag{A.4.6}$$

where $\epsilon_\mu = \pm 1$. If t be the number of negative ϵ_μ, we have $t = 0$ in the case that M is Riemannian and $t = 1$ if M is Lorentzian. We have now

$$\langle \mathrm{d}x^I, \mathrm{d}x^J \rangle = \begin{cases} \epsilon_I, & I = J; \\ 0, & I \neq J, \end{cases} \tag{A.4.7}$$

where $\epsilon_I = \epsilon_{\mu_1}\epsilon_{\mu_2}\ldots\epsilon_{\mu_r}$. If $I \sqcup \bar{I}$ is the concatenation of I and $\bar{I}$, we denote with ζ_I the sign of the permutation $(1,2,\ldots,m) \to I \sqcup \bar{I}$. Then the local version of the Hodge operator is given by

$$*\mathrm{d}x^I = \epsilon_I \zeta_I \mathrm{d}x^{\bar{I}}, \tag{A.4.8}$$

which can also be written as

$$*\left(\mathrm{d}x^{\mu_1} \wedge \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}\right) = \frac{\sqrt{|g|}}{(m-r)!}\epsilon^{\mu_1\mu_2\ldots\mu_r}{}_{\nu_{r+1}\ldots\nu_m}\mathrm{d}x^{\nu_{r+1}} \wedge \ldots \wedge \mathrm{d}x^{\nu_m}. \tag{A.4.9}$$

This local form allows us to give a formula for how the Hodge star transforms under a conformal transformation $g \mapsto e^{2f}g$.

Proposition A.4.7. Let $\bar{g} = e^{2f}g$ be a conformal rescaled metric. If $\alpha \in \Omega^r(U,F)$, we have $*_{\bar{g}}\alpha = e^{(m-2r)f} *_g \alpha$. In special if $2r = m$, we have $*_{\bar{g}}\alpha = *_g\alpha$.

Proof. First consider the case that α is a basis element of $\Omega^r(U)$. So $\alpha = \mathrm{d}x^{\mu_1} \wedge \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$ with $1 \leq \mu_1 < \mu_2 \ldots < \mu_r \leq m$. Then $*_g\alpha$ is exactly the left-hand side of (A.4.9), while we can rewrite the right-hand side so that we obtain

$$*_g\alpha = \frac{\sqrt{|g|}}{(m-r)!}g^{\mu_1\nu_1}g^{\mu_2\nu_2}\ldots g^{\mu_r\nu_r}\epsilon_{\nu_1\ldots\nu_m}\mathrm{d}x^{\nu_{r+1}} \wedge \ldots \wedge \mathrm{d}x^{\nu_m}.$$

Now by (A.4.6), g can locally be written as a diagonal matrix, and therefore $\bar{g}$ is also a diagonal matrix with the same entries as g but multiplied with a factor e^{2f}. So we find that $\det\bar{g} = e^{2mf}\det g$, whence $\sqrt{|\bar{g}|} = e^{mf}\sqrt{|g|}$. Then we find

$$\begin{aligned} *_{\bar{g}}\alpha &= \frac{\sqrt{|\bar{g}|}}{(m-r)!}\bar{g}^{\mu_1\nu_1}\bar{g}^{\mu_2\nu_2}\ldots\bar{g}^{\mu_r\nu_r}\epsilon_{\nu_1\ldots\nu_m}\mathrm{d}x^{\nu_{r+1}} \wedge \ldots \wedge \mathrm{d}x^{\nu_m} \\ &= e^{(m-2r)f}\frac{\sqrt{|g|}}{(m-r)!}g^{\mu_1\nu_1}g^{\mu_2\nu_2}\ldots g^{\mu_r\nu_r}\epsilon_{\nu_1\ldots\nu_m}\mathrm{d}x^{\nu_{r+1}} \wedge \ldots \wedge \mathrm{d}x^{\nu_m} \\ &= e^{(m-2r)f} *_g \alpha. \end{aligned}$$

Now, since the identity holds for basis elements, it holds for general $\alpha \in \Omega^r(U)$. Furthermore, if follows also that the identity holds for $\alpha \in \Omega^r(U,F)$, since $*\alpha = A \otimes *\eta$ if $A \in F$ and $\eta \in \Omega^r(U)$ such that $\alpha = A \otimes \eta$. ∎

Proposition A.4.8. The map $*^2 : \Omega^r(U,F) \to \Omega^r(U,F)$ is given by $*^2 = (-1)^{r(m-r)+t}\mathrm{Id}$, where $t=0$ if M is Riemannian and $t=1$ if M is Lorentzian.

Proof. We have for $\mathrm{d}x^I \in \Omega^r(U)$

$$\begin{aligned} *^2\mathrm{d}x^I &= \epsilon_I\zeta_I * \mathrm{d}x^{\bar{I}} \\ &= \epsilon_I\epsilon_{\bar{I}}\zeta_I\zeta_{\bar{I}}\mathrm{d}x^I, \end{aligned}$$

where $\epsilon_I\epsilon_{\bar{I}} = (-1)^t$ and $\zeta_I\zeta_{\bar{I}} = (-1)^{|I||\bar{I}|} = (-1)^{r(m-r)}$. Since $*\alpha = A \otimes *\eta$ for $A \in F$ and $\eta \in \Omega^r(U)$ such that $\alpha = A \otimes \eta$, it follows that the identity is also valid for $\Omega^r(U,F)$. □

Theorem A.4.9. Let $\mathrm{d}^* : \Omega^r(U,F) \to \Omega^{r-1}(U,F)$ be the adjoint of d:

$$\langle \mathrm{d}\beta, \alpha\rangle = \langle \beta, \mathrm{d}^*\alpha\rangle. \tag{A.4.10}$$

Then we have

$$\mathrm{d}^* = (-1)^{(m-r)(r-1)+1+t} * \mathrm{d}* \tag{A.4.11}$$

on the subspace of F-valued r-forms with compact support (which is whole $\Omega^r(U,F)$ if M is compact).

Proof. Let $\alpha \in \Omega^r(U,F)$ and $\beta \in \Omega^{r-1}(U,F)$. Then by Proposition A.2.2 we have

$$\mathrm{d}(\beta \wedge *\alpha) = \mathrm{d}\beta \wedge *\alpha - (-1)^r\beta \wedge \mathrm{d} * \alpha. \tag{A.4.12}$$

Since $\mathrm{d} * \alpha$ is an $(m-r+1)$ form, we have

$$\begin{aligned} **(\mathrm{d}*\alpha) &= (-1)^{(m-r+1)(m-[m-r+1])+t}\mathrm{d}*\alpha \\ &= (-1)^{(m-r+1)(r-1)+t}\mathrm{d}*\alpha \\ &= (-1)^{mr-r^2+r-m+r-1+t}\mathrm{d}*\alpha \\ &= (-1)^{mr-r^2-m-1+t}\mathrm{d}*\alpha, \end{aligned} \tag{A.4.13}$$

since $(-1)^{2r}=1$. Hence we find $\mathrm{d}*\alpha = (-1)^{mr-r^2-m-1+t} **(\mathrm{d}*\alpha)$, so from (A.4.12) we find

$$\mathrm{d}(\beta \wedge *\alpha) = \mathrm{d}\beta \wedge *\alpha - \beta \wedge *[(-1)^{(m-r)(r-1)+1+t} * \mathrm{d} * \alpha], \tag{A.4.14}$$

where we used $(-1)^r(-1)^{mr-r^2-m-1+t} = (-1)^{(m-r)(r-1)-1+t}$ and $(-1)^2=1$. Then with Stokes' theorem, we find

$$\begin{aligned} \int_M \mathrm{d}\beta \wedge *\alpha - \int_M \beta \wedge *(-1)^{(m-r)(r-1)+1+t} * \mathrm{d}*\alpha] &= \int_M \mathrm{d}(\beta\wedge *\alpha) \\ &= \int_{\partial M} \beta \wedge *\alpha \\ &= 0, \end{aligned}$$

which proves the theorem. □

Proposition A.4.10. With respect to this inner product defined above $[\mathcal{A},\cdot]^* : \Omega^r(U,\mathfrak{g}) \to \Omega^{r-1}(U,\mathfrak{g})$ satisfies $[\mathcal{A},\cdot]^* = (-1)^{(m-r)(r-1)+1+t} * [\mathcal{A}, *\cdot]$.

Proof. Let $\beta \in \Omega^{r-1}(U,\mathfrak{g})$ and $\gamma \in \Omega^r(U,\mathfrak{g})$. Then we can write $\mathcal{A} = \mathcal{A}^aT_a$, $\beta = \beta^bT_b$ and $\gamma = T_c\gamma^c$. We have $[T_a,T_b] = f_{ab}{}^cT_c$, hence by (A.1.6) we can write

$$[\mathcal{A},\beta] = f_{ab}{}^dT_d\mathcal{A}^a \wedge \beta^b; \tag{A.4.15}$$
$$[\mathcal{A},*\gamma] = f_{ac}{}^dT_d\mathcal{A}^a \wedge *\gamma^c. \tag{A.4.16}$$

So we find

$$\begin{aligned} ([\mathcal{A},\beta],\gamma) &= (f_{ab}{}^d T_d \mathcal{A}^a \wedge \beta^b, T_c\gamma^c) \\ &= f_{ab}{}^d \mathrm{tr}(T_d T_c) \int \mathcal{A}^a \wedge \beta^b \wedge *\gamma^c. \end{aligned}$$

At the other hand we have

$$\begin{aligned} (\beta, *[\mathcal{A}, *\gamma]) &= \Big(T_b\beta^b, f_{ac}{}^d T_d * (\mathcal{A}^a \wedge *\gamma^c)\Big) \\ &= f_{ac}{}^d \mathrm{tr}(T_b T_d) \int \beta^b \wedge ** (\mathcal{A}^a \wedge *\gamma^c) \\ &= (-1)^{(m-r+1)(r-1)+t} f_{ac}{}^d \mathrm{tr}(T_b T_d) \int \beta^b \wedge \mathcal{A}^a \wedge *\gamma^c \\ &= (-1)^{(m-r)(r-1)+t} f_{ac}{}^d \mathrm{tr}(T_b T_d) \int \mathcal{A}^a \wedge \beta^b \wedge *\gamma^c \end{aligned}$$

where we used in the third equality that $\mathcal{A} \wedge *\gamma$ and $*\gamma \wedge \mathcal{A}$ are $(m-r+1)$ forms, for which $** = (-1)^{(m-r+1)(r-1)+t}$, and in the last equality we used that the interchange of β^b and $\mathcal{A}^a$ gives a factor $(-1)^{r-1}$, since β is an $(r-1)$ form. We also used $(-1)^{2(r-1)} = 1$. Now Since $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$ for arbitrary $n \times n$-matrices A, B and C, we find

$$\begin{aligned} f_{ab}{}^d \mathrm{tr}(T_d T_c) &= \mathrm{tr}([T_a, T_b] T_c) \\ &= \mathrm{tr}\Big((T_a T_b - T_b T_a) T_c\Big) \\ &= \mathrm{tr}(T_a T_b T_c) - \mathrm{tr}(T_b T_a T_c) \\ &= \mathrm{tr}(T_b T_c T_a) - \mathrm{tr}(T_b T_a T_c) \\ &= \mathrm{tr}(T_b [T_c, T_a]) \\ &= f_{ca}{}^d \mathrm{tr}(T_b T_d) \\ &= -f_{ac}{}^d \mathrm{tr}(T_b T_d). \end{aligned}$$

So we find $([\mathcal{A},\beta],\gamma) = (-1)^{(m-r)(r-1)+1+t}(\beta, *[\mathcal{A}, *\gamma])$, which proves the proposition. □

This proposition combined with Theorem A.4.9 proves the following:

Theorem A.4.11. Let $\mathcal{A} \in \Omega^1(U, \mathfrak{g})$ and let $\mathcal{D}_\mathcal{A}$ be the covariant derivative $\mathcal{D}_\mathcal{A} = \mathrm{d} + [\mathcal{A}, \cdot]$. Then we have

$$\mathcal{D}^*_\mathcal{A} = (-1)^{(m-r)(r-1)+1+t} * \mathcal{D}_\mathcal{A} * . \tag{A.4.17}$$

B Complex geometry

B.1 Complex manifolds

Definition B.1.1. A manifold M is called *complex* if it has a atlas of charts to open subsets in $\mathbb{C}^n$ such that the transition maps are holomorphic. We say that M has complex dimension $\dim_{\mathbb{C}} M = n$, whereas the real dimension is $\dim_{\mathbb{R}} M = 2n$. A manifold M is called *almost complex* if there exists a $(1,1)$-tensor field J which acts as vector bundle isomorphism $J : TM \to TM$ such that $J^2 = -1$. The map J is called an *almost complex structure* on M.

Definition B.1.2. Let V be a real vector space. Then we define its *complexification* $V^{\mathbb{C}}$ by

$$V^{\mathbb{C}} = V \otimes \mathbb{C} = \{X + iY : X, Y \in V\}.$$

Notice that V is a subspace of $V^{\mathbb{C}}$, since every $X \in V$ can be identified with $X + i0 \in V^{\mathbb{C}}$. Let M be a complex manifold with $\dim_{\mathbb{C}} M = n$. Then we have $2n$ real coordinates $x^1, \dots, x^n, y^1, \dots, y^n$, which allows us to introduce the following vectors, which span $T_xM^{\mathbb{C}}$:

$$\begin{aligned}
\frac{\partial}{\partial z^\mu} &= \frac{1}{2}\left(\frac{\partial}{\partial x^\mu} - i\frac{\partial}{\partial y^\mu}\right)\\
\frac{\partial}{\partial \bar{z}^\mu} &= \frac{1}{2}\left(\frac{\partial}{\partial x^\mu} + i\frac{\partial}{\partial y^\mu}\right).
\end{aligned}$$

We shall often use the abbreviations ∂_μ and $\partial_{\bar{\mu}}$ instead of $\frac{\partial}{\partial z^\mu}$ and $\frac{\partial}{\partial \bar{z}^\mu}$. The corresponding dual basis of one-forms on $T_x^*M^{\mathbb{C}} = (T_xM^{\mathbb{C}})^*$ is given by

$$\begin{aligned}
\mathrm{d}z^\mu &= \mathrm{d}x^\mu + i\mathrm{d}y^\mu\\
\mathrm{d}\bar{z}^\mu &= \mathrm{d}x^\mu - i\mathrm{d}y^\mu.
\end{aligned}$$

As the term almost complex suggests we have

Lemma B.1.3. Every complex manifold is almost complex.

Proof. We define the map $J_x : T_xM \to T_xM$ by

$$\begin{aligned}
J\left(\frac{\partial}{\partial x^\mu}\right) &= \frac{\partial}{\partial y^\mu};\\
J\left(\frac{\partial}{\partial y^\mu}\right) &= -\frac{\partial}{\partial x^\mu}.
\end{aligned}$$

Clearly we have $J^2 = -1$. $\square$

Note that J depends holomorphic on x and if we extend J linearly to $T_xM^{\mathbb{C}}$, we have $J(\partial_\mu) = i\partial_\mu$, while $J(\partial_{\bar{\mu}}) = -i\partial_{\bar{\mu}}$. Now, J has eigenvalues $\pm i$ on $T_xM^{\mathbb{C}}$, so we can decompose $T_xM^{\mathbb{C}}$ into the eigenspaces of J.

Definition B.1.4. We define $T_xM^\pm$ to be the eigenspace of J corresponding to the eigenvalue i. That is $T_xM^\pm = \{Z \in T_xM^{\mathbb{C}} : J(Z) = \pm iZ\}$. A vector $Z|_x$ is called *holomorphic* if $Z|_x \in T_xM^+$ and *anti-holomorphic* if $Z|_x \in T_xM^-$.

A basis of T_xM^+ is given by the ∂_μ, while a basis of T_xM^- is given by the $\partial_{\bar{\mu}}$. It is convenient to introduce projection operators.

Definition B.1.5. We define the operators $P_x^\pm : T_xM^{\mathbb{C}} \to T_xM^\pm$ by $P_x^\pm = \frac{1}{2}(\mathrm{Id} \mp iJ)$.

Then we have $JP^{\pm}(Z) = \frac{1}{2}\Big(J(Z) \mp iJ^2(Z)\Big) = \frac{1}{2}(J(Z) \pm iZ) = \pm iP^{\pm}(Z)$, so $Z_{\pm} = P^{\pm}Z \in T_xM^{\pm}$. Hence indeed all $Z|_x \in T_xM^{\mathbb{C}}$ can uniquely be decomposed into holomorphic and anti-holomorphic vectors $Z = Z_+ + Z_-$. Thus we have found

Proposition B.1.6. We have the following decomposition: $T_xM^{\mathbb{C}} = T_xM^+ \oplus T_xM^-$.

Definition B.1.7. We define $\mathscr{X}(M)^+ = \{Z \in \mathscr{X}(M)^{\mathbb{C}} : P^+(Z) = Z\}$ to be the space of *holomorphic vector fields*. Similarly, $\mathscr{X}(M)^- = \{Z \in \mathscr{X}(M)^{\mathbb{C}} : P^-(Z) = Z\}$ is defined as the space of *anti-holomorphic vector fields*.

Proposition B.1.8. We can decompose every vector field into a holomorphic and an anti-holomorphic part: $\mathscr{X}(M)^{\mathbb{C}} = \mathscr{X}(M)^+ \oplus \mathscr{X}(M)^-$.

Proof. This follows directly from the decomposition of $T_xM^{\mathbb{C}}$ and the fact that J and thus $P^{\pm}$ depends holomorphically on x. □

B.2 Complex differential forms

Definition B.2.1. Let $\Omega^p(M)$ be the space of real p-forms, then we define the space of complexified p-forms by

$$\Omega^p(M)^{\mathbb{C}} = \{\omega + i\eta : \omega, \eta \in \Omega^p(M)\}.$$

We shall denote the space of complex p-forms at $x \in M$ with $\Omega^p_x(M)^{\mathbb{C}}$. If $\zeta \in \Omega^p(M)^{\mathbb{C}}$ with $\zeta = \omega + i\eta$ where $\omega, \eta \in \Omega^p(M)$, we extent the differential operator d to $\Omega^p(M)^{\mathbb{C}}$ by

$$\mathrm{d}\zeta = \mathrm{d}\omega + i\mathrm{d}\eta.$$

We readily find that $\mathrm{d} : \Omega^p(M)^{\mathbb{C}} \to \Omega^{p+1}(M)$ satisfies (A.1.1) and (A.2.1) with in the first equation $X_1, \ldots, X_{p+q} \in \mathscr{X}(M)^{\mathbb{C}}$ and in both equations $\alpha \in \Omega^p(M)^{\mathbb{C}}$ and $\beta \in \Omega^q(M)^{\mathbb{C}}$. Furthermore, we remark that d is a real operator, since $\overline{\mathrm{d}\zeta} = \mathrm{d}\omega - i\mathrm{d}\eta = \mathrm{d}\bar{\zeta}$.

Definition B.2.2. Let M be a complex manifold with complex dimension n and let $V_1, \ldots, V_k$ with $k = p + q$ be vector fields in $\mathscr{X}(M)^{\mathbb{C}}$ such that V_i either in $\mathscr{X}(M)^+$ or $\mathscr{X}(M)^-$. Then we say that a complex k-form ω is of *bidegree* (p, q), or equivalently ω is a (p, q)-form, if $\omega(V_1, \ldots, V_k) = 0$ unless p of the V_i are in $\mathscr{X}(M)^+$ and q of the V_i are in $\mathscr{X}(M)^-$. The space of (p, q)-forms is denoted with $\Omega^{p,q}(M)$.

Proposition B.2.3. Every (p, q)-form can be written as

$$\omega = \frac{1}{p!q!}\omega_{\mu_1\ldots\mu_p\nu_1\ldots\nu_q}\mathrm{d}z^{\mu_1} \wedge \ldots \wedge \mathrm{d}z^{\mu_p} \wedge \mathrm{d}\bar{z}^{\nu_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{\nu_q}.$$

Proof. This follows directly from the fact that $\langle \mathrm{d}z^{\mu}, \partial_{\bar{\nu}}\rangle = 0$, and so $\mathrm{d}z^{\mu}$ is a $(1, 0)$-form. Similarly, $\mathrm{d}\bar{z}^{\mu}$ is a $(0, 1)$-form, and so the set $\{\mathrm{d}z^{\mu_1} \wedge \ldots \wedge \mathrm{d}z^{\mu_p} \wedge \mathrm{d}\bar{z}^{\nu_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{\nu_q}\}$ is a basis for the (p, q)-forms. □

Corollary B.2.4. We have the decomposition

$$\Omega^k(M)^{\mathbb{C}} = \bigoplus_{p+q=k} \Omega^{p,q}(M).$$

If ω is a (p, q)-form, then $\mathrm{d}\omega$ is a mixture of $(p + 1, q)$- and $(p, q + 1)$-forms. Therefore, we introduce the following operators:

Definition B.2.5. We define $\partial : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M)$ and $\bar{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$ as the operators satisfying $\mathrm{d} = \partial + \bar{\partial}$. If we define $P^{p,q} : \Omega^k(M)^{\mathbb{C}} \to \Omega^{p,q}(M)$ to be the projection operator onto the space of (p, q)-forms, we have $\partial = P^{p+1,q} \circ \mathrm{d}$ and $\bar{\partial} = P^{p,q+1} \circ \mathrm{d}$. The operators ∂ and $\bar{\partial}$ are called *Dolbeault operators*.

B.3 Hermitian manifolds

Let M a complex manifold. Then we can extend Riemannian metrics on TM to $TM^{\mathbb{C}}$:

Proposition B.3.1. Let g be a Riemannian metric on a complex manifold M. Then we can extend g to $TM^{\mathbb{C}}$ by

$$g_x(Z,W) = g_x(X,U) - g_x(Y,V) + i\Big(g_x(X,V) + g_x(Y,U)\Big),$$

where $Z = X + iY, W = U + iV \in \mathscr{X}(M)^C$ with $X, Y, U, V \in T_xM^{\mathbb{C}}$.

Remark. The components of g are given by

$$\begin{aligned} g_{\mu\nu}(x) &= g_x(\partial_\mu, \partial_\nu); \\ g_{\mu\bar{\nu}}(x) &= g_x(\partial_\mu, \partial_{\bar{\nu}}); \\ g_{\bar{\mu}\nu}(x) &= g_x(\partial_{\bar{\mu}}, \partial_\nu); \\ g_{\bar{\mu}\bar{\nu}}(x) &= g_x(\partial_{\bar{\mu}}, \partial_{\bar{\nu}}). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} g_{\mu\nu} &= g_{\nu\mu}; \\ \overline{g_{\mu\nu}} &= g_{\bar{\mu}\bar{\nu}}; \\ \overline{g_{\mu\bar{\nu}}} &= g_{\bar{\mu}\nu}. \end{aligned}$$

Definition B.3.2. A metric on a complex manifold is called *Hermitian* if it satisfies

$$g_x(J_xX, J_xY) = g_x(X,Y)$$

for all vectors $X, Y \in T_xM^{\mathbb{C}}$ and all $x \in M$. A complex manifold with a Hermitian metric is called a *Hermitian manifold*.

Proposition B.3.3. Every complex manifold M admits a Hermitian metric.

Proof. Let g be a Riemannian metric on M. Then $\hat{g}(X,Y) = \frac{1}{2}\Big(g(X,Y) + g(JX,JY)\Big)$ is an Hermitian metric. $\square$

Lemma B.3.4. Let g be a Hermitian metric on a complex manifold M. Then $g_{\mu\nu} = g_{\bar{\mu}\bar{\nu}} = 0$.

Proof. We have

$$g_{\mu\nu} = g(\partial_\mu, \partial_\nu) = g(J\partial_\mu, J\partial_\nu) = -g(\partial_\mu, \partial_\nu) = -g_{\mu\nu},$$

so we find that $g_{\mu\nu} = 0$. Since $g_{\bar{\mu}\bar{\nu}} = \overline{g_{\mu\nu}}$, we also find that $g_{\bar{\mu}\bar{\nu}} = 0$. $\square$

Definition B.3.5. Let M be a complex manifold with Hermitian metric g. Then we define the *Kähler form* ω of g by

$$\omega_x(X,Y) = g_x(J_xX, Y),$$

where $X, Y \in T_xM^{\mathbb{C}}$.

Proposition B.3.6. The Kähler form ω is anti-symmetric, and so indeed a two-form, which is of bidegree $(1,1)$. Furthermore, Ω is invariant under the action of J.

Proof. Let $X, Y \in T_x M^{\mathbb{C}}$. Then

$$\omega(X,Y) = g(JX,Y) = g(J^2X, JY) = -g(X,JY) = -g(JY,X) = -\omega(Y,X).$$

In order to calculate the bidegree, we have $\omega(\partial_\mu, \partial_\nu) = g(J\partial_\mu, \partial\nu) = ig_{\mu\nu} = 0$, so $\omega_{\mu\nu} = 0$. Similarly, we find $\omega_{\bar{\mu}\bar{\nu}} = 0$, and $\omega_{\mu\bar{\nu}} = -\omega_{\bar{\mu}\nu} = ig_{\mu\bar{\nu}}$. Hence we find

$$\begin{aligned}\omega &= \omega_{\mu\bar{\nu}}\mathrm{d}z^\mu \otimes \mathrm{d}\bar{z}^\nu + \omega_{\bar{\mu}\nu}\mathrm{d}\bar{z}^\mu \otimes \mathrm{d}z^\nu \\ &= ig_{\mu\bar{\nu}}\mathrm{d}z^\mu \otimes \mathrm{d}\bar{z}^\nu - ig_{\bar{\mu}\nu}\mathrm{d}\bar{z}^\mu \otimes \mathrm{d}z^\nu \\ &= ig_{\mu\bar{\nu}}\mathrm{d}z^\mu \wedge \mathrm{d}\bar{z}^\nu.\end{aligned} \tag{B.3.1}$$

Finally, ω is invariant under the action of J, since

$$\omega(JX,JY) = g(J^2X,JY) = g(J^3X,J^2Y) = g(JX,Y) = \omega(X,Y).$$

$\square$

References

[1] M.F. Atiyah, *Geometry of Yang-Mills Fields*, Accademia Nazionale Dei Lincei Scuola Normale Superiore (1979).

[2] M.F. Atiyah, N.J. Hitchin, I.M. Singer, *Self-duality in four dimensional Riemannian Geometry*, Proceedings of the Royal Society of London, Series A, **362**, 425-61 (1978).

[3] M. Barot, *Representations of Quivers*, Lecture notes (2006).

[4] C.W. Bernard, N.H. Christ, A.H. Guth, E.J. Weinberg, *Pseudoparticle parameters for arbitrary gauge groups*, Physical Review D, **16**, 2967-77 (1977).

[5] M. Bianchi, F. Fucito, G. Ross, *Explicit Construction of Yang-Mills Instantons on ALE Space*, arXiv:hep-th/9601162v1 (1996).

[6] R. Bott, *An application of the Morse Theory to the topology of Lie-groups*, Bull. Soc. Math. France, **84**, 251-81 (1963).

[7] G. E. Bredon, *Topology and Geometry*, Springer-Verlag (1993).

[8] N.H. Christ, E.J. Weinberg, N.K. Stanton, *General self-dual Yang-Mills solutions*, Physical Review D, **18**, 2013-25 (1978).

[9] S. Coleman, *The uses of instantons*, Aspects of Symmetry 265-350, Cambridge Univservisty Press (1985).

[10] A. Collinucci, A. Wijns, *Topology of fibre bundles and global aspects of gauge theories*, arXiv:hep-th/0611201v1 (2006).

[11] J.B. Conway, *A Course in Functional Analysis*, Second Edition, Springer-Verlag (1990).

[12] H. Derksen, *Quiver representations*, Lecture notes (2001).

[13] S.K. Donaldson, P.B. Kronheimer, *The Geometry of Four-Manifolds*, Oxford University Press (1990).

[14] J.J. Duistermaat, J.A.C. Kolk, *Lie Groups*, Springer-Verlag (2000).

[15] P.M.N. Feehan, *Geometry of the Moduli Space of Self-Dual Connections over the Four-Sphere*, Columbia University (1992).

[16] D.S. Freed, K.K. Uhlenbeck, *Instantons and Four-Manifolds*, Springer-Verlag (1984).

[17] D.J. Griffiths, *Introduction to Electrodynamics, Third Edition*, Prentice-Hall (1999)

[18] B.C. Hall, *Lie Groups, Lie Algebras, and Representations - An Elementery Introduction*, Springer-Verlag (2003).

[19] J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag (1972).

[20] T.A. Ivey, J.M. Landsberg, *Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems*, AMS (2003).

[21] M. Jardin, R. Vidal Martins, *The ADHM Variety and Perverse Coherent Sheaves*, arXiv:1005.0807 (2010)

[22] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry Volume 1*, Interscience Publishers (1963).

[23] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry Volume 2*, Interscience Publishers (1969).

[24] P.B. Kronheimer, H. Nakajima, *Yang-Mills instantons on ALE gravitational instantons*, Math. Ann. 288, 263-307 (1990)

[25] M. Nakahara, *Geometry, Topology and Physics*, Taylor & Francis Group, Second Edition (2003).

[26] H. Nakajima, *Instantons on ALE spaces, quiver varieties, and Kac-Moody Algebras*, Duke Mathematical Journal Vol. 76, No. 2 (1994).

[27] J.M. Lee, *Introduction to Smooth Manifolds*, Springer-Verlag (2003).

[28] L.H. Ryder, *Quantum Field Theory, Second Edition*, Cambridge University Press (1996).

[29] M. Spivak, *A Comprehensive Introduction to Differential Geometry Volume One*, Publish or Perish, Third Edition with corrections (2005).

[30] M. Spivak, *A Comprehensive Introduction to Differential Geometry Volume Two*, Publish or Perish, Third Edition (1999).

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

[32] M.R. Sepanski, *Compact Lie Groups*, Springer-Verlag (2007).

[33] Y. Tian, C.-J. Zhu, *Comments on Noncommutative ADHM Construction*, arXiv:hep-th/0210163v2 (2002).

[34] S. Vandoren, P. Van Nieuwenhuizen, *Lectures on instantons*, arXiv:0802.1862v1 (2008).

Index