Compactness in gauge theory

Patrícia Muñoz Ewald

Dissertation submitted to the Institute of Mathematics and Statistics of the University of São Paulo for the degree of Master of Science

Program: Mathematics
Advisor: Prof. Dr. Cristian Ortiz

During the development of this dissertation the author was financially supported by CNPq

São Paulo, June 2021

Compactness in gauge theory

This is the original version of the dissertation developed by the candidate Patrícia Muñoz Ewald as it was submitted to the Judging Committee.

Ao meu Tata.

Acknowledgements

To Cristian, for the support, patience, and for believing in me.

To my friends, for making life fun and full of coffee.

To my family, for making me who I am.

To Vitor, for everything.

To CNPq for the financial support.

Abstract

In 1982, Karen Uhlenbeck published two foundational papers in gauge theory, which quickly led to Simon Donaldson's Fields medal winning result on topology of four-manifolds, and to the beginning of an era of using gauge theoretic techniques as tools for proving theorems. In 2019, she became the first (and thus far only) woman to receive the Abel prize, for these and other groundbreaking works in geometric analysis.

In one of these works, entitled *Connections with L^p bounds on curvature*, Uhlenbeck proved two very important technical results on the existence of a *good* gauge, and the sequential compactness of weak connections with bounded curvature. In this work, we prove these results and then address their immediate consequence: the uniform convergence of weak Yang-Mills connections with bounded curvature.

Keywords: *gauge theory, geometric analysis, fibre bundles, connections, Uhlenbeck compactness, Yang-Mills equation.*

Resumo

Em 1982, Karen Uhlenbeck publicou dois artigos fundamentais em teoria de gauge, que rapidamente levaram Simon Donaldson ao resultado em topologia de 4-variedades que lhe rendeu a medalha Fields, e ao início de uma era de utilização de técnicas de teoria de gauge em demonstrações. Em 2019, se tornou a primeira (e por enquanto, única) mulher a receber o prêmio Abel, por esses e outros trabalhos revolucionários em análise geométrica.

Em um desses trabalhos, *Connections with L^p bounds on curvature*, Uhlenbeck provou dois resultados técnicos muito importantes sobre a existência de uma boa escolha de gauge, e sobre a compacidade sequencial de conexões fracas com curvatura uniformemente limitada. Neste trabalho, provamos esses resultados e em seguida nos voltamos à uma consequência imediata: a convergência uniforme de sequências de conexões Yang-Mills fracas com curvatura uniformemente limitada.

Palavras-chave: *teoria de gauge, análise geométrica, fibrados, conexões, compacidade de Uhlenbeck, equação de Yang-Mills.*

Contents

Introduction

This work is very technical. It deals with something called gauge theory, and the main results concern existence of objects with good properties, and convergence of sequences in abstract spaces. The motivation for these results is of type "because they are powerful tools for proofs in gauge theory", which may seem tautological and empty for a non mathematician unaware of the larger context. Hence, before getting into the main results and the mathematics of it, we will spend some effort to motivate the area as a whole, and we aim to answer the following three questions in increasing level of abstractness:

1) What is gauge theory?

2) What are the motivations to study gauge theory, mathematical or otherwise?

3) What is this specific manuscript about?

What is gauge theory?

To put it shortly, gauge theory is the study of connections on principal bundles. Unpacking this terse definition is the goal of the first half of chapter 1, and the reader could simply skim it to get a feel for the objects of interest. Nevertheless, we will attempt here to be a bit more intuitive.

A fibre bundle consists of a family of spaces, all isomorphic to each other and called fibres, which is parametrized by a base manifold, such that at each point of the manifold there is one fibre. If the fibres are, for instance, one dimensional vector spaces, then imagining a bundle of wires is an appropriate visualization. Now, if we call the base manifold M, and let the fibres be isomorphic to F, it is *not* correct to think of the bundle as a product $M \times F$, even though at each point the fibre is isomorphic to $\{\text{point}\} \times F$. A fibre bundle may be twisted somehow. When the bundle is globally a product, it is called trivial, and we demand that all fibre bundles be locally trivial.

Because of this twisting of the bundle, there is no natural way of moving along fibres, which in particular means that there is no natural notion of differentiation. A connection is essentially a choice of how to navigate this twist. It defines a covariant derivative on the bundle, a way to differentiate sections. It makes it possible to define a way to move between fibres, called parallel transport. On a principal bundle, a connection can be defined as a choice of horizontal subbundle inside the tangent bundle of the total space. If the bundle is trivial and there are natural choices, then these choices define a canonical connection.

Finally, a gauge transformation is nothing more than an automorphism of the bundle; locally, it can also be seen as a choice of trivialization. These objects are both troublesome and helpful. On the one hand, when we identify all isomorphic bundles, a great deal of redundancy is introduced and solving equations might be hard. On the other hand, looking for gauge invariance becomes a good way of finding interesting objects.

Why gauge theory?

In biology, there is a concept called *convergent evolution*, which is the independent evolution of similar traits in organisms which are not closely related.[1] A similar phenomenon sometimes arises between physics and pure mathematics, whereas two areas develop independently of each other and eventually converge on the same idea; a notable example is general relativity and differential geometry. Where the analogy differs is that when the two interpretations become aware of each other and come into contact, the ensuing interaction often leads to profound results. Gauge theory is a prime example of this wonderful phenomenon.

The fundamentals of the general theory of fibre bundles were developed mainly from 1935 to 1950, and in 1951, N. Steenrod published the first book on the subject, *Topology of Fibre Bundles* [Ste51]. At the same time, the notion of gauge invariance was already well known to physicists in the context of electromagnetism, but it was not very fleshed out and it was not given much importance. Unaware of the existing mathematical theory, in 1954 physicists C.-N. Yang and R. Mills published their acclaimed paper, *Conservation of isotopic spin and isotopic gauge invariance* [YM54], in which they generalize the abelian principle of gauge invariance of eletromagnetism to the non-abelian case of isospin, and in the process "discover" connections and curvature.

The realization of the geometric meaning of gauge fields would only come around 20 years later. Indeed, Yang himself affirmed during an interview that he only learned of the concept of a connection around 1970.[2] In the abstract of a 1975 paper of Yang and T.-T. Wu, they state [WY75]:

> "Generalizations to non-Abelian groups are carried out, and results in identification with the mathematical concept of connections on principal fiber bundles."

Nevertheless, Yang and Mills' paper ushered in a proliferation of gauge theories in physics which culminated in the formulation of the Standard Model of particle physics. The Standard Model is the quantum field theory which describes three of four fundamental forces of the universe, and is one of the greatest achievements of contemporary physics.

The year 1975 also saw the introduction of instantons (then called pseudoparticles), special solutions of the Yang–Mills equations in four dimensions [BPST75]. This finally piqued the interest of mathematicians, and led M. Atiyah and others to publish papers looking at constructions and the moduli space of instantons, [AW77, ADHM78, AHS78]. Already at this point there was plenty of interaction between different areas of mathematics: using R. Penrose's twistor theory, the instanton problem was seen to a problem in complex analysis, and then in algebraic geometry. The main developments of this time are beautifully explained in lecture notes by Atiyah, [Ati79].

As the success and intrinsic interest of Yang–Mills theory in mathematics became apparent, others followed suit. For the next few years, several works appeared studying solutions to the Yang–Mills equations. In particular, K. Uhlenbeck, one of the founders of

[1] Two examples of this are wings evolving in birds and bats, or dorsal fins and flippers evolving in fish and marine mammals.

[2] See the quote at the beginning of [BM94, chapter 3].

modern geometric analysis, because interested in the subject and in 1982 published two works considered fundamental for the analysis of Yang–Mills theory.

As a result of Uhlenbeck's and others' work, S. Donaldson obtained a groundbreaking result in 1983. In his paper [Don83], Donaldson used the moduli space of instantons to prove a theorem on the topology of four-manifolds, which building on work from M. Freedman essentially proved the existence of non-smoothable topological manifolds in dimension four. For their work, Donaldson and Freedman were awarded the Fields Medal in 1986.

From that point onwards, gauge theory became a powerful tool in mathematics, especially for finding topological invariants. Donaldson expanded his work on topological invariants of four-manifolds, founding an area called Donaldson theory, and A. Floer soon came up with instanton invariants for three-manifolds, marking the beginning of Floer theory [Flo88].

As time passed, new equations emerged to coexist with the Yang–Mills equation, via methods such as dimensional reduction or the addition of supersymmetry. As the landscape broadened, so did the applications. One important example is the Seiberg–Witten equation, which was shown to provide easier proofs to many of Donaldson's results, and is used to define certain types of Floer homologies; a very important and somewhat recent application in the context of contact topology[3] is C. Taubes's proof of the Weinstein conjecture [Tau07]. In the same year, A. Kapustin and E. Witten published the Kapustin–Witten equations [KW07], which promise applications to number theory via the geometric Langlands program.

More recently, two of the plenary lectures of the 2018 International Congress of Mathematicians were given by prominent gauge theorists discussing the future of the area. One lecture was given by Donaldson, on recent developments in Kähler geometry and exceptional holonomy. The other lecture, given by P. Kronheimer and T. Mrowka, stresses the utility of gauge theory as a tool, and proposes a proof of the infamous four colour theorem in graph theory using instanton Floer homology for knots.

Gauge theory is a cornerstone of particle physics, and a mathematical theory with far-reaching applications. We hope to have convinced the reader of its ongoing significance. There is still much to be done on both sides, and thus it seems fair to conclude that this is something worth studying.

The main results

With all of this said, we now discuss the contents of this dissertation.

In this work, we studied Uhlenbeck's highly acclaimed paper *Connections with L^p bounds on curvature* [Uhl82]. We mainly followed the exposition on [Weh04]. The main results can be very roughly (and imprecisely) stated as follows.

Theorem (Gauge fixing)**.** *For a local connection with small curvature, there exists a good gauge.*

Theorem (Weak compactness)**.** *A sequence of connections with limited curvature is somehow equivalent to a sequence which converges weakly.*

Theorem (Strong compactness)**.** *A sequence of Yang–Mills connections with limited curvature is somehow equivalent to a sequence which converges uniformly.*

[3]A contact manifold is the analogue of a symplectic manifold in odd dimensions.

Let us now be more precise. Let P be a principal G-bundle over a compact manifold M with (possibly empty) boundary, where G is a compact Lie group (this will **always** be our setup, unless stated otherwise). We are interested in studying connections on P, and the most convenient way is to interpret connections as $\mathfrak{g}$-valued 1-forms on P, and we denote the space of connections by $\mathcal{A}(P)$. The curvature of a connection is a 2-form. The gauge transformations, defined as G-equivariant maps from P to G, act on the connections and their curvatures, and form a group, which we denote $\mathcal{G}(P)$.

Both $\mathcal{A}(P)$ and $\mathcal{G}(P)$ are infinite dimensional objects, with $\mathcal{A}(P)$ being an affine space and $\mathcal{G}(P)$ a Lie group. In order to use tools from functional analysis, we first need to turn them into a Banach space and a Banach manifold. To do so, Sobolev spaces of sections of fibre bundles are introduced, and we define the Sobolev spaces $\mathcal{A}^{k,p}(P)$ of weak connections and $\mathcal{G}^{k,p}(P)$ of weak gauge transformations on P.

We call our "good" gauge the Uhlenbeck gauge condition, which consists of a gauge fixing differential equation (the Coulomb gauge), and estimates on the norm of the connection using the norm of the curvature. Then, the first result is precisely stated as follows.

Theorem (Gauge fixing, cf. definition 2.1 and theorem 2.2)**.** *Suppose that $1 < q \leq p < \infty$ such that $q \geq \frac{n}{2}$, $p > \frac{n}{2}$, and in case $q < n$, $p \leq \frac{nq}{n-q}$. Then there exist constants $\tilde{C}$ and $\tilde{\varepsilon} > 0$ such that the following holds:*

For every point in M, there is a neighbourhood $U \subseteq M$ such that for every connection $A \in \mathcal{A}^{1,p}(U)$ with $\|F_A\|_q^q \leq \tilde{\varepsilon}$ there exists a gauge transformation $u \in \mathcal{G}^{2,p}(U)$ such that $\tilde{A} := u^*A$ is in Uhlenbeck gauge.

The gauge fixing lemma is one of the main ingredients in the proof of the compactness theorems. We can now state the first one precisely.

Theorem (Weak compactness, cf. theorem 3.5)**.** *Let $\frac{n}{2} < p < \infty$. A sequence of connections in $\mathcal{A}^{1,p}(P)$ with uniform L^p-bound on the sequence of curvatures has a subsequence which is gauge equivalent to a weakly convergent sequence, with gauge transformations in $\mathcal{G}^{2,p}(P)$.*

For a connection $A \in \mathcal{A}(P)$, let F_A be its curvature. The Yang–Mills functional is defined as

$$\mathcal{YM}(A) = \int_M |F_A|^2 ,$$

and a connection which is a critical point of this functional is a weak Yang–Mills connection. Finally, our last result is:

Theorem (Strong compactness, cf. theorem 3.9)**.** *Let $1 < p < \infty$ be such that $p > \frac{n}{2}$ and in case $n = 2$, $p \geq \frac{4}{3}$. Suppose a sequence of connections $(A^i)_{i\in\mathbb{N}} \subseteq \mathcal{A}^{1,p}(P)$ is such that the A^i are weak Yang–Mills connections and $\|F_{A^i}\|_p$ is uniformly bounded. Then there exists a subsequence (with same label $i \in \mathbb{N}$) and a sequence of gauge transformations $(u^i)_{i\in\mathbb{N}} \subseteq \mathcal{G}^{2,p}(P)$ such that $u^{i*}A^i$ converges strongly with all derivatives to a smooth Yang–Mills connection.*

This work is organized as follows: In **chapter** 1, we give precise definitions of vector and principal bundles, giving special attention to the equivalent definitions of connections and gauge transformations. After defining Sobolev spaces of sections of fibre bundles and quoting without proof several results from analysis, we redefine the spaces of connections and gauge transformations in this new context and prove several lemmas that show that these objects are well defined and well behaved. We finish the chapter with a discussion on the very important particular case of Yang-Mills connections.

In **chapters** 2 and 3, we state and prove the main results, giving further motivation as well.

There are two **appendices**: appendix A goes into more detail on gauge theory from the physicist's point of view, following a historical thread and in particular showing how to reinterpret eletromagnetism as a $U(1)$ Yang-Mills gauge theory; appendix B contains a collection of results from analysis that are used throughout the text but would involve going off too big of a tangent to introduce and study in detail.

Chapter 1

Gauge theory

The goal of this first chapter is to establish our framework. If gauge theory is the study of connections on principal bundles, then it is vital to have a solid foundation on the theory of fibre bundles. We begin by giving a quick introduction to vector bundles and principal bundles in sections 1.1 and 1.2, paying special attention to the many equivalent definitions of connections and gauge transformations, and culminating in the introduction of the Yang–Mills functional. There are three subsections which are separated from the rest of the text, because they contain a series of important conventions and definitions; in particular, the local formulations are all gathered in a single subsection, as we will be working on local trivializations for most of this text. In section 1.3, we take a slight detour and introduce Sobolev spaces of sections of fibre bundles, in preparation for the analysis of section 1.4. In this final section, we define the Sobolev spaces of connections and gauge transformations, as well as weak (non-smooth) Yang–Mills connections; we prove several lemmas, some of which are very technical, but which show that the objects we defined are well-defined and well-behaved, and provide the foundations for the bigger proofs ahead.

1.1 Vector bundles

In this first short section we define vector bundles and connections. At the end of the section we especialize to the case of Riemannian manifolds in order to introduce notation and objects which will be used extensively throughout the rest of the text. For an introduction to this subject see e.g. [Cra15].

Definition 1.1. A vector bundle of rank r over a manifold M consists of

- a manifold E,
- a surjective map $\pi : E \to M$, and
- for each $x \in M$, a vector space structure on the fibres $E_x := \pi^{-1}(x)$,

satisfying a local triviality condition: around each $x \in M$ there is a neighbourhood U and a diffeomorphism

$$\phi : E|_U := \pi^{-1}(U) \to U \times \mathbb{K}^r$$

sending each fibre E_x isomorphically to $\{x\} \times \mathbb{K}^r$, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

A **section** of the vector bundle is a map $s : M \to E$ such that $\pi \circ s(x) = x$, that is, it takes points on M to the corresponding fibre. It is a smooth section if s is a smooth map between manifolds, and we denote the space of sections of a vector bundle E by $\Gamma(E)$; it is easily seen to be a vector space over $\mathbb{K}$ with the operations defined pointwise, and in fact it is a module over $C^\infty(M)$.

An important notion is that of a trivial vector bundle. The (real or complex) trivial vector bundle of rank r over M is simply $\pi : M \times \mathbb{K}^r \to M$, where π is the projection on the first factor. A vector bundle is said to be **trivializable** if it is isomorphic to the corresponding trivial bundle. A vector bundle morphism [1] between two vector bundles E and F over M is a smooth map $u : E \to F$ such that it preserves the fibres and the restriction $u_x : E_x \to F_x$ is linear; it is an isomorphism when u is a diffeomorphism, or equivalently when the u_x are linear isomorphisms.

Just as manifolds have local coordinate charts which make them more tractable using tools from calculus, vector bundles have frames. A **frame** for a rank r vector bundle is a collection of r sections,

$$e = (e^1, \ldots, e^r)$$

such that for each $x \in M$, $e^1(x), \ldots, e^r(x)$ is a basis for the vector space E_x. Vector bundles may not have globally defined frames, and indeed a choice of frame is equivalent to a trivialization (a choice of isomorphism to the trivial bundle). Nevertheless, local frames always exist: the local triviality condition in the definition is equivalent to the existence of a local frame around every point in M.

New vector bundles can be constructed from previously known ones. Any operation that is natural in the context of vector spaces (e.g., direct sum, dual, tensor product) extends to vector bundles. For instance, if $E \to M$ is a vector bundle, we can define its dual $E^* \to M$ by defining

$$E_x^* = (E_x)^*$$

for all $x \in M$; if $F \to M$ is another vector bundle, we can define the bundle $\mathrm{Hom}(E, F) \to M$ with fibres

$$\mathrm{Hom}(E, F)_x = \mathrm{Hom}(E_x, F_x);$$

similarly, since we will always be working over vector bundles with finite rank, we can define the tensor product as

$$(E \otimes F)_x = \mathrm{Hom}(E_x^*, F_x).$$

Moreover, it is possible to take pullbacks of vector bundles. Given a smooth map $f : M \to N$ and a vector bundle $E \to N$, we can form the pullback bundle $f^*E \to M$ by letting

$$(f^*E)_x = E_{f(x)}.$$

Example 1.2. A classical example of a vector bundle is the tangent bundle $TM \to M$, and its sections are vector fields, $\Gamma(TM) = \mathcal{X}(M)$. A manifold whose tangent bundle is trivializable is called parallelizable. More interestingly, differential forms are sections of a vector bundle, $\Omega^k(M) = \Gamma(\Lambda^k T^* M)$.

[1] One says that morphisms like this "cover the identity", because it is also possible to define morphism between vector bundles over different bases, say $E \to M$ and $F \to N$. In this case a map $f : M \to N$ is also needed, and then $u_x : E_x \to F_{f(x)}$.

Definition 1.3. A **connection** on a vector bundle $E \to M$ is a bilinear map

$$\nabla : \mathcal{X} \times \Gamma(E) \longrightarrow \Gamma(E)$$
$$(X, s) \mapsto \nabla_X(s)$$

satisfying

- $\nabla_{fX} s = f \nabla_X s$,
- $\nabla_X(fs) = f\nabla_X s + X(f)s$ (Leibniz rule)

for all $X \in \mathcal{X}(M)$, $s \in \Gamma(E)$ and $f \in C^\infty(M)$.

On a given trivialization of the vector bundle with frame $e = (e_1, \ldots, e_r)$, the connection is uniquely characterized by a **connection matrix** $A := (A_{ij})$, which is an r-by-r matrix of 1-forms, $A_{ij} \in \Omega^1(M)$,

$$\nabla^A_X(e_j) = \sum_{i=1}^r A_{ij}(X)e_i,$$

and we denote by ∇^A the connection associated to the connection matrix A. Using the Leibniz rule, on a local section

$$s = \sum_{i=1}^r f^i e_i$$

we have

$$\nabla^A_X s(x) = \sum_i \mathrm{d}f^i(X_x)e_i(x) + \sum_{i,j} f^j(x)A_{ij}(X_x)e_i(x). \tag{1.1}$$

Another way to interpret a connection is as a **covariant derivative**,

$$\mathrm{d}_A : \Gamma(E) \longrightarrow \Omega^1(M, E), \quad \mathrm{d}_A(s)(X) := \nabla^A_X(s).$$

From the properties of the connection it is immediate that the covariant derivative is linear and satisfies the Leibniz rule

$$\mathrm{d}_A(fs) = f\,\mathrm{d}_A s + \mathrm{d}f \otimes s.$$

Moreover, from (1.1) we see that on a local section s defined in terms of a local frame e as above,

$$\mathrm{d}_A s = \sum_i \mathrm{d}f^i e_i + \sum_{i,j} f^j A_{ij} e_i,$$

which leads to the frequently used notation

$$\mathrm{d}_A = \mathrm{d} + A$$

for the local representation of the covariant derivative.

There are two usual ways to extend the covariant derivative from sections to more general k-forms on the vector bundle, and we define two operators

$$\nabla^A,\ \mathrm{d}_A : \Omega^k(M, E) \to \Omega^{k+1}(M, E)$$

as follows: for $X_0, \dots, X_k \in TM$ and $\omega \in \Omega^k(M, E)$,

$$\nabla^A \omega\,(X_0, \dots, X_k) := \nabla^A_{X_0}(\omega(X_1, \dots, X_k)) - \omega(\nabla^M_{X_0} X_1, \dots, X_k) - \dots - \omega(X_1, \dots, \nabla^M_{X_0} X_k), \tag{1.2}$$

and

$$\mathrm{d}_A \omega\,(X_0, \dots, X_k) := \sum_{i=0}^{k} (-1)^i \nabla^A_{X_i}(\omega(X_0, \dots, \hat{X}_i, \dots, X_k)) + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_k), \tag{1.3}$$

where ∇^M is the Levi-Civita connection on $TM \to M$. [2] These operators are related to one another by

$$\mathrm{d}_A \omega\,(X_0, \dots, X_k) = \sum_{i=0}^{k} (-1)^i\, \nabla^A \omega\,(X_i, X_0, \dots, \hat{X}_i, \dots, X_k), \tag{1.4}$$

and as expected when $k = 0$ (that is, in sections of the bundle), $\mathrm{d}_A = \nabla^A$, so that the notation $\mathrm{d}_A s(X) = \nabla^A_X s$ is still consistent.

Just as we can define new bundles from old using operations which are natural to vector spaces and pullbacks, the same can be done with connections. Let $E \to M$ be a vector bundle, and let $f : N \to M$ be a smooth map. Given a connection ∇ on E we can define a pullback connection $f^*\nabla$ on f^*E as follows: for $s \in \Gamma(E)$, $x \in N$ and $X \in TN$,

$$(f^*\nabla)_X (f^*s)(x) = \nabla_{\mathrm{d}_x f(X)} s(f(x)).$$

This can be used to define a further useful property of a connection: it defines a way to move from one fibre E_x to another along paths on M. Given a path on the base, say $\gamma : I \to M$ for some interval $I \subseteq \mathbb{R}$, we can define paths on E above γ as a section $u \in \Gamma(\gamma^*E)$, and then

$$\begin{aligned} u : I &\longrightarrow E, \\ u(t) &\in E_{\gamma(t)}. \end{aligned}$$

Such a path u is said to be parallel to γ if

$$(\gamma^*\nabla)_{\frac{\mathrm{d}}{\mathrm{d}t}} u = 0.$$

Given γ, ∇ and an initial point $s \in E_{\gamma(t_0)}$ for $t_0 \in I$, there is a unique path u_s such that $u_s(t_0) = s$ and u_s is parallel to γ. A collection of such paths taking each point in $E_{\gamma(t_0)}$ to $E_{\gamma(t_1)}$ for $I = [t_0, t_1]$ is called **parallel transport**,

$$\begin{aligned} P^{t_1, t_0}_{\gamma} : E_{\gamma(t_0)} &\to E_{\gamma(t_1)} \\ s &\mapsto u_s(t_1), \end{aligned} \tag{1.5}$$

and this is a linear isomorphism of the fibres it connects.

[2]See the discussion below on Riemannian manifolds for the definition of the Levi-Civita connection.

One last thing that needs to be discussed is vector bundle **metric**. A metric on a vector bundle $E \to M$ is a family

$$h = \{h_x\}_{x \in M}$$

of inner products on the vector spaces E_x, which vary smoothly on M in the sense that, for two sections $s, s' \in \Gamma(E)$,

$$h(s, s')(x) = h_x(s(x), s'(x))$$

is smooth as a function $M \to \mathbb{K}$. Given a vector bundle $E \to M$ with metric h, a connection on E is said to be compatible with h if, for instance,

$$Xh(s, s') = h(\nabla_X s, s') + h(s, \nabla_X s')$$

holds for every $X \in \mathcal{X}(M)$ and $s, s' \in \Gamma(E)$.

Proposition 1.4 ([Cra15], proposition 1.41)**.** *Every vector bundle admits a metric, and for every metric there always exists a compatible connection.*

1.1.1 Useful notions on Riemannian manifolds

Before moving on from vector bundles, let us discuss a bit about the very important and special case of the tangent bundle. A **Riemannian metric** on a manifold M is simply a metric on the tangent bundle, therefore by the previous proposition we know that every smooth manifold admits a Riemannian metric. We will denote a Riemannian metric by g.

When $E = TM$, the sections of the vector bundle are also vector fields, and so a connection is an operator

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M),$$

which makes it possibly to talk about torsion,

$$T_\nabla(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

A connection is said to be torsion-free if $T_\nabla = 0$. Because of the existence of torsion, in this case it is possible to single out a canonical connection:

Proposition 1.5 ([Cra15], theorem 1.43)**.** *On a Riemannian manifold there exists a unique connection compatible with the metric and torsion-free. It is called the* ***Levi-Civita*** *connection.*

It will be useful to define the **Christoffel symbols** for the Levi-Civita connection. On $TM \to M$ it is natural to use a local frame

$$\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$$

induced by a local coordinate chart $(U, x_1, \dots, x_n)$. Let $\partial_i := \frac{\partial}{\partial x_i}$ and $\nabla_i := \nabla_{\partial_i}$, then the Christoffel symbols Γ^k_{ij} are defined as [3]

$$\nabla_i \partial_j = \Gamma^k_{ij} \partial_k.$$

The fact that ∇ is torsion-free is reflected in the symmetry of the symbols,

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

[3] We use the Einstein summation convention for repeated indices.

Moreover, representing the metric g locally by a matrix

$$g_{ij} := g(\partial_i, \partial_j),$$

and letting (g^{ij}) be its inverse, the compatibility of ∇ with the metric becomes

$$\Gamma_{ij}^k = \frac{1}{2}\sum_l \left(\frac{\partial g_{jl}}{\partial x_i} + \frac{\partial g_{il}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_l}\right) g^{lk}.$$

The Christoffel symbols are also useful to show the dependence of the covariant derivative on the metric on the base manifold. For instance, we can write locally for $A \in \Omega^1(M, E)$:

$$(\nabla A)_{ij} := \nabla A(\partial_i, \partial_j) = \nabla_i^E (A(\partial_j)) - A(\nabla_j^M \partial_i) = \nabla_i^E A_j - \Gamma_{ij}^k A_k.$$

Besides a canonical choice of connection, the metric on M also gives natural definitions to operations such as integration and inner products, and then a further choice of vector bundle metric extends these notions to bundle valued objects. We give these definitions now.

First, the metric defines a **volume form** $\mathrm{d}vol$ and then for $f \in C_0(M)$ we can define

$$\int_M f := \int_M f \,\mathrm{d}vol. \tag{1.6}$$

Locally, the volume form is given by $\mathrm{d}vol = \sqrt{|\det g|}\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n$. For simplicity we are assuming that M is oriented in order to have a globally defined volume form, however that is not necessary for defining integration, see e.g. [Aub82, chapter 1, sections 9 and 11].

Moreover, the metric induces a pointwise inner product on tensors $\alpha, \beta \in \otimes^k T_x^* M$, which can be written as

$$\langle \alpha, \beta \rangle = g^{i_1 j_1} \cdots g^{i_k j_k} \alpha_{i_1 \cdots i_k} \beta_{j_1 \cdots j_k}.$$

This is independent of the local coordinates. Note that a Riemannian metric is itself a tensor, $g \in \Gamma(\otimes^2 T^* M)$.

The previous definitions can now be used to define the **Hodge star**, $* : \Omega^k(M) \to \Omega^{n-k}(M)$, as the only map that satisfies

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \,\mathrm{d}vol, \quad \forall \alpha, \beta \in \Omega^k(M).$$

We can also define an inner product on the space of k-forms: for $\alpha, \beta \in \Omega^k(M)$,

$$\langle \alpha, \beta \rangle := \int_M \langle \alpha, \beta \rangle \,\mathrm{d}vol = \int_M \alpha \wedge *\beta.$$

We will sometimes denote $\mathrm{d}vol =: *1$. Note that $*$ is its own inverse up to a sign: for $\alpha \in \Omega^k(M)$, $*^2\alpha = (-1)^{k(n-k)}\alpha$. Occasionally, when the sign itself is not important, we will simply use $*^2 = \pm 1$.

The Hodge star is also used to define the **codifferential**,

$$\begin{aligned} \mathrm{d}^* : \Omega^k(M) &\longrightarrow \Omega^{k-1}(M) \\ \alpha &\mapsto -(-1)^{n(k-1)} * \mathrm{d} * \alpha. \end{aligned}$$

The operator d^* is also called the **formal adjoint** to the exterior derivative because of the following: for $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^{k+1}(M)$,

$$\begin{aligned}\int_M \mathrm{d}(\alpha \wedge *\beta) &= \int_M \mathrm{d}\alpha \wedge *\beta + (-1)^k \alpha \wedge \mathrm{d} * \beta \\ &= \int_M \mathrm{d}\alpha \wedge *\beta + (-1)^k (-1)^{(n-k+1)(k-1)} \alpha \wedge (**) \,\mathrm{d} * \beta \\ &= \int_M \mathrm{d}\alpha \wedge *\beta + (-1)^{n(k-1)} \alpha \wedge *(* \,\mathrm{d} *)\beta \\ &= \int_M \mathrm{d}\alpha \wedge *\beta + -\alpha \wedge * \,\mathrm{d}^*\beta \\ &= \langle \mathrm{d}\alpha, \beta \rangle - \langle \alpha, \mathrm{d}^*\beta \rangle.\end{aligned}$$

Then by Stokes's theorem, if either M has no boundary or one of the forms vanishes on ∂M,

$$\langle \mathrm{d}\alpha, \beta \rangle = \langle \alpha, \mathrm{d}^*\beta \rangle.$$

Finally, it is possible to extend these notions to vector bundle valued differential forms. Let $E \to M$ be such a vector bundle, with bundle metric h. Given a local frame e on a trivializing neighbourhood $U \subseteq M$, for $\alpha \in \Omega^k(M, E)$ and $\beta \in \Omega^l(M, E)$ we can write $\alpha = \alpha^i \otimes e_i$ and $\beta = \beta^i \otimes e_i$, for $\alpha^i, \beta^i \in \Omega^\bullet(M)$. Then h induces a pairing

$$\langle \cdot \wedge \cdot \rangle : \Omega^k(M, E) \times \Omega^l(M, E) \to \Omega^{k+l}(M),$$

which is given locally by

$$\langle \alpha \wedge \beta \rangle := (\alpha^i \wedge \beta^j)\, h(e_i, e_j).$$

There is also a natural extension for the Hodge star, given locally by

$$*\alpha := (*\alpha^i) \otimes e_i.$$

Note that now for any $\alpha, \beta \in \Omega^k(M, E)$ we can associate a top-form on M given by $\langle \alpha \wedge *\beta \rangle$, and thus we define an inner product on $\Omega^k(M, E)$,

$$\langle \alpha, \beta \rangle = \int_M \langle \alpha \wedge *\beta \rangle.$$

For the purposes of later use, we will actually refer to

$$\langle \alpha, \beta \rangle := *\langle \alpha \wedge *\beta \rangle \in \Omega^n(M) \tag{1.7}$$

as the **(pointwise) inner product on** $\Omega^k(M, E)$.

Moreover, if A is a connection on the vector bundle $E \to M$, we may also define the formal adjoint of the covariant derivative d_A,

$$\mathrm{d}_A^* : \Omega^k(M, E) \to \Omega^{k-1}(M, E),$$

$$\mathrm{d}_A^* = -(-1)^{n(k-1)} * \mathrm{d}_A * .$$

1.2 Principal bundles

Principal bundles are the underlying objects in gauge theory. Our main objects of study, connections and gauge transformations, will be defined in this context. There is much more that can be said about principal bundles and their relationship to vector bundles,[4] but our focus in this section will be on the many ways the bundle, connections and gauge transformations can be defined. For organizational purposes, and because we will refer back to them often, all of the local results are collected in a separate subsection.

Definition 1.6. A principal G-bundle is a manifold P along with

- a G action on P, $P \times G \to P$, $(p, g) \mapsto pg$,
- a surjective map $\pi : P \to M$ that is G-invariant, i.e. $\pi(pg) = \pi(p)$ for all $p \in P$ and $g \in G$,

and such that a local triviality condition is satisfied: for all $x \in M$, there exists a neighbourhood U and a diffeomorphism

$$\begin{aligned}\Phi : \pi^{-1}(U) &\longrightarrow U \times G \\ p &\mapsto (\pi(p), \phi(p))\end{aligned}$$

taking a fibre $\pi^{-1}(x)$ to $\{x\} \times G$ and which is G-equivariant, that is $\phi(pg) = \phi(p)g$.

Equivalently, $\pi : P \to M$ is a principal G-bundle if π is a submersion and there is a free and proper G-action on P which is fibre preserving. The action is also transitive on the fibres, that is, for any $p, q \in \pi^{-1}(x)$ there exists $g \in G$ such that $q = pg$. Unlike vector bundles, principal bundles do not generally have global sections; indeed, a global section is equivalent to a trivialization of the bundle, $P \simeq M \times G$.

A more direct point of view is to use a **bundle atlas** $(U_\alpha, \Phi_\alpha = (\pi, \phi_\alpha))_{\alpha \in A}$ to write any $p \in P$ as

$$p = \lfloor \alpha, x, g \rceil$$

for $\alpha \in A$, $x = \pi(p)$ and $g = \phi_\alpha(p) \in G$. This is an equivalence class, and we will want that on non-empty intersections $U_\alpha \cap U_\beta$

$$\lfloor \alpha, x, \phi_\alpha(p) \rceil = \lfloor \beta, x, \phi_\beta(p) \rceil.$$

We define so called **transition functions** $\phi_{\alpha\beta} : U_\alpha \cap U_\beta \to G$ as

$$\phi_{\alpha\beta}(x) = \phi_\alpha(p)\phi_\beta(p)^{-1}$$

for any $p \in \pi^{-1}(x)$; this is well defined, as for any $q \in \pi^{-1}(x)$ there exists $g \in G$ such that $q = pg$ and

$$\phi_\alpha(pg)\phi_\beta^{-1}(pg) = \phi_\alpha(p)g(\phi_\beta(p)g)^{-1} = \phi_\alpha(p)\phi_\beta(p)^{-1}.$$

Then the equivalence relation that defines $\lfloor \cdot, \cdot, \cdot \rceil$ will be

$$(\alpha, x, g) \sim (\beta, x, h) \iff h = \phi_{\beta\alpha}(x)g.$$

[4] Frame bundles and G-structures, for instance. See [Cra15].

The G-action is defined naturally as

$$pg = \lfloor \alpha, \pi(p), \phi_\alpha(p)g \rceil.$$

Note that these functions satisfy **cocycle conditions**, on $U_\alpha \cap U_\beta \cap U_\gamma$

$$\phi_{\alpha\gamma}(x) = \phi_{\alpha\beta}(x)\phi_{\beta\gamma}(x), \tag{1.8}$$

and also $\phi_{\alpha\alpha} = \mathbb{1}$ and $\phi_{\alpha\beta}(x)^{-1} = \phi_{\beta\alpha}(x)$.

The open cover of M and the transition functions defined on the intersections encode the whole principal bundle. Indeed, given such a cover $M = \bigcup_{\alpha\in A} U_\alpha$ and G-valued transition functions $\{\phi_{\alpha\beta}\}_{\alpha,\beta\in A}$, the equivalence relation $(\alpha, x, g) \sim (\beta, x, \phi_{\beta\alpha}(x)g)$ gives rise to a principal G-bundle

$$\pi : \{\lfloor \alpha, x, g \rceil : \alpha \in A, x \in M, g \in G\} \to M.$$

If these transition functions originated from a pre-existing bundle P, then this procedure reconstructs P. More generally, it is possible to check if two principal G-bundles are isomorphic by observing their transition functions on the same open cover of the base bundle. A G-bundle isomorphism is a bundle isomorphism [5] that also preserves the group action.

Lemma 1.7. *Let $M = \bigcup_{\alpha\in A} U_\alpha$ be an open cover of M and let $\{\phi_{\alpha\beta}\}$, $\{\psi_{\alpha\beta}\}$, $\alpha, \beta \in A$ be two sets of transition functions for two principal G-bundles over M. Then these bundles are isomorphic if, and only if, there exist a cover $M = \bigcup_{\alpha\in A} V_\alpha$ with $V_\alpha \subseteq U_\alpha$ and local functions $g_\alpha : V_\alpha \to G$ such that $\psi_{\alpha\beta}(x) = g_\alpha(x)\phi_{\alpha\beta}(x)g_\beta^{-1}(x)$.*

Isomorphic bundles will usually be identified, and so an object of great importance is the group of G-**bundle automorphisms** of a principal bundle P,

$$\mathrm{Aut}(P) := \{\psi : P \to P : \pi \circ \psi = \psi \text{ and } \psi(pg) = \psi(p)g\}.$$

To each $\psi \in \mathrm{Aut}(P)$ we will associate a map called a gauge transformation. To properly discuss these transformations and connections on principal bundles, we need to look at associated bundles.

Let F be some other manifold with a representation $\rho : G \to \mathrm{Diff}(F)$ which gives a G-action on it. Then we may define an associated bundle to a principal G-bundle P as the set of equivalence classes $\lfloor p, f \rceil = \lfloor pg, \rho(g^{-1})f \rceil$ for all $g \in G$, and we denote it $E(P, F) = (P \times F)/G$. If $F = V$ is a vector space, then $\rho : G \to GL(V)$ is a representation of G in the usual sense, and $E(P, V)$ is a vector bundle. We will now define and give properties of two bundles associated to P which will be especially important.

We denote by $\tilde{\pi} : \mathrm{Ad}(P) \to M$ the **associated bundle with fibre** G

$$\mathrm{Ad}(P) := \frac{P \times G}{G},$$

where G acts on itself by conjugation, $(g, h) \mapsto hgh^{-1}$. Thus we have the fibres

$$\tilde{\pi}^{-1}(x) = \mathrm{Ad}(P)_x = \{\lfloor p, g \rceil : g \in G\},$$

[5] A diffeomorphism of the total spaces which preserves the fibres.

for $p \in \pi^{-1}(x)$, and note that $\lfloor pg, h \rceil = \lfloor p, ghg^{-1} \rceil$. We can identify the sections of $\operatorname{Ad}(P)$ with the G-equivariant functions,

$$C^\infty(P,G)^G := \left\{ u : P \to G : u(pg) = g^{-1}u(p)g = \operatorname{Ad}_{g^{-1}} u(p) \right\},$$

in the following way:

- a section $s \in \Gamma \operatorname{Ad}(P)$ induces a function u by $s(\pi(p)) = \lfloor p, u(p) \rceil$, and u is equivariant because
$$\lfloor p, u(p) \rceil = \lfloor pg, u(pg) \rceil = \lfloor p, \operatorname{Ad}_g u(pg) \rceil$$
implies $u(p) = gu(pg)g^{-1}$;
- a function $u \in C^\infty(P,G)^G$ induces a section $s(x) = \lfloor p, u(p) \rceil$, and this does not depend on the choice of $p \in \pi^{-1}(x)$, as for any other $q \in \pi^{-1}(x)$ we write $q = pg$ for some $g \in G$, and so
$$\lfloor q, u(q) \rceil = \lfloor pg, u(pg) \rceil = \lfloor pg, \operatorname{Ad}_{g^{-1}} u(p) \rceil = \lfloor p, u(p) \rceil.$$

Furthermore, the sections of $\operatorname{Ad}(P)$ form a group under pointwise multiplication,

$$(s \cdot s')(x) := \lfloor p, u(p)u'(p) \rceil.$$

The G-equivariant functions can further be identified with the G-bundle automorphisms of P. This identification can be written as follows:

- A function $u \in C^\infty(P,G)^G$ induces an automorphism $\psi : P \to P$, $\psi(p) = pu(p)$ which is clearly fibre preserving, and is equivariant:
$$\psi(pg) = pgu(pg) = pu(p)g = \psi(p)g.$$
- Conversely, there is a unique map $u : P \to G$ such that $\psi(p) = pu(p)$ for a given $\psi \in \operatorname{Aut}(P)$. The equivariance of ψ gives
$$\psi(pg) = p(gu(pg)) = p(u(p)g) = \psi(p)g,$$
and because the action of G is free, $gu(pg) = u(p)g$, which establishes $u \in C^\infty(P,G)^G$.

This identification also introduces a group multiplication in $\operatorname{Aut}(P)$ via composition of automorphisms,

$$\psi_1 \circ \psi_2(p) = p \cdot u_2(p)u_1(p).$$

Thus we have the group isomorphisms

$$\Gamma \operatorname{Ad}(P) \simeq C^\infty(P,G)^G \simeq \operatorname{Aut}(P).$$

Definition 1.8 (Gauge transformation)**.** We will call $u \in C^\infty(P,G)^G$ a gauge transformation, and denote the group of gauge transformations $\mathcal{G}(P)$.

We will have more to say about gauge transformations, their action and local representation, further into the chapter.

Let $\mathfrak{g}$ be the Lie algebra of G, and let G act on $\mathfrak{g}$ via the adjoint action

$$\begin{aligned}\mathrm{Ad} : G &\to \mathrm{End}(\mathfrak{g}) \\ \xi &\mapsto \mathrm{Ad}_g(\xi) = g\xi g^{-1}.\end{aligned}$$

Then $\mathrm{ad}(P) \to M$ is the **associated vector bundle with standard fibre** $\mathfrak{g}$,

$$\mathrm{ad}(P) := \frac{P \times \mathfrak{g}}{G}.$$

Just as with $\mathrm{Ad}(P)$, we can identify the sections of $\mathrm{ad}(P)$ with G-equivariant functions $P \to \mathfrak{g}$,

$$C^\infty(P, \mathfrak{g})^G := \left\{ f : P \to \mathfrak{g} : f(pg) = \mathrm{Ad}_{g^{-1}} f(p) \right\},$$

such that any section can be written $s(x) = \lfloor p, f(p) \rceil$, for any $p \in \pi^{-1}(x)$. The space of sections $\Gamma\,\mathrm{ad}(P)$ has a natural Lie algebra structure induced by the bracket on $\mathfrak{g}$,

$$[s, s'](x) = \lfloor p, [f(p), f'(p)] \rceil.$$

The importance of this associated bundle will become clear after the next definition.

There is a canonical vertical subbundle $T^V P \subseteq TP$ given by $T^V P = \ker \mathrm{d}\pi$, composed of vectors tangent to the fibres $P_x \simeq G$. Each vertical tangent space is isomorphic to the Lie algebra,

$$T_p^V P = \ker \mathrm{d}_p\pi \simeq \mathfrak{g},$$

and the tangent vectors will be denoted $p\xi$, as defined in subsection 1.2.2. Every complement of these vertical spaces is isomorphic to $Im(\mathrm{d}_p\pi) = T_{\pi(p)}M$, but there is in general no canonical choice for these horizontal spaces; a connection on a principal bundle represents precisely a choice of equivariant horizontal distribution H such that $TP = T^V P \oplus H$.

Definition 1.9 (Connection)**.** A connection on P is a 1-form $A \in \Omega^1(P, \mathfrak{g})$ satisfying

- $A_{pg}(vg) = g^{-1}A_p(v)g = \mathrm{Ad}_{g^{-1}} A$, for all $g \in G$ and $v \in T_pP$,
- $A_p(p\xi) = \xi$, for all $\xi \in \mathfrak{g}$,

that is, it is G-equivariant and takes fixed values on vertical tangent vectors. We denote the set of smooth connections $\mathcal{A}(P)$.

To see the relation between both notions, observe that we can write

$$TP = T^V P \oplus \ker A.$$

Note that $\Omega^1(P, \mathfrak{g})$ means that the one-forms take values in $\mathfrak{g}$, which is a finite dimensional vector space. In the language of vector bundles, $\Omega^1(P, \mathfrak{g}) = \Gamma(\Lambda T^*P \otimes P \times \mathfrak{g})$, where $P \times \mathfrak{g} \to P$ is a trivializable vector bundle of rank $\dim \mathfrak{g}$. From either point of view it is clear that many global results on forms will readily generalize to connections; for instance, it makes sense to write $\mathrm{d}A$ in this case, when in general connections are needed to define covariant derivatives.

It is easy to see from the definition that the difference of two connections is a **basic form**: G-equivariant and horizontal [6]. On the other hand, it is known that the space of basic k-forms $\Omega^k_{bas}(P, \mathfrak{g})$ is isomorphic to $\Omega^k(M, \mathrm{ad}(P))$: for $\tau \in \Omega^k_{bas}(P, \mathfrak{g})$, the corresponding $\tilde{\tau} \in \Omega^k(M, \mathrm{ad}(P))$ is uniquely defined by

$$\lfloor p, \tau_p(X_1, \ldots, X_k)\rceil = \tilde{\tau}_{\pi(p)}(\,\mathrm{d}_p\pi(X_1), \ldots, \,\mathrm{d}_p\pi(X_k)) \;\in \mathrm{ad}(P)_{\pi(p)} \tag{1.9}$$

for any $X_1, \ldots, X_k \in T_pP$. Thus, the space of connections is an affine space, and fixing a reference connection $\tilde{A} \in \mathcal{A}(P)$ we can write

$$\mathcal{A}(P) = \tilde{A} + \Omega^1(M, \mathrm{ad}(P)).$$

A similar isomorphism exists for general associated vector bundles $E(P, V)$, and it allows a connection on the principal bundle to induce one on the associated bundle. Suppose $\rho : G \to GL(V)$ is a representation and $E(P, V)$ is the corresponding bundle. There is an isomorphism $h : \Omega^k(M, E(P, V)) \xrightarrow{\sim} \Omega^k_{bas}(P, V)$, and furthermore we can differentiate the representation, $\mathrm{d}_{\mathbb{1}}\rho : \mathfrak{g} \to \mathrm{End}(V)$. For $A \in \mathcal{A}(P)$ we can then define

$$\begin{aligned} \mathrm{d}_A : \Omega^k_{bas}(P, V) &\longrightarrow \Omega^{k+1}_{bas}(P, V) \\ \omega &\mapsto \mathrm{d}f + \mathrm{d}_{\mathbb{1}}\rho(A)(\omega), \end{aligned}$$

and this in turn will induce a connection on $E(P, V)$,

$$\nabla^A := h^{-1} \circ \mathrm{d}_A \circ h : \Omega^k(M, E(P, V)) \to \Omega^{k+1}(M, E(P, V)).$$

In the particular case of $\mathrm{ad}(P)$, which is the associated vector bundle $E(P, \mathfrak{g})$ with $\rho = \mathrm{Ad} : G \to GL(\mathfrak{g})$, we have $\mathrm{d}_{\mathbb{1}}\rho(\xi)\eta = \mathrm{ad}(\xi)\eta = [\xi, \eta]$ and so a connection $A \in \mathcal{A}(P)$ defines

$$\begin{aligned} \mathrm{d}_A : \Omega^k_{bas}(P, \mathfrak{g}) &\longrightarrow \Omega^{k+1}_{bas}(P, \mathfrak{g}) \\ \tau &\mapsto \mathrm{d}\tau + [A \wedge \tau], \end{aligned} \tag{1.10}$$

where $[\cdot \wedge \cdot]$ is the wedge product of two forms with the Lie bracket used to combine the values in $\mathfrak{g}$.[7] This will then induce on $\mathrm{ad}(P)$ a covariant derivative

$$\nabla^A : \Gamma(\mathrm{ad}(P)) \longrightarrow \Omega^1(M, \mathrm{ad}(P))$$

which can be written explicitly in the following way: for $s \in C^\infty(P, \mathfrak{g})^G \simeq \Gamma(\mathrm{ad}(P))$, $X \in T_xM$ and $Y \in T_pP$ such that $\pi(p) = x$ and $\mathrm{d}_p\pi(Y) = X$,

$$\nabla^A s(X) = \lfloor p, \,\mathrm{d}_p s(Y) + [A(Y), s(p)]\rceil \in \mathrm{ad}(P)_x. \tag{1.11}$$

This, of course, extends to $\nabla^A : \Omega^k(M, \mathrm{ad}(P)) \to \Omega^{k+1}(M, \mathrm{ad}(P))$ exactly as in (1.2).

Now, while it is true that for the usual (de Rham) exterior derivative we have $\mathrm{d}^2 = 0$, this will not hold in general for d_A defined above. This failure can be measured by the **curvature** of the connection,

$$F_A = \,\mathrm{d}A + \frac{1}{2}[A \wedge A] \in \Omega^2_{bas}(P, \mathfrak{g}), \tag{1.12}$$

[6] It kills vertical tangent vectors.

[7] For example, for $\alpha, \beta \in \Omega^1_{bas}(P, \mathfrak{g})$ and $X, Y \in T_pP$, $[\alpha \wedge \beta](X, Y) = [\alpha(X), \beta(Y)] - [\alpha(Y), \beta(X)]$, and then note $[\alpha \wedge \alpha] = 2[\alpha, \alpha]$.

and we obtain $\mathrm{d}_A \mathrm{d}_A \tau = [F_A \wedge \tau]$ for all $\tau \in \Omega^k_{bas}(P, \mathfrak{g})$. A connection for which $F_A = 0 = \mathrm{d}_A^2$ is called a **flat connection**. The curvature satisfies the Bianchi identity,

$$\mathrm{d}_A F_A = 0.$$

Moreover, it can be seen as a differential form in $\Omega^2(M, \mathrm{ad}(P))$.

We now calculate the effect of a gauge transformation on a connection and its curvature. For that, define the Maurer-Cartan form $\theta_g = \mathrm{d}_g L_{g^{-1}} : T_g G \to \mathfrak{g}$, for $L_g(h) = gh$.

Lemma 1.10 (The gauge action). *Let $u \in \mathcal{G}(P)$ and $\psi(p) = pu(p)$ its corresponding automorphism. Then the action of a gauge transformation on a connection is defined as $u^*A := \psi^*A$ and can be written*

$$u^*A = u^{-1}Au + u^*\theta \tag{1.13}$$

$$= u^{-1}Au + u^{-1}\,\mathrm{d}u. \tag{1.14}$$

Proof. Define the multiplication σ on $P \times G$ as $\sigma(p, g) = pg = \tilde{R}_g(p) = \tilde{L}_p(g)$ and write $\psi(p) = \sigma(Id, u)(p)$. Then for $v \in T_pP$,

$$\begin{aligned} \mathrm{d}_p\psi(v) &= \mathrm{d}_{(p,u(p))}\sigma \circ \mathrm{d}_p(Id, u)(v) \\ &= \mathrm{d}_{(p,u(p))}\sigma(v,\, \mathrm{d}_p u(v)) \\ &\overset{(*)}{=} \mathrm{d}_p\tilde{R}_{u(p)}(v) + \mathrm{d}_{u(p)}\tilde{L}_p(\mathrm{d}_p u(v)), \end{aligned}$$

where $(*)$ follows because $T_{(p,g)}(P \times G) \simeq T_pP \oplus T_gG$ and using curves one can easily show that $\mathrm{d}_{(p,g)}\sigma(v, 0) = \mathrm{d}_p\tilde{R}_g(v)$ and $\mathrm{d}_{(p,g)}\sigma(0, X) = \mathrm{d}_g\tilde{L}_p(X)$. Since $\tilde{L}_p(h) = ph = pgg^{-1}h$, we may write $\tilde{L}_p = \tilde{L}_{pg} \circ L_{g^{-1}}$, so that

$$\mathrm{d}_{u(p)}\tilde{L}_p(\mathrm{d}_p u(v)) = \mathrm{d}_{\mathbb{1}}\tilde{L}_{pu(p)}\,\mathrm{d}_{u(p)}L_{u(p)^{-1}}(\mathrm{d}_p u(v)) = \mathrm{d}_{\mathbb{1}}\tilde{L}_{pu(p)}\theta_{u(p)}(\mathrm{d}_p u(v))$$

For $p\xi = \mathrm{d}_{\mathbb{1}}\tilde{L}_p(\xi)$ and $vg = \mathrm{d}_p\tilde{R}_g(v)$ as defined in subsection 1.2.2,

$$\mathrm{d}_p\psi(v) = vu(p) + pu(p)\theta_{u(p)}(\mathrm{d}_p u(v)).$$

Now, using this expression and calculating using the properties of the connection,

$$\begin{aligned} \psi^*A(v) &= A_{\psi(p)}(\mathrm{d}_p\psi(v)) \\ &= A_{pu(p)}(vu(p) + pu(p)\underbrace{\theta_{u(p)}(\mathrm{d}_p u(v))}_{\in\mathfrak{g}}) \\ &= u(p)^{-1}A_p(v)u(p) + \theta_{u(p)}(\mathrm{d}_p u(v)), \end{aligned}$$

and rewriting $\theta_{u(p)}(\mathrm{d}_p u(v))$ as $(u^*\theta)_p(v)$ or $\mathrm{d}_{u(p)}L_{u(p)^{-1}}(\mathrm{d}_p u(v)) = u(p)^{-1}\,\mathrm{d}_p u(v)$ from $gv_h = \mathrm{d}_h L_g(v)$, we have the expressions we wanted for the gauge action. □

From this, it is also possible to show the effect of a gauge transformation on the curvature,

$$F_{u^*A} = u^*F_A = u^{-1}F_A u. \tag{1.15}$$

Finally, we define a pointwise inner product on $\Omega^k(M, \mathrm{ad}(P))$ as in (1.7). We assume the Riemannian metric on M is given, and the metric on $\mathrm{ad}(P)$ is induced by the Ad-invariant inner product on $\mathfrak{g}$ defined in subsection 1.2.2,

$$\langle \lfloor p, \xi \rceil, \lfloor p, \zeta \rceil \rangle_{\mathrm{ad}(P)} := \langle \xi, \zeta \rangle_{\mathfrak{g}}.$$

Now, viewing the curvature as being in $\Omega^2(M, \mathrm{ad}(P))$ and writing $|\cdot| = \sqrt{\langle \cdot\,, \cdot \rangle} : M \to \mathbb{R}$, we can define the **Yang–Mills functional**

$$\mathcal{YM}(A) := \int_M |F_A|^2 \,\mathrm{d}vol = \int_M *\langle F_A \wedge *F_A \rangle. \tag{1.16}$$

Due to the gauge action on F_A and the invariance of the metric on $\mathrm{ad}(P)$, this functional is invariant under gauge transformations,

$$\mathcal{YM}(u^*A) = \mathcal{YM}(A), \quad \forall u \in \mathcal{G}(P).$$

We will want to study its extrema. For that, note that because $\mathcal{A}(P)$ is an affine space with vector space $\Omega^1(M, \mathrm{ad}(P))$, it is enough to check the variation of the curvature along lines $A + t\beta$, for $\beta \in \Omega^1(M, \mathrm{ad}(P))$. Now,

$$F_{A+t\beta} = F_A + t\,\mathrm{d}_A\beta + \frac{1}{2}t^2[\beta \wedge \beta],$$

and thus momentarily writing $\langle \cdot, \cdot \rangle$ for the integrated (not pointwise) inner product it is straightforward to calculate

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathcal{YM}(A + t\beta) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle F_{A+t\beta}, F_{A+t\beta} \rangle \\ &= 2\langle F_A,\, \mathrm{d}_A\beta \rangle,\end{aligned}$$

and so an extremum of the $\mathcal{YM}$ functional has to satisfy the **weak Yang–Mills equation**,

$$\int_M \langle F_A,\, \mathrm{d}_A\beta \rangle = 0, \quad \forall \beta \in \Omega^1(M, \mathrm{ad}(P)).$$

If everything is smooth, this is equivalent to the **(strong) Yang–Mills equation**,

$$\begin{cases} \mathrm{d}_A^* F_A = 0 & \text{on } M, \\ *F_A|_{\partial M} = 0 & \text{on } \partial M, \end{cases}$$

where in the case of a manifold without boundary this is just the usual Yang–Mills equation. This will be proved further ahead in a more general context as lemma 1.37. Just as the functional, these equations are invariant under gauge transformations, and because of this it is said that the solutions to the Yang–Mills equations come in gauge orbits.

1.2.1 Local formulations and results

In this subsection we will derive local representations and formulas for the objects we have defined, as these will be the forms which we will mainly use throughout the text.

Recall the description of a principal G-bundle $\pi : P \longrightarrow M$ via local trivializations and transition functions given at the beginning of this section. On a trivializing open set $U_\alpha \subseteq M$ we have a bundle chart

$$\begin{aligned} \Phi_\alpha : \pi^{-1}(U_\alpha) &\longrightarrow U_\alpha \times G \\ p &\mapsto (\pi(p), \phi_\alpha(p)), \end{aligned}$$

and the transition functions defined on intersections $U_\alpha \cap U_\beta$

$$\phi_{\alpha\beta}(x) = \phi_\alpha(p)\phi_\beta(p)^{-1},$$

which obey the cocycle conditions (1.8). Moreover, we write the induced trivializations on the associated bundles $\mathrm{Ad}(P)$ and $\mathrm{ad}(P)$ over U_α with the same notation, as there is very little chance of confusion in context. For $\tilde{\pi} : \mathrm{Ad}(P) \longrightarrow M$,

$$\begin{aligned} \tilde{\Phi}_\alpha : \tilde{\pi}^{-1}(U_\alpha) &\longrightarrow U_\alpha \times G \\ \lfloor p, g \rceil &\mapsto (\pi(p), \tilde{\phi}_\alpha(\lfloor p, g \rceil)), \end{aligned}$$

with $\tilde{\phi}_\alpha(\lfloor p, g \rceil) = \phi_\alpha(p) g \phi_\alpha(p)^{-1}$. And for $\tilde{\pi} : \mathrm{ad}(P) \to M$,

$$\begin{aligned} \tilde{\Phi}_\alpha : \tilde{\pi}^{-1}(U_\alpha) &\longrightarrow U_\alpha \times \mathfrak{g} \\ \lfloor p, \xi \rceil &\mapsto (\pi(p), \tilde{\phi}_\alpha(\lfloor p, \xi \rceil)), \end{aligned}$$

with $\tilde{\phi}_\alpha(\lfloor p, \xi \rceil) = \phi_\alpha(p) \xi \phi_\alpha(p)^{-1}$. Throughout this subsection we will assume this setting.

For a gauge transformation $u \in \mathcal{G}(P)$, let $\tilde{u} \in \Gamma(\mathrm{Ad}(P))$ be the corresponding section, such that $\tilde{u}(x) = \lfloor p, u(p) \rceil$ for $x = \pi(p)$. Then we can use this to define u locally on U_α,

$$u_\alpha := \tilde{\phi}_\alpha \circ \tilde{u} : U_\alpha \to G, \tag{1.17}$$

and this acts on $U_\alpha \times G$ by $(x, g) \mapsto (x, g u_\alpha(x))$. Thus, for any $p \in \pi^{-1}(x)$,

$$u_\alpha(x) = \phi_\alpha(p) u(p) \phi_\alpha(p)^{-1}$$

is well defined, and we can recover u on U_α,

$$u(p) = \phi_\alpha(p)^{-1} u_\alpha(x) \phi_\alpha(p).$$

If we assume this to be valid for all $\alpha \in A$, then using the transition functions on $U_\alpha \cap U_\beta$ we get the transition identity

$$u_\beta = \phi_{\alpha\beta}^{-1} u_\alpha \phi_{\alpha\beta}. \tag{1.18}$$

Likewise, any collection of G-valued functions $\{u_\alpha\}_{\alpha \in A}$ satisfying (1.18) uniquely defines a global gauge transformation in the same way as we used to recover $u(p)$. Equivalently, a global gauge transformation must satisfy

$$\phi_{\alpha\beta} = u_\alpha^{-1} \phi_{\alpha\beta} u_\beta.$$

The local description of connections on open sets $U \subseteq M$ is very similar, using the isomorphism $\Omega^k_{bas}(P, \mathfrak{g}) \simeq \Omega^k(M, \mathrm{ad}(P))$ and then the second component of the local trivialization for $\mathrm{ad}(P)$. For $\tau \in \Omega^k_{bas}(P, \mathfrak{g})$ let $\tilde{\tau} \in \Omega^k(M, \mathrm{ad}(P))$ be the corresponding form given by (1.9). This defines τ locally,

$$\tau_\alpha := \tilde{\phi}_\alpha \circ \tilde{\tau} \in \Omega^k(U_\alpha, \mathfrak{g}),$$

and similar to the case of the gauge transformation, on an intersection $U_\alpha \cap U_\beta$ these local forms will satisfy

$$\tau_\beta = \phi_{\alpha\beta}^{-1} \tau_\alpha \phi_{\alpha\beta},$$

and the global form can be reconstructed as

$$\tau(Y_1, \ldots, Y_k) = \phi_\alpha(p)^{-1} \tau_\alpha(\mathrm{d}_p\pi(Y_1), \ldots, \mathrm{d}_p\pi(Y_k))\phi_\alpha(p), \ \forall Y_1, \ldots, Y_k \in T_pP.$$

In the case of connections it is necessary to choose a reference connection in order to use this isomorphism, and there is no canonical choice if the bundle is not flat. However, locally on $\pi^{-1}(U_\alpha)$ there is a natural choice, namely $\tilde{A}_\alpha := \phi_\alpha^{-1}\,\mathrm{d}\phi_\alpha$. Then for the local representative $A_\alpha \in \Omega^1(U_\alpha, \mathfrak{g})$ of $A \in \mathcal{A}(P)$ we write

$$A_\alpha(\mathrm{d}_p\pi(Y)) = \phi_\alpha(p)A(Y)\phi_\alpha(p)^{-1} - \mathrm{d}_p\phi_\alpha(Y)\phi_\alpha(p)^{-1}, \ \forall Y \in T_pP. \tag{1.19}$$

Assuming that A can be recovered over U_α as

$$A(Y) = \phi_\alpha(p)^{-1}A_\alpha(\mathrm{d}_p\pi(Y))\phi_\alpha(p) + \phi_\alpha(p)^{-1}\,\mathrm{d}_p\phi_\alpha(Y)$$

we will once again get a transition identity for the local representatives of connections over intersections. This is stated in the following lemma, along with the local formula for the gauge action, which shows that locally a gauge transformation can be thought of as a change of trivialization.

Lemma 1.11. *For a connection $A \in \mathcal{A}(P)$, its local representatives A_α, A_β have to meet the following on $U_\alpha \cap U_\beta$:*

$$A_\beta = \phi_{\alpha\beta}^{-1} A_\alpha \phi_{\alpha\beta} + \phi_{\alpha\beta}^{-1}\,\mathrm{d}\phi_{\alpha\beta}.$$

Moreover, the local effect of a gauge transformation is

$$(u^*A)_\alpha = u_\alpha^{-1} A_\alpha u_\alpha + u_\alpha^{-1}\,\mathrm{d}u_\alpha.$$

Proof. Both of the affirmations are checked with straightforward calculations. For the change in trivialization, omitting the evaluations at p, Y and $\mathrm{d}_p\pi(Y)$, we calculate

$$\begin{aligned}
A_\beta &= \phi_\beta A \phi_\beta^{-1} - \mathrm{d}\phi_\beta \cdot \phi_\beta^{-1} \\
&= \phi_\beta(\phi_\alpha^{-1} A_\alpha \phi_\alpha + \phi_\alpha^{-1}\,\mathrm{d}\phi_\alpha)\phi_\beta^{-1} - \mathrm{d}\phi_\beta \cdot \phi_\beta^{-1} \\
&= \phi_{\alpha\beta}^{-1} A_\alpha \phi_{\alpha\beta} + \phi_{\alpha\beta}^{-1}\,\mathrm{d}\phi_\alpha \cdot \phi_\beta^{-1} - \mathrm{d}\phi_\beta \cdot \phi_\beta^{-1} \\
&= \phi_{\alpha\beta}^{-1} A_\alpha \phi_{\alpha\beta} + \phi_{\alpha\beta}^{-1}(\mathrm{d}\phi_\alpha \cdot \phi_\beta^{-1} - \phi_{\alpha\beta}\,\mathrm{d}\phi_\beta \cdot \phi_\beta^{-1}) \\
&= \phi_{\alpha\beta}^{-1} A_\alpha \phi_{\alpha\beta} + \phi_{\alpha\beta}^{-1}(\mathrm{d}\phi_\alpha \cdot \phi_\beta^{-1} + \phi_\alpha \phi_\beta^{-1} \phi_\beta\,\mathrm{d}\phi_\beta^{-1}) \\
&= \phi_{\alpha\beta}^{-1} A_\alpha \phi_{\alpha\beta} + \phi_{\alpha\beta}^{-1}\,\mathrm{d}\phi_{\alpha\beta}.
\end{aligned}$$

Then for the local gauge action

$$(u^*A)_\alpha = \phi_\alpha(u^*A)\phi_\alpha^{-1} - \mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1}$$

$$
\begin{aligned}
&= \phi_\alpha(u^{-1}Au + u^{-1}\,\mathrm{d}u)\phi_\alpha^{-1} - \mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1} \\
&= \phi_\alpha(\phi_\alpha^{-1}u_\alpha^{-1}\phi_\alpha A\phi_\alpha^{-1}u_\alpha\phi_\alpha + \phi_\alpha^{-1}u_\alpha^{-1}\phi_\alpha\,\mathrm{d}(\phi_\alpha^{-1}u_\alpha\phi_\alpha))\phi_\alpha^{-1} - \mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1} \\
&= u_\alpha^{-1}(\phi_\alpha A\phi_\alpha^{-1})u_\alpha \\
&\qquad + u_\alpha^{-1}\phi_\alpha\,\mathrm{d}\phi_\alpha^{-1} \cdot u_\alpha + u_\alpha^{-1}\,\mathrm{d}u_\alpha + \mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1} \\
&\qquad\qquad - \mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1} \\
&= u_\alpha^{-1}A_\alpha u_\alpha + u_\alpha^{-1}\,\mathrm{d}\phi_\alpha \cdot \phi_\alpha^{-1}u_\alpha + u_\alpha^{-1}\phi_\alpha\,\mathrm{d}\phi_\alpha^{-1} \cdot u_\alpha + u_\alpha^{-1}\,\mathrm{d}u_\alpha \\
&= u_\alpha^{-1}A_\alpha u_\alpha + u_\alpha^{-1}\,\mathrm{d}u_\alpha.
\end{aligned}
$$

□

Finally, the covariant exterior derivative $\mathrm{d}_A : \Omega^k_{bas}(P, \mathfrak{g}) \to \Omega^{k+1}_{bas}(P, \mathfrak{g})$ induced by a connection $A \in \mathcal{A}(P)$ as in (1.10) can be written locally on U_α, taking the local representatives $A_\alpha \in \Omega^1(U_\alpha, \mathfrak{g})$ and $\tau_\alpha \in \Omega^k(U_\alpha, \mathfrak{g})$,

$$(\mathrm{d}_A\tau)_\alpha = \mathrm{d}\tau_\alpha + [A_\alpha \wedge \tau_\alpha];$$

its formal adjoint $\mathrm{d}_A^* : \Omega^k_{bas}(P, \mathfrak{g}) \to \Omega^{k+1}_{bas}(P, \mathfrak{g})$ is given locally as

$$(\mathrm{d}_A^*\tau)_\alpha = \mathrm{d}^*\tau_\alpha - (-1)^{(n-k)(k-1)} * [A_\alpha \wedge *\tau_\alpha];$$

and the curvature (1.12) can be written

$$(F_A)_\alpha = \mathrm{d}A_\alpha + \frac{1}{2}[A_\alpha \wedge A_\alpha] \in \Omega^2(U_\alpha, \mathfrak{g}),$$

and transforms under a change of trivialization in the same way it does under a gauge transformation, $(F_A)_\beta = \phi_{\alpha\beta}^{-1}(F_A)_\alpha\phi_{\alpha\beta}$.

A word on notation. Occasionally, when working locally but not specifying a bundle atlas, we will drop subscripts α and a connection $A \in \mathcal{A}(P)$ will be represented by $A \in \Omega^1(U, \mathfrak{g})$, or a gauge transformation $u \in \mathcal{G}(P)$ will be represented by $u : U \to G$. When specifying a global reference connection $\tilde{A}$, a connection $A \in \mathcal{A}(P)$ will be denoted $A = \tilde{A} + \alpha$, and $\alpha \in \Omega^1_{bas}(P, \mathfrak{g})$ will be its representative.

1.2.2 Several results on G, $\mathfrak{g}$ and the action on P

In this subsection we fix some notation pertaining to the action of G on P and $\mathfrak{g}$, define the inner product that will be fixed on $\mathfrak{g}$ and some other notions that depend on it, such as the metric on G, and prove a result on the derivative of the adjoint action which will be needed later.

We begin with some notation: For $p \in P$ and $g, h \in G$,

$$
\begin{aligned}
gh &= L_g(h), \quad L_g : G \to G, \\
hg &= R_g(h), \quad R_g : G \to G, \\
pg &= \tilde{L}_p(g), \quad \tilde{L}_p : G \to P, \\
pg &= \tilde{R}_g(p), \quad \tilde{R}_g : P \to P.
\end{aligned}
$$

For $p \in P$, $\xi \in \mathfrak{g}$, and $v \in T_pP$,

$$p\xi := \mathrm{d}_{\mathbb{1}}\tilde{L}_p(\xi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p\exp(t\xi) \in T_pP,$$

$$vg := \mathrm{d}_p \tilde{R}_g(v) \qquad\qquad \in T_{pg}P,$$

and for $g, h \in G$, $\xi \in \mathfrak{g}$ and $v \in T_hG$,

$$g\xi := \mathrm{d}_{\mathbb{1}} L_g(\xi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g\exp(t\xi) \in T_gG,$$
$$gv := \mathrm{d}_h L_g(v) \qquad\qquad \in T_{gh}G.$$

We need to fix a certain inner product on $\mathfrak{g}$. The following lemma is the main reason that we restrict to compact Lie groups; its proof can be found in [Kna96, proposition 4.24], and the subsequent properties of the inner product and the induced metrics can be found in [Weh04, remark A.3].

Lemma 1.12. *Let G be a compact Lie group. Then the Lie algebra $\mathfrak{g}$ admits an inner product which is invariant under the adjoint action of G on $\mathfrak{g}$, that is, for all $\xi, \eta \in \mathfrak{g}$ and $g \in G$,*

$$\langle g\xi g^{-1}, g\eta g^{-1}\rangle = \langle \xi, \eta\rangle.$$

This inner product satisfies, for all $\xi, \eta, \zeta \in \mathfrak{g}$,

$$\langle [\xi, \eta], \zeta\rangle = \langle \xi, [\eta, \zeta]\rangle, \tag{1.20}$$

and it can be rescaled in such a way that the associated norm $|\xi| = \sqrt{\langle \xi, \xi\rangle}$ satisfies, for all $\xi, \eta \in \mathfrak{g}$,

$$|[\xi, \eta]| \leq |\xi| \cdot |\eta|. \tag{1.21}$$

We will fix this inner product on $\mathfrak{g}$ throughout all of this text.

The inner product on $\mathfrak{g}$ also induces a **metric on** TG,

$$\langle X, Y\rangle_G := \langle g^{-1}X, g^{-1}Y\rangle_{\mathfrak{g}}, \ \forall X, Y \in T_gG,$$

where $g^{-1}X$ is understood in the sense defined above. It is clear that with this metric left and right multiplication are isometries of G. If we denote by $\exp_g$ the exponential map with base point $g \in G$, then for all $\xi \in \mathfrak{g}$ and $g \in G$,

$$\exp_g(g\xi) = g\exp(\xi),$$
$$\exp(g^{-1}\xi g) = g^{-1}\exp(\xi)g.$$

Moreover, the geodesics are 1-parameter subgroups,

$$\exp((s+t)\xi) = \exp(s\xi)\exp(t\xi),$$

for all $s, t \in \mathbb{R}$ and $\xi \in \mathfrak{g}$.

Furthermore, we can define a **geodesic distance on** G,

$$\mathrm{d}_G(g, h) := \inf\left\{|X|, X \in T_gG \text{ and } h = \exp_g(X)\right\}, \tag{1.22}$$

which is invariant under left and right multiplication, and this can be used to define a metric on $C^0(U, G)$. For maps u and v from some domain U to G, it will be denoted by

$$\mathrm{d}(u, v) := \sup_{x\in U} \mathrm{d}_G(u(x), v(x)),$$

and this too will be invariant under left and right multiplication by continuous maps.

We also define a **convex geodesic ball** of radius R around $\mathbb{1} \in G$ to be such that

(i) the exponential map is a bijection between $B_R(\mathbb{1}) \subseteq \mathfrak{g}$ and $B_R(\mathbb{1}) \subseteq G$,

(ii) and for all $g, h \in B_R(\mathbb{1})$ there is a unique minimal geodesic from g to h that lies entirely within $B_R(\mathbb{1})$.

For the existence of such balls see e.g. [GHL90, 2.89, 2.90], Moreover, because left multiplications are isometries of G, there exist convex geodesic balls of same radius around any $g \in G$.

Finally, we prove a small lemma on the (covariant) derivative of the adjoint action on k-forms.

Lemma 1.13. *For $\tau \in \Omega^k(U, \mathfrak{g})$ and $u : U \to G$,*

$$\mathrm{d}(\mathrm{Ad}_u \tau) = \mathrm{Ad}_u \, \mathrm{d}\tau + [\,\mathrm{d}u \cdot u^{-1} \wedge \mathrm{Ad}_u \tau]. \tag{1.23}$$

Moreover, for the covariant derivative ∇ induced by the canonical flat connection on the trivial vector bundle $U \times \mathfrak{g} \to U$,

$$\nabla(u\tau u^{-1}) = u(\nabla\tau)u^{-1} + [\,\mathrm{d}u \cdot u^{-1}, u\tau u^{-1}]. \tag{1.24}$$

Proof. For a fixed point $p \in U$ and some $X \in TU$,

$$\begin{aligned}
\mathrm{d}_p(u\tau u^{-1})(X) &= \mathrm{d}(\mathrm{Ad}_{u \cdot u(p)^{-1}}(u(p)\tau u(p)^{-1}))(X) \\
&\overset{(1)}{=} \Big(\mathrm{d}_{\mathbb{1}}\,\mathrm{Ad} \circ \mathrm{d}_p(u \cdot u(p)^{-1}) \cdot u(p)\tau u(p)^{-1} + \mathrm{Ad}_{u \cdot u(p)^{-1}}\, \mathrm{d}_p(u(p)\tau u(p)^{-1})\Big)(X) \\
&\overset{(2)}{=} \Big([\,\mathrm{d}_p u \cdot u(p)^{-1}, u(p)\tau u(p)^{-1}] + u\, \mathrm{d}_p \tau u^{-1}\,(p)\Big)(X) \\
&= \Big([\,\mathrm{d}u \cdot u^{-1}, u\tau u^{-1}] + u\, \mathrm{d}\tau \cdot u^{-1}\Big) X(p),
\end{aligned}$$

where (1) follows from looking at Ad as a map on a product manifold, $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$, similar to what was done in the calculation of the gauge group action, lemma 1.10, and (2) is just $\mathrm{d}_{\mathbb{1}}\,\mathrm{Ad}(\xi)\eta = \mathrm{ad}(\xi)\eta = [\xi, \eta]$.

We now use (1.23) to prove the second identity (1.24). With the canonical flat connection $\nabla^E = \mathrm{d}$, evaluating at vector fields $X_j \in TU$ and letting $\tilde{\tau} := \tau(X_1, \ldots, X_k)$,

$$\begin{aligned}
\nabla(u\tau u^{-1})(X_0, \ldots, X_k) &= \nabla^E_{X_0}(u\tilde{\tau}u^{-1}) - u\tau u^{-1}(\nabla_{X_0}X_1, \ldots, X_k) \\
&\qquad - \cdots - u\tau u^{-1}(X_1, \ldots, \nabla_{X_0}X_k) \\
&= \mathrm{d}(\mathrm{Ad}_u \tilde{\tau})X_0 - u\tau(\nabla_{X_0}X_1, \ldots, X_k)u^{-1} \\
&\qquad - \cdots - u\tau(X_1, \ldots, \nabla_{X_0}X_k)u^{-1} \\
&= [\,\mathrm{d}u(X_0)u^{-1}, u\tilde{\tau}u^{-1}] + u\,\mathrm{d}\tilde{\tau}(X_0)u^{-1} - u\tau(\nabla_{X_0}X_1, \ldots, X_k)u^{-1} \\
&\qquad - \cdots - u\tau(X_1, \ldots, \nabla_{X_0}X_k)u^{-1} \\
&= [\,\mathrm{d}u \cdot u^{-1}, u\tau u^{-1}](X_0, \ldots, X_k) + u\nabla\tau(X_0, \ldots, X_k)u^{-1} \\
&= \Big([\,\mathrm{d}u \cdot u^{-1}, u\tau u^{-1}] + u(\nabla\tau)u^{-1}\Big)(X_0, \ldots, X_k).
\end{aligned}$$

□

1.3 An analytic interlude

Because it is modelled on $\Omega^1(M, \mathrm{ad}(P))$, it is easily seen that $\mathcal{A}(P)$ is an infinite dimensional affine space, and $\mathcal{G}(P)$ can be shown to be an infinite dimensional Lie group. It is known that C^∞ spaces are not the ideal setting for doing functional analysis, as they are not Banach spaces. What is more, the Yang–Mills equations are partial differential equations; in PDE theory, the C^k spaces are not necessarily the best domains to look for solutions, and it turns out to be more profitable to work with so-called weak derivatives and weak solutions. For these reasons, we will rework our framework into the less regular realm of Sobolev spaces. From now on, we trade smoothness in for powerful analytic results.

In this section, our aim is to give the definitions and most relevant results on Sobolev spaces of sections of fibre bundles. For an introduction to the theory of Sobolev spaces on $\mathbb{R}^n$, we recommend [Eva10, chapter 5]; for definitions and results on Sobolev spaces of functions on Riemannian manifolds, see [Aub82, chapter 2]. For Sobolev spaces of sections of vector bundles, see [Nic07, section 10.4.2]. We mainly follow [Weh04, appendix B].

Let (M, g) be a compact Riemannian manifold, $E \to M$ a vector bundle, and choose a bundle metric on E and a compatible connection. We have shown that with these choices it is possible to define: a covariant derivative

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

given by the connection on E and the Levi-Civita connection for the Riemannian metric g as in (1.2); a pointwise inner product on $\Omega^k(M, E)$,

$$\langle \alpha, \beta \rangle = *\langle \alpha \wedge *\beta \rangle : M \to \mathbb{R},$$

given by the bundle metric and the Riemannian metric as in (1.7), which we then use to define $|\alpha| := \sqrt{\langle \alpha, \alpha \rangle}$; and a way to integrate functions over M given by the Riemannian metric, using the volume form as in (1.6).

Definition 1.14 (Sobolev space of sections of vector bundles)**.** For $k \in \mathbb{N}_0$ and $1 \leq p < \infty$, the Sobolev space $W^{k,p}(M, E)$ of sections of the vector bundle $E \to M$ is defined as the completion of $\Gamma(E)$ with respect to the $W^{k,p}$-norm,

$$\|\alpha\|_{k,p} = \sum_{j=1}^{k} \left\| \nabla^j \alpha \right\|_p,$$

where

$$\left\| \nabla^j \alpha \right\|_p = \left(\int_M \left| \nabla^j \alpha \right|^p \right).$$

Several properties of the usual Sobolev spaces on Euclidean space will generalize to Sobolev spaces of sections. From the definition, it is clear that $W^{k,p}(M, E)$ are Banach spaces, and therefore many important results from functional analysis are valid, in particular those from Appendix B. Moreover, this definition directly gives the density of smooth sections, which makes it possible to give proofs using approximating sequences of smooth sections, for which stronger results are valid; see, e.g., lemma 1.20 below. Other properties and results will be consequences of the following characterization:

Remark 1.15. Consider a finite bundle atlas $(U_i, \Phi_i)_{i=1}^N$ of $E \to M$, with $\Phi_i : E|_{U_i} \to V_i \times \mathbb{R}^m$ for $\phi : U_i \to V_i \subseteq \mathbb{R}^n$ a coordinate chart of M and m the rank of E. A section $\alpha \in \Gamma(E)$ which is locally $\alpha : U_i \to E|_{U_i}$ is represented by $\Phi_{i*}\alpha : V_i \to \mathbb{R}^m$, or with m components $(\Phi_{i*}\alpha)_j : V_i \to \mathbb{R}$. Then for $k \in \mathbb{N}_0$ and $1 \leq p < \infty$, the $W^{k,p}$-norm defined above is equivalent to

$$\sum_{i=1}^{N}\sum_{j=1}^{m} \|(\Phi_{i*}\alpha)_j\|_{k,p(V_i)},$$

where the $W^{k,p}$ norm on functions $\mathbb{R}^n \to \mathbb{R}$ is the usual one. Therefore α lies in $W^{k,p}(M,E)$ if and only if its local components $(\Phi_{i*}\alpha)_j$ are $W^{k,p}$-functions for all coordinate patches $i = 1, \dots, N$.

Note that it is essential that the base manifold is compact for the sum above to be finite and the norms equivalent.

From this characterization it also becomes clear that when M is compact, the space $W^{k,p}(M,E)$ will not depend on the choices involved, as the $W^{k,p}$-norms induced will be equivalent. The choices were: the metric on M, the bundle metric on E, and the compatible connection on E.

Remark 1.16 (On the norm of a metric)**.** In chapter 2, we will several times look at the $W^{k,\infty}$-norm of a Riemannian metric g. The metric is a tensor, $g \in \Gamma(\oplus^2 T^*M)$, and therefore as a section of a vector bundle it makes sense to define its $W^{k,p}$-norm as above, but with respect to what metric on M? As we have noted, over a compact manifold all of the norms will be equivalent, so any fixed choice is valid. In chapter 2 we will mostly be working over open sets in Euclidean (half) space, and thus we may canonically choose the Euclidean metric in this case.

By far the most important result from the usual theory of Sobolev spaces which is also valid for sections of bundles is the Sobolev embeddings and estimates. Before stating the result, we need to define the norm on the spaces C^j of continuous functions, for $j \in \mathbb{N}_0$. For $\alpha \in C^0(M, \otimes^k T^*M \otimes E)$ a k-form,

$$\|\alpha\|_\infty := \sup_{x\in M} |\alpha(x)|,$$

where $|\cdot|$ comes from the pointwise inner product defined on $\Omega^k(M,E)$ above; then for a section $\alpha \in C^j(M,E)$,

$$\|\alpha\|_{j,\infty} := \sup_{k\leq j} \left\|\nabla^k \alpha\right\|_\infty.$$

Finally, note that $W^{0,p}(M,E)$ is simply $L^p(M,E)$, and these L^p-spaces are also included in the following result, as $j = 0$ is allowed.

A word on notation. We will denote the norm on the Sobolev space $W^{k,p}(M,E)$ over the Riemannian manifold (M,g) as $\|\cdot\|_{g;\,k,p\,(M)}$, and when the metric and the space are clear from context, they will be omitted from the norm.

Theorem 1.17 (Sobolev embeddings and estimates)**.** *Let $E \to M$ be a vector bundle over a compact Riemannian n-manifold, $j < k \in \mathbb{N}$ and $1 \leq p, q < \infty$.*

(i) *If $k - \frac{n}{p} \geq j - \frac{n}{q}$ then the inclusion*

$$W^{k,p}(M,E) \hookrightarrow W^{j,q}(M,E)$$

is continuous, i.e., there exists a constant C_W such that for $\alpha \in W^{k,p}(M,E)$,

$$\|\alpha\|_{j,q} \leq C_W \|\alpha\|_{k,p} .$$

(ii) *If $k - \frac{n}{p} > j - \frac{n}{q}$, this inclusion is a compact map.*

(iii) *Furthermore, if $k - \frac{n}{p} > j$, there is a continuous embedding $W^{k,p}(M) \hookrightarrow C^j(M)$, i.e., there exists a constant C_W such that for $\alpha \in W^{k,p}(M,E)$,*

$$\|\alpha\|_{j,\infty} \leq C_W \|\alpha\|_{k,p} .$$

Moreover, this inclusion is compact.

The generalization of the result from bounded domains in $\mathbb{R}^n$ to the case of vector bundles is straighforward using remark 1.15, see the discussion after Theorem A.2 in [Weh04].

These embeddings will cause a few hypotheses to appear particularly frequently: the inclusion

$$W^{k,p} \hookrightarrow C^0$$

will lead to the condition $kp > n$; and the inclusion

$$W^{1,p} \hookrightarrow L^{2p}$$

will lead to $p \geq \frac{n}{2}$. When looking at the Sobolev spaces of connections and gauge transformations, we will generally be working with the $W^{1,p}$ and $W^{2,p}$ spaces, and for $k = 2$ the hypothesis $p > \frac{n}{2}$ guarantees both of these embeddings.

Of central importance to the results in this text is the Banach–Alaoglu theorem, which we state and use in the following form:

Theorem 1.18 (Banach–Alaoglu). *Let $k \in \mathbb{N}$ and $1 < p < \infty$, and let $E \to M$ be a vector bundle over a compact Riemannian manifold. Then every bounded sequence in $W^{k,p}(M,E)$ has a weakly convergent subsequence.*

The generalization from the result for bounded domains in $\mathbb{R}^n$ also follows from remark 1.15. As for the relation of this result to the usual Banach–Alaoglu theorem from functional analysis, note that Sobolev spaces are reflexive.

A recurrent argument used in proofs is to make use of the Banach–Alaoglu theorem to find a weakly convergent subsequence for a bounded sequence, and then use some compact Sobolev embedding to find a further subsequence which converges uniformly.

Finally, in order to deal with manifolds with boundary, we need the following result.

Theorem 1.19 (Trace theorem). *Let M be a compact Riemannian n-manifold and let $1 \leq p < \infty$. The restriction to the boundary ∂M is a bounded linear operator $W^{1,p}(M) \to L^p(\partial M)$.*

Proof. See [Weh04, theorem B.10]. □

This means that for any $h \in W^{1,p}(M)$,

$$\|h|_{\partial M}\|_{p(\partial M)} \leq C \|h\|_{1,p(M)} .$$

In particular, any sequence of smooth functions that approximates h in the $W^{k,p}(M)$-norm restricts to a sequence in the boundary that approximates $h|_{\partial M}$ in the $L^p(\partial M)$-norm, since

$$\|(h_j - h)|_{\partial M}\|_{p(\partial M)} \leq c \|h_j - h\|_{1,p(M)} \leq c \|h_j - h\|_{k,p(M)} \to 0.$$

Using the trace theorem and approximating sequences, we can generalize Stokes's theorem to the non-smooth case. We state and prove this as a lemma, and make implicit use of it whenever we use Stokes's theorem.

Lemma 1.20 (Sobolev Stokes's Lemma)**.** *Let M be a compact n-dimensional manifold with boundary, $E \to M$ a trivial vector bundle and $\alpha \in W^{k,p}(M, \otimes^{n-1} T^*M \otimes E)$ an $(n-1)$-form. Then*

$$\int_{\partial M} \alpha|_{\partial M} = \int_M \mathrm{d}\alpha.$$

Proof. We can approximate α in the $W^{1,p}(M)$-norm by a sequence $(\alpha_j)_{j \in \mathbb{N}} \subseteq \Omega^{n-1}(M, E)$ of smooth forms, and using the trace theorem $(\alpha_j|_{\partial M}) \subseteq \Omega^{n-1}(\partial M, E)$ also approximates $\alpha|_{\partial M}$ in the $L^p(\partial M)$-norm. Then, because of the triviality of the bundle, Stokes's theorem can be applied to each component of the smooth forms,

$$\int_M \mathrm{d}\alpha = \lim_{j \to \infty} \int_M \mathrm{d}\alpha_j = \lim_{j \to \infty} \int_{\partial M} \alpha_j|_{\partial M} = \int_{\partial M} \alpha|_{\partial M} .$$

□

Finally, we will make use of a few product inequalities, but we state them in appendix B so as not to clutter the more important results here.

Let us shift our focus now to maps between manifolds, say M and X. Suppose that M is a compact n-dimensional manifold, and X is an ℓ-manifold, and fix on X an embedding $\Phi : X \to \mathbb{R}^{2\ell+1}$, an atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$, and a metric. We then get the following two results.

Proposition 1.21. *Let M and X be as above, and let $k \in \mathbb{N}$ and $1 \leq p < \infty$ be such that $kp > n$. For $u \in C^0(M, X)$ the following are equivalent:*

- *(i)* $\phi_\alpha \circ u \in W^{k,p}(u^{-1}(U_\alpha, \mathbb{R}^\ell)$ *for all* $\alpha \in A$,
- *(ii)* $\Phi \circ u \in W^{k,p}(M, \mathbb{R}^{2\ell+1})$,
- *(iii)* $u = \exp_s(V)$ *for some* $s \in C^\infty(M, X)$ *and* $V \in W^{k,p}(M, s^*TX)$.

In case $X = G$ is a Lie group with Lie algebra $\mathfrak{g}$, the last item can be reformulated as $u = s \cdot \exp(\xi)$ for $s \in C^\infty(M, G)$ and $\xi \in W^{k,p}(M, \mathfrak{g})$, and there is another equivalence,

- *(iv)* $u^{-1}\,\mathrm{d}u \in W^{k-1,p}(M, T^*M \otimes \mathfrak{g})$.

Proposition 1.22. *For a sequence $(u_i)_{i \in \mathbb{N}}$ and some u, all of which satisfy the equivalent conditions in the previous proposition, and under the same assumptions, the following are equivalent:*

- *(i)* *u_i converges to u in the C^0-topology and $\phi_\alpha \circ u_i$ converges to $\phi_\alpha \circ u$ with respect to the $W^{k,p}$-norm for all $\alpha \in A$,* *(ii)* $\Phi \circ u_i$ *converges to* $\Phi \circ u$ *with respect to the* $W^{k,p}$*-norm,*

(iii) *there exist* $s \in C^\infty(M, X)$, $V \in W^{k,p}(M, s^*TX)$, *and for sufficiently large* $i \in \mathbb{N}$ *there are* $V_i \in W^{k,p}(M, s^*TX)$ *such that* $u = \exp_s(V)$, $u_i = \exp_s(V_i)$ *and the* V_i *converge to* V *in the* $W^{k,p}$*-norm,*

and in case $X = G$ *is a Lie group with Lie algebra* $\mathfrak{g}$, *once again the last item can be reformulated with* $u = s \cdot \exp(\xi)$, $u_i = s \cdot \exp(\xi_i)$ *and* $\xi_i \to \xi$ *in* $W^{k,p}(M, \mathfrak{g})$, *and there is another equivalence,*

(iv) u_i *converges to* u *in the* C^0*-topology and* $u_i^{-1}\,\mathrm{d}u_i$ *converges to* $u^{-1}\,\mathrm{d}u$ *in* $W^{k-1,p}(M, T^*M \otimes \mathfrak{g})$.

Proof. See [Weh04, lemmata B.5 and B.7]. □

A key lemma in the proof of the equivalences in proposition 1.22 is the following, which will also be used elsewhere in this text and for that reason we state it. This lemma is the reason for the assumption $kp > n$ in the previous results, and this hypothesis is needed here because of the embedding $W^{k,p} \hookrightarrow C^0$.

Lemma 1.23. *Let* $U \subseteq M$ *be an open subset of a compact* n*-manifold,* $1 \leq p < \infty$ *and* $k, m, N \in \mathbb{N}$ *such that* $kp > n$. *If* $f \in C^k(V \subseteq \mathbb{R}^m, \mathbb{R}^N)$, *then composition with* f *is a continuous map, that is*

$$\begin{aligned} W^{k,p}(U, V) &\longrightarrow W^{k,p}(U, \mathbb{R}^N) \\ u &\mapsto f \circ u. \end{aligned}$$

Proof. See [Weh04, lemma B.8]. This result too is stated there for bounded domains in $\mathbb{R}^n$ and then generalizes via remark 1.15. □

At last, we may give the definitions for Sobolev spaces of maps of manifolds and of sections of fibre bundles.

Definition 1.24 (Sobolev space of maps of manifolds)**.** For M and X manifolds as described above, the Sobolev space $W^{k,p}(M, X)$ is given as a set by functions $u \in C^0(M, X)$ such that u satisfies the equivalent statements of lemma 1.21, and the topology on this space is given by defining its convergent sequences using the equivalent statements of lemma 1.22.

Definition 1.25 (Sobolev space of sections of fibre bundles)**.** Let $X \hookrightarrow F \xrightarrow{\pi} M$ be a fibre bundle, and fix a bundle atlas $(U_\alpha, \tau_\alpha)_{\alpha \in A}$. In every local trivialization $\pi \times \tau_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times X$ a section u is represented by $\tau_\alpha \circ u : U_\alpha \to X$. Then define the Sobolev space $W^{k,p}(M, F)$ of sections of this fibre bundle to consist of all sections u such that

$$\tau_\alpha \circ u \in W^{k,p}(U_\alpha, X)$$

for all $\alpha \in A$, following the definition above for maps of manifolds. The topology is once again given by the equivalent conditions in proposition 1.22 via convergence in a bundle atlas.

We make a few observations. First, proposition 1.21 shows that $W^{k,p}(M, X)$ is independent of the choice of atlas (used in condition (i)), embedding (used in condition (ii)) and metric (used in condition (iii)), and therefore it is well defined for $kp > n$. Moreover, the definition via embeddings makes it possible to extend the Sobolev embeddings in theorem 1.17 to the Sobolev spaces defined above.

Partial differential equations and elliptic operators

When working with partial differential equations (PDEs) in an abstract context, there are a few aspects of solutions that one is usually concerned with: existence, uniqueness and regularity. Existence, of course, has to do with whether there is a solution to the given problem in the chosen domain, which is usually a Sobolev space; if a solution exists, it is often good to know whether it is unique, at least up to some equivalence class (up to a measure zero set, for instance); regularity means, roughly, the degree of smoothness of a solution.

In this work, the main PDE we are concerned with is the Yang–Mills equation. However, in the proofs of the main theorems we will also need to quote results for the Neumann problem and the $\mathrm{d} + \mathrm{d}^*$ operator.

For the Yang–Mills equation, we will need a regularity result (proposition 1.39) which will be crucial in the proof of the strong compactness theorem 3.9; this will be discussed in more detail in the next section, when discussing weak Yang–Mills connections. For the Neumann problem, we will need results on existence and uniqueness, as well as an estimate; these will be used only once, in the proof of theorem 2.2, and can be found in appendix B. For the $\mathrm{d} + \mathrm{d}^*$ operator, we will need the following estimate, which will be used in chapter 2 in the motivation and proof of theorem 2.2.

Theorem 1.26 ([Weh04], theorem 5.1)**.** *Let M be a compact manifold with (possibly empty) boundary. For $1 < p < \infty$, if $A \in W^{1,p}(M, T^*M)$ satisfies $*A|_{\partial M} = 0$ and $H^1(M;\mathbb{R}) = 0$, then there is a constant C such that*

$$\|A\|_{1,p} \leq C(\|\mathrm{d}A\|_p + \|\mathrm{d}^*A\|_p).$$

Moreover, this constant depends $W^{1,\infty}$-continuously on the metric on M. This is also valid for 1-forms with values in a finite dimensional vector space.

All of these results are proved individually and in great detail in the book [Weh04], and indeed have many chapters dedicated to them. However, it is worth commenting that all of the operators associated to these equations are part of a class of special operators with good properties, called *elliptic* operators. More often than not, elliptic theory is considered standard in geometric analysis texts. For instance, in Uhlenbeck's original paper [Uhl82], after the statement of theorem 1.3, she writes:[8]

> "Regularity of solutions of Yang–Mills equations for connections $A \in \mathcal{A}^{1,p}$, $2p \geq \dim M$ follows rather easily from [the gauge fixing] theorem. (...) The system of equations (...) is uniformly elliptic. Now standard techniques apply."

After this, she immediately concludes the main regularity result for Yang–Mills connections, her corollary 1.4 and our theorem 1.40. Another example, from Donaldson and Kronheimer's book [DK97, page 55]:

> "The operator $\mathrm{d} + \mathrm{d}^*$ is elliptic, its kernel decomposes according to degree and so if, as we suppose, $H^1(M)$ is zero, all the 1-forms are orthogonal to the kernel. So elliptic theory gives inequalities
>
> $$\|A\|_{k,2} \leq \text{const.} \left(\|\mathrm{d}^*A\|_{k-1,2} + \|\mathrm{d}A\|_{k-1,2}\right).$$

[8]The mathematical notation in the quotes was changed to agree with the one fixed in this work, so as not to cause confusion.

> ...
>
> While we have been discussing above the case of a closed base manifold, similar ideas can be applied on manifolds with boundary or on complete manifolds, given appropriate boundary or decay conditions."

We will not delve into elliptic theory, as that is beyond the scope of this work, but we leave here a few references. For elliptic theory on compact manifolds, a good introduction which avoids doing too much analysis can be found in [Nic07, chapter 10]; a more succinct overview which delves into the analytical aspects is [Wel80, chapter IV]. For elliptic boundary value problems (i.e., on manifolds with boundary), see [Sch95] or [Hö7, section 20.1].

1.4 Analysis in gauge theory

With all of the analytical machinery defined, we can turn again to gauge theory. In all that follows, let $P \to M$ be a principal G-bundle over a compact Riemannian n-manifold.

Definition 1.27 (Sobolev space of connections)**.** Fix a smooth reference connections $\tilde{A} \in \mathcal{A}(P)$, and note that it gives a covariant derivative $\nabla^{\tilde{A}}$ on the associated vector bundle $\mathrm{ad}(P) \to M$. Then we define the (affine) Sobolev space of connections as

$$\mathcal{A}^{k,p}(P) := \tilde{A} + W^{k,p}(M, T^*M \otimes \mathrm{ad}(P)),$$

modelled after the Sobolev space of sections $W^{k,p}(M, T^*M \otimes \mathrm{ad}(P))$,[9] which is a vector space.

Once again, note that since M is compact, $\mathcal{A}^{k,p}(P)$ will not depend on the choices of Riemannian metric and reference connection, even though the norm on $W^{k,p}(M, T^*M \otimes \mathrm{ad}(P))$ does depend on these choices.

Recalling the definitions given in subsection 1.2.1, in a local trivialization $\Phi = \pi \times \phi : \pi^{-1}(U) \to U \times G$ over some $U \subseteq M$ we have a natural reference connections $\tilde{A} = \phi^{-1}\,\mathrm{d}\phi$ and can represent a connection by $A = \tilde{A} + \alpha$ for $\alpha \in \Omega^1(U, \mathfrak{g})$. The reference connection can be used to define $\mathcal{A}^{k,p}(P|_U)$, and we affirm that the norm on $W^{k,p}(U, T^*U \otimes \mathrm{ad}(P|_U))$ is equal to the norm on $W^{k,p}(U, T^*U \otimes \mathfrak{g})$. Thus we are able to locally define the Sobolev space of connections as

$$\mathcal{A}^{k,p}(U) := W^{k,p}(U, T^*U \otimes \mathfrak{g}), \tag{1.25}$$

such that

$$\mathcal{A}^{k,p}(P|_U) = \phi^{-1}\,\mathrm{d}\phi + \mathcal{A}^{k,p}(U).$$

Definition 1.28 (Sobolev space of gauge transformations)**.** For $kp > n$, we define the Sobolev space of gauge transformations as the Sobolev space of sections of the bundle $\mathrm{Ad}(P) \to M$,[10]

$$\mathcal{G}^{k,p}(P) := W^{k,p}(M, \mathrm{Ad}(P)).$$

[9] See definition 1.14.
[10] See definition 1.25.

Similar to the smooth case, this is naturally isomorphic to the Sobolev space of G-equivariant maps $W^{k,p}(P,G)^G$, and from proposition 1.21, a map $u \in W^{k,p}(P,G)^G$ can be written as $u = s \cdot \exp(\xi)$ for $s \in C^\infty(P,G)^G$ a smooth gauge transformation, and $\xi \in W^{k,p}(P,\mathfrak{g})^G$. On a local trivialization over $U \subseteq M$, a gauge transformation is represented by a map $u : U \to G$, and thus locally we can identify $\mathcal{G}^{k,p}(P|_U)$ with

$$\mathcal{G}^{k,p}(U) := W^{k,p}(U,G), \tag{1.26}$$

and furthermore proposition 1.21 will yield

$$u^{-1}\,\mathrm{d}u \in W^{k-1,p}(U, T^*U \otimes \mathfrak{g}) = \mathcal{A}^{k-1,p}(U)$$

for $u \in \mathcal{G}^{k,p}(U)$.

Just as we defined the Yang–Mills functional to be an L^2-energy, we may generalize this and define an L^q-energy of a connection $A \in \mathcal{A}(P)$, for $1 \leq q < \infty$,

$$\mathcal{E}_q(A) := \int_M |F_A|^q = \|F_A\|_q^q. \tag{1.27}$$

We may extend these functionals to Sobolev spaces of connections, and they will be well behaved.

Lemma 1.29. *When $\frac{n}{2} \leq q < \infty$, $\mathcal{E}_q$ is a continuous functional on $\mathcal{A}^{1,q}(P)$, and for every smooth reference connection $\tilde{A}$ there exists a constant C such that for all $A = \tilde{A} + \alpha \in \mathcal{A}^{1,p}(P)$,*

$$\mathcal{E}_q(A)^{\frac{1}{q}} \leq \mathcal{E}_q(\tilde{A})^{\frac{1}{q}} + 2\,\|\alpha\|_{1,q} + C\,\|\alpha\|_{1,q}^2.$$

Proof. The curvature of $A = \tilde{A} + \alpha$ is

$$\begin{aligned} F_A &= \mathrm{d}(\tilde{A}+\alpha) + \frac{1}{2}[(\tilde{A}+\alpha)\wedge(\tilde{A}+\alpha)] \\ &= \mathrm{d}\tilde{A} + \mathrm{d}\alpha + \frac{1}{2}([\tilde{A}\wedge\tilde{A}] + [\alpha\wedge\alpha] + 2[\tilde{A}\wedge\alpha]) \\ &= F_{\tilde{A}} + \mathrm{d}\alpha + [\tilde{A}\wedge\alpha] + \frac{1}{2}[\alpha\wedge\alpha] \\ &= F_{\tilde{A}} + \mathrm{d}_{\tilde{A}}\alpha + \frac{1}{2}[\alpha\wedge\alpha]. \end{aligned}$$

From (1.4), we know that for any $X, Y \in TM$

$$\mathrm{d}_{\tilde{A}}\alpha(X,Y) = \nabla^{\tilde{A}}\alpha(X,Y) - \nabla^{\tilde{A}}\alpha(Y,X),$$

which implies $|\,\mathrm{d}_{\tilde{A}}\alpha| \leq 2\left|\nabla^{\tilde{A}}\alpha\right|$.

Writing $\alpha = \alpha_i\,\mathrm{d}x_i$ on a local frame $\{\partial x_i\}$,

$$[\alpha\wedge\alpha] = \sum_{i,j}[\alpha_i,\alpha_j]\,\mathrm{d}x_i\,\mathrm{d}x_j = \sum_{i<j}2[\alpha_i,\alpha_j]\,\mathrm{d}x_i\,\mathrm{d}x_j \tag{1.28}$$

and so

$$\left|\frac{1}{2}[\alpha\wedge\alpha]\right|^2 = \sum_{i<j}|[\alpha_i,\alpha_j]|^2 \overset{(*)}{\leq} \sum_{i<j}|\alpha_i|^2|\alpha_j|^2 \leq \sum_{i,j}|\alpha_i|^2|\alpha_j|^2 = |\alpha|^4,$$

where we used (1.21) in $(*)$, and then

$$\left\| \frac{1}{2}[\alpha \wedge \alpha] \right\|_q = \left(\int \left| \frac{1}{2}[\alpha \wedge \alpha] \right|^q \right)^{\frac{1}{q}} \leq \left(\int |\alpha|^{2q} \right)^{\frac{1}{q}} = \|\alpha\|_{2q}^2 .$$

Finally,

$$\begin{aligned} \left| \mathcal{E}_q(A)^{\frac{1}{q}} - \mathcal{E}_q(\tilde{A})^{\frac{1}{q}} \right| &= \left| \left\| F_{\tilde{A}} + \mathrm{d}_{\tilde{A}}\alpha + \frac{1}{2}[\alpha \wedge \alpha] \right\|_q - \|F_{\tilde{A}}\|_q \right| \\ &\leq \left| \|\mathrm{d}_{\tilde{A}}\alpha\|_q + \left\| \frac{1}{2}[\alpha \wedge \alpha] \right\|_q \right| \\ &\leq 2 \left\| \nabla^{\tilde{A}}\alpha \right\|_q + \|\alpha\|_{2q}^2 \\ &\leq 2 \|\alpha\|_{1,q} + C_W \|\alpha\|_{1,q}^2 , \end{aligned}$$

where the last inequality follows because of the embedding $W^{1,q} \hookrightarrow L^{2p}$ which holds with the assumption $q \geq \frac{n}{2}$. □

Locally, the energy of a connection $A \in \mathcal{A}^{1,q}(U)$ is denoted the same way and given by

$$\mathcal{E}_q(A) = \|F_A\|_{q(U)}^q ,$$

where $F_A \in L^q(U, \Lambda^2 T^*U \otimes \mathfrak{g})$ is the local representative of the curvature. Then the estimate in the previous lemma becomes

$$\mathcal{E}_q(A)^{\frac{1}{q}} = \|F_A\|_q \leq 2 \|A\|_{1,q} + C \|A\|_{1,q}^2 . \tag{1.29}$$

We already know that the group of gauge transformations is indeed a group; we now prove that the group operations are continuous, and therefore the $\mathcal{G}^{k,p}$ are topological groups, for $kp > n$.

Lemma 1.30. *Let $k \in \mathbb{N}$ and $1 \leq p < \infty$ be such that $kp > n$; then group multiplication and inversion are continuous maps on $\mathcal{G}^{k,p}(P)$.*

Proof. Let $(U_\alpha, \tau_\alpha)_{\alpha \in A}$ be a bundle atlas of $\mathrm{Ad}(P) \to M$. By definition, a gauge transformation $u \in \mathcal{G}^{k,p}(P)$ is a continuous section of this bundle such that for all $\alpha \in A$, $u_\alpha := \tau_\alpha \circ u : U_\alpha \to G$ is in $W^{k,p}(U_\alpha, G)$. If $(V_\beta, \phi_\beta)_{\beta \in B}$ is an atlas of G, then this means that $\phi_\beta \circ u_\alpha \in W^{k,p}(u_\alpha^{-1}(V_\beta), \mathbb{R}^\ell)$, for all $\beta \in B$ and ℓ the dimension of G. We will prove the continuity of the inversion map, and the multiplication will be analogous.

Let $i : G \to G$ be the inversion map, $i(g) = g^{-1}$. Then u^{-1} can be defined by $(u^{-1})_\alpha = i \circ u_\alpha$ for all $\alpha \in A$. We would like to show that $(u^{-1})_\alpha \in W^{k,p}(U_\alpha, G)$. For $\gamma \in B$,

$$\phi_\gamma \circ (u^{-1})_\alpha = \phi_\gamma \circ i \circ u_\alpha = (\phi_\gamma \circ i \circ \phi_\beta^{-1}) \circ (\phi_\beta \circ u_\alpha).$$

Now we can use Lemma 1.23 with $f = \phi_\gamma \circ i \circ \phi_\beta^{-1} \in C^\infty(\phi_\beta(V_\beta) \subseteq \mathbb{R}^\ell, \mathbb{R}^\ell)$ to conclude that $\phi_\gamma \circ (u^{-1})_\alpha \in W^{k,p}((u^{-1})_\alpha^{-1}(V_\gamma), \mathbb{R}^\ell)$. Since $\alpha \in A$ and $\gamma \in B$ were arbitrary, this proves $u^{-1} \in \mathcal{G}^{k,p}(P)$. □

Moreover, the gauge action on the appropriate Sobolev space of connections is continuous.

Lemma 1.31. *Let $k \in \mathbb{N}$ and $1 \le p < \infty$ be such that $kp > n$; then the gauge action*

$$\mathcal{G}^{k,p}(P) \times \mathcal{A}^{k-1,p}(P) \to \mathcal{A}^{k-1,p}(P)$$
$$(u, A) \mapsto u^*A$$

is a continuous map. Moreover, for every trivializing neighbourhood $U \subseteq M$ there is a constant C such that for $u \in \mathcal{G}^{k,p}(U)$ and $A \in \mathcal{A}^{k-1,p}(U)$ the following holds:

$$\|u^*A\|_{k-1,p} \le \left\|u^{-1}\,\mathrm{d}u\right\|_{k-1,p} + C\,\|A\|_{k-1,p}\left(1 + \left\|u^{-1}\,\mathrm{d}u\right\|_{k-2,2p}\right)^{k-1}.$$

Proof. First, note that it suffices to prove continuity of the action on an arbitrary trivializing neighbourhood, say $U \subseteq M$. We wish to prove that for two sequences $(u_i) \subseteq \mathcal{G}^{k,p}(U)$ and $(A_i) \subseteq \mathcal{A}^{k-1,p}(U)$ converging to $u \in \mathcal{G}^{k,p}(U)$ and $A \in \mathcal{A}^{k-1,p}(U)$, respectively,

$$(u_i)^{-1}A_iu_i + (u_i)^{-1}\,\mathrm{d}u_i = u_i^*A_i \longrightarrow u^*A.$$

We treat the cases $k = 1$ and $k \ge 2$ separately. In both cases, note that by definition (lemma 1.22), (u_i) converges to u in $C^0(U, G)$ and $(u_i)^{-1}\,\mathrm{d}u_i \overset{(k-1,p)}{\longrightarrow} u^{-1}\,\mathrm{d}u$.

For $k = 1$, the inequality is simply

$$\|u^*A\|_{p(U)} \le \left\|u^{-1}\,\mathrm{d}u\right\|_{p(U)} + \left\|u^{-1}Au\right\|_{p(U)},$$

which is just the triangle inequality for the norm of the gauge action. Then when looking at $\|u^*A - u_i^*A_i\|_{p(U)}$, we have shown that the first term converges, and the second term

$$\begin{aligned}\left\|u^{-1}Au - u_i^{-1}A_iu_i\right\|_p &\le \left\|u^{-1}Au - u^{-1}A_iu\right\|_p + \left\|u^{-1}A_iu - u_i^{-1}A_iu_i\right\|_p \\ &\le \|A - A_i\|_p + \|\mathrm{Ad}_u - \mathrm{Ad}_{u_i}\|\,\|A_i\|_p\end{aligned}$$

will also converge because of the C^0-convergence $u_i \to u$. For the second inequality, we used the invariante of the metric on $\mathfrak{g}$ under conjugation and the fact that Ad_g is a bounded linear operator on $\mathfrak{g}$.

For $k \ge 2$, once again writing

$$\|u^*A\|_{k-1,p(U)} \le \left\|u^{-1}\,\mathrm{d}u\right\|_{k-1,p(U)} + \left\|u^{-1}Au\right\|_{k-1,p(U)},$$

the first term is known to converge, and for the second term we use lemma 1.32 below with $\tau = A, \tau_i = A_i$ and $\ell = k - 1$ to get convergence and an estimate which yields the desired inequality. □

Lemma 1.32. *Let $U \subseteq M$ be a trivializing neighbourhood of $P \to M$, $0 \le \ell \le k$ be integers and $1 \le p < \infty$ be such that $kp > n$ and $p \ge \frac{n}{2}$. Then the following holds.*

Let $(u_i) \subseteq \mathcal{G}^{k,p}(U)$ and $(\tau_i) \subseteq W^{\ell,p}(\Lambda^\ell T^*U \otimes \mathfrak{g})$ be sequences that converge to $u \in \mathcal{G}^{k,p}(U)$ and $\tau \in W^{\ell,p}(\Lambda^\ell T^*U \otimes \mathfrak{g})$, respectively and in the appropriate topologies. Then

$$(u_i)^{-1}\tau_iu_i \overset{\ell,p}{\longrightarrow} u^{-1}\tau u.$$

Moreover, there exists a constant C such that

$$\left\|u^{-1}\tau u\right\|_{\ell,p} \le C\,\|\tau\|_{\ell,p}\left(1 + \left\|u^{-1}\,\mathrm{d}u\right\|_{\ell-1,2p}\right)^{\ell}.$$

Proof. The proof will go through by induction on ℓ.

For $\ell = 0$, the estimate is provided by the invariance of the metric on $\mathfrak{g}$ under conjugation and is simply

$$\left\|u^{-1}\tau u\right\|_p = \|\tau\|_p .$$

Convergence follows from the C^0 convergence of the u_i (lemma 1.22).

Assume the lemma to hold for $\ell - 1 \geq 0$. The case $\ell = 0$ already provides convergence for the L^p term in the norm and it remains to show that the derivative, which we write using (1.24) and $\mathrm{d}u^{-1} \cdot u = -u^{-1}\,\mathrm{d}u$,

$$\nabla(u_i^{-1}\tau_i u_i) = u_i^{-1}\nabla\tau_i u_i - [u_i^{-1}\,\mathrm{d}u_i, u_i^{-1}\tau_i u_i],$$

converges to $\nabla(u^{-1}\tau u)$ in the $W^{\ell-1,p}$-norm. The first term in the derivative can be seen to converge from the induction hypothesis, using the lemma for $\nabla\tau$ and $\ell - 1$. For the second term, the Lie bracket, note the calculation below,

$$\begin{aligned}
&\|[u_i^{-1}\,\mathrm{d}u_i, u_i^{-1}\tau_i u_i] - [u^{-1}\,\mathrm{d}u, u^{-1}\tau u]\|_{\ell-1,p} \\
&\quad = \left\|[u_i^{-1}\,\mathrm{d}u_i - u^{-1}\,\mathrm{d}u, u_i^{-1}\tau_i u_i] + [u^{-1}\,\mathrm{d}u, u_i^{-1}\tau_i u_i - u^{-1}\tau u]\right\|_{\ell-1,p} \\
&\quad \leq \left\|[u_i^{-1}\,\mathrm{d}u_i - u^{-1}\,\mathrm{d}u, u_i^{-1}\tau_i u_i]\right\|_{\ell-1,p} + \left\|[u^{-1}\,\mathrm{d}u, u_i^{-1}\tau_i u_i - u^{-1}\tau u]\right\|_{\ell-1,p} \\
&\quad \overset{(1)}{\leq} \left\|\left|u_i^{-1}\,\mathrm{d}u_i - u^{-1}\,\mathrm{d}u\right| \cdot \left|u_i^{-1}\tau_i u_i\right|\right\|_{\ell-1,p} + \left\|\left|u^{-1}\,\mathrm{d}u\right| \cdot \left|u_i^{-1}\tau_i u_i - u^{-1}\tau u\right|\right\|_{\ell-1,p} \\
&\quad \overset{(2)}{\leq} \left\|u_i^{-1}\,\mathrm{d}u_i - u^{-1}\,\mathrm{d}u\right\|_{\ell-1,2p} \left\|u_i^{-1}\tau_i u_i\right\|_{\ell-1,2p} \\
&\qquad\qquad + \left\|u^{-1}\,\mathrm{d}u\right\|_{\ell-1,2p} \left\|u_i^{-1}\tau_i u_i - u^{-1}\tau u\right\|_{\ell-1,2p} \\
&\quad \overset{(3)}{\leq} C_W \underbrace{\left\|u_i^{-1}\,\mathrm{d}u_i - u^{-1}\,\mathrm{d}u\right\|_{\ell,p}}_{\overset{(4)}{\longrightarrow} 0} \underbrace{\left\|u_i^{-1}\tau_i u_i\right\|_{\ell-1,2p}}_{\overset{(5)}{\longrightarrow}\|u^{-1}\tau u\|_{\ell-1,2p}} \\
&\qquad\qquad + C'_W \underbrace{\left\|u^{-1}\,\mathrm{d}u\right\|_{\ell,p}}_{\text{const.}} \underbrace{\left\|u_i^{-1}\tau_i u_i - u^{-1}\tau u\right\|_{\ell-1,2p}}_{\overset{(6)}{\longrightarrow} 0} \longrightarrow 0,
\end{aligned}$$

where (1) follows from (1.20); (2) is lemma B.8 with $r = s = 2p$; (3) is the Sobolev embedding $W^{\ell,p} \hookrightarrow W^{\ell-1,2p}$ which is valid for $p \geq \frac{n}{2}$; (4) follows from Lemma 1.22, noting that $\ell \leq k - 1$; (5) and (6) follow taking the lemma with $\ell - 1$ (which is valid by induction hypothesis) but with $(k-1, 2p)$ instead of (k, p), since by the same embedding $W^{\ell,p} \hookrightarrow W^{\ell-1,2p}$, the original sequence (τ_i) also converges in $W^{\ell-1,2p}$-norm and (u_i) also converges in $\mathcal{G}^{k-1,2p}$. This proves convergence for ℓ from $\ell - 1$.

For the estimate, assume it valid for $\ell - 1$. Then, denoting by C all constants which do not come from a Sobolev estimate, and by C_W all of the ones that do,

$$\begin{aligned}
\left\|u^{-1}\tau u\right\|_{\ell,p} &\leq \left\|u^{-1}\tau u\right\|_p + \left\|\nabla(u^{-1}\tau u)\right\|_{\ell-1,p} \\
&\overset{(1)}{\leq} \|\tau\|_p + \left\|u^{-1}\nabla\tau u\right\|_{\ell-1,p} + \left\|u^{-1}\,\mathrm{d}u\right\|_{\ell-1,2p} \left\|u^{-1}\tau u\right\|_{\ell-1,2p} \\
&\overset{(2)}{\leq} \|\tau\|_p + C\,\|\nabla\tau\|_{\ell-1,p} \left(1 + \left\|u^{-1}\,\mathrm{d}u\right\|_{\ell-2,2p}\right)^{\ell-1}
\end{aligned}$$

$$
\begin{aligned}
&\quad + \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-1,2p} C \left\|\tau\right\|_{\ell-1,2p} \left(1 + \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-2,4p}\right)^{\ell-1} \\
&\leq C \left\|\tau\right\|_{\ell,p} \left(1 + \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-1,2p}\right)^{\ell-1} \\
&\quad + \left(1 + \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-1,2p}\right) C_W \left\|\tau\right\|_{\ell,p} \left(1 + C_W \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-1,2p}\right)^{\ell-1} \\
&\leq C \left\|\tau\right\|_{\ell,p} \left(1 + \left\| u^{-1}\,\mathrm{d}u \right\|_{\ell-1,2p}\right)^{\ell},
\end{aligned}
$$

where (1) follows from arguments similar to those already used for the convergence above, (2) is the induction hypothesis, and the Sobolev estimates for $W^{\ell,p} \hookrightarrow W^{\ell-1,2p}$ and $W^{\ell-1,2p} \hookrightarrow W^{\ell-2,4p}$ hold due to $p \geq \frac{n}{2}$. □

The following results will be critical for proving the weak and strong compactness theorems.

Lemma 1.33. *Let $k \in \mathbb{N}$ and $1 \leq p < \infty$ be such that $kp > n$ and $p > \frac{n}{2}$. Let $(A^i)_{i\in\mathbb{N}} \subseteq \mathcal{A}^{k-1,p}(P)$ and $(u^i)_{i\in\mathbb{N}} \subseteq \mathcal{G}^{k,p}(P)$ be two sequences such that both $\|A^i\|_{k-1,p}$ and $\|u^{i*}A^i\|_{k-1,p}$ are uniformly bounded. Then the following holds:*

(i) For every trivialization over some domain $U_\alpha \subseteq M$, there exists a uniform bound on $\|(u^i_\alpha)^{-1}\,\mathrm{d}u^i_\alpha\|_{k-1,p(U_\alpha)}$.

(ii) There exists a subsequence of (u^i) that converges in the C^0-topology to some limit in $\mathcal{G}^{k,p}(P)$.

Proof. See [Weh04, lemma A.8]. □

Lemma 1.34. *Let $1 < p < \infty$ and $k \in \mathbb{N}_0$ such that $kp > n$ and $p \geq \frac{n}{2}$. Let $A, A' \in \mathcal{A}^{k,p}(P)$. If there is a continuous gauge transformation u such that $A' = u^*A$, then $u \in \mathcal{G}^{k+1,p}(P)$.*

Proof. Look at the local representatives on some chart $U \subseteq M$, such that $A, A' \in \mathcal{A}^{k,p}(U)$ and $u \in C^0(U, G)$, then write $A' = u^{-1}Au + u^{-1}\,\mathrm{d}u$ (the gauge action looks the same locally, see lemma 1.11). In case $k = 0$, it suffices to look at

$$
\left\| u^{-1}\,\mathrm{d}u \right\|_p \leq \left\| A' - u^{-1}Au \right\|_p \leq \|A'\|_p + \|A\|_p ,
$$

where in the second inequality we used that the inner product on $\mathfrak{g}$ is Ad-invariant. Then $u^{-1}\,\mathrm{d}u \in L^p(U, T^*U \otimes \mathfrak{g})$, and so by proposition 1.21, $u \in \mathcal{G}^{1,p}(P)$.

For $k > 1$, we will use the estimate in lemma 1.32 above. First, note that for all $j \leq k$,

$$
\left\| u^{-1}\,\mathrm{d}u \right\|_{2^j p} \leq \|A'\|_{2^j p} + \|A\|_{2^j p} \leq \|A'\|_{k,p} + \|A\|_{k,p}
$$

since $kp > n > \frac{2^j-1}{2^j}n$. Then,

$$
\left\| u^{-1}\,\mathrm{d}u \right\|_{k,p} \leq \|A'\|_{k,p} + \left\| u^{-1}Au \right\|_{k,p} ,
$$

and from lemma 1.32, $\|u^{-1}Au\|_{k,p}$ is bounded by $\|A\|_{k,p}$ and $\|u^{-1}\,\mathrm{d}u\|_{k-1,2p}$. The norm $\|A'\|_{k,p}$ is finite by assumption, and we bound the second term as follows: we iterate the estimate above, using the embeddings

$$
W^{k,p} \hookrightarrow W^{k-1,2p} \hookrightarrow W^{k-2,4p} \hookrightarrow \cdots \hookrightarrow W^{k-j,2^j p}
$$

to bound the $\|A'\|$ term, which hold since $p \geq \frac{n}{2} \geq \frac{n}{2^j}$ for $j \geq 1$, and using lemma 1.32 to bound the $\|u^{-1}Au\|$ term. This process will finally end with $\|u^{-1}\,\mathrm{d}u\|_{k,p}$ bounded by a finite amount of terms $\|A'\|_{k,p}$ and $\|u^{-1}\,\mathrm{d}u\|_{2^k p}$, which we showed above is finite. Thus, $u \in \mathcal{G}^{k+1,p}(P)$. □

That is, bounds on connections give bounds on gauge transformations relating them.

Non-smooth Yang–Mills connections

In the smooth case, we defined the Yang–Mills functional (1.16),

$$\mathcal{YM}(A) = \int_M |F_A|^2\,,$$

showed that critical points satisfy the weak Yang–Mills equation, and noted that smooth solutions to the weak equation also satisfy the strong Yang–Mills equation with boundary condition. It is possible to extend the functional to connections with less regularity. For $\mathcal{YM}$ to be well-defined, we need F_A of class L^2. Let $A \in \mathcal{A}^{1,p}(P)$, and consider the curvature:

$$F_A = \mathrm{d}A + \frac{1}{2}[A \wedge A].$$

It is clear that $\mathrm{d}A \in L^p(P, \mathfrak{g})$, yielding the necessary hypothesis $p \geq 2$, which at this point we are used to; as for the second term, note that $\|[A \wedge A]\|_2 \leq \|A\|_4^2$, hence we need the Sobolev embedding $W^{1,p} \hookrightarrow L^4$, which holds when $p \geq \frac{4n}{4+n}$.

That said, we can ask even less of the connection if we only want it to satisfy the weak Yang–Mills equation 1.30.

Definition 1.35. For $1 \leq p < \infty$ such that $p > \frac{n}{2}$, and in case $n = 2$ assume in addition $p \geq \frac{4}{3}$, a connection $A \in \mathcal{A}^{1,p}(P)$ is called **weak Yang–Mills** if it satisfies

$$\int_M \langle F_A,\, \mathrm{d}_A\beta\rangle = 0, \quad \forall \beta \in \Omega^1(M, \mathrm{ad}(P)). \tag{1.30}$$

The conditions in the definition guarantee that $\langle F_A,\, \mathrm{d}_A\beta\rangle \in L^1(M)$. With $p \geq \frac{n}{2}$, the embedding $W^{1,p} \hookrightarrow L^{2p}$ guarantees that $[A \wedge A]$ is in L^p, and therefore so is the curvature. Then we need $\mathrm{d}_A\beta$ of class L^{p*}, for

$$\frac{1}{p*} := 1 - \frac{1}{p}.$$

Since β is smooth, we need $W^{1,p} \hookrightarrow L^{p*}$ for $A \in \mathcal{A}^{1,p}$, and the condition for this is $p \geq \frac{2n}{n+1}$; for $n = 1$ this is met due to $p \geq 1$, for $n \geq 3$ this is met due to $p \geq \frac{n}{2} \geq \frac{2n}{n+1}$, and then for $n = 2$ we need to assume $p \geq \frac{4}{3}$.

As we hinted to, the Yang–Mills functional might not be defined or finite for a weak Yang–Mills connection: we do not explicitly ask for $p \geq 2$, and while $p \geq \frac{n}{4}$ guarantees this for $n \geq 4$, this might fail if $n \leq 3$; moreover, the condition $p \geq \frac{4n}{4+n}$ fails for $n = 3$.

The following lemma shows that (1.30) is preserved under gauge transformations.

Lemma 1.36. *Let $A \in A^{1,p}(P)$ be a weak Yang–Mills connection, and fix a compact set $K \subseteq M$. Then for every gauge transformation $u \in \mathcal{G}^{2,p}(P|_K)$, $u^*A \in \mathcal{A}^{1,p}(P|_K)$ is also weak Yang–Mills.*

Proof. First, we show that (1.30) still holds if we the test forms are not smooth but instead $\beta \in W^{2,p}(M, \mathrm{ad}(P))$. Since $F_A \in L^p$, we check that $\mathrm{d}_A\beta \in L^{p*}$. From the local formula

$$(\mathrm{d}_A\beta)_\alpha = \mathrm{d}\beta_\alpha + [A_\alpha \wedge \beta_\alpha],$$

the first term is in L^{p*} since $\mathrm{d}\beta_\alpha \in W^{1,p}$ and $W^{1,p} \hookrightarrow L^{p*}$, and then

$$\|[A_\alpha \wedge \beta_\alpha]\|_{p*} \leq c\, \||A_\alpha|\, |\beta_\alpha|\|_{p*} \leq c\, \|A_\alpha\|_{p*}\, \|\beta_\alpha\|_\infty \leq C_W\, \|A_\alpha\|_{1,p}\, \|\beta_\alpha\|_{2,p}\,,$$

where $W^{2,p} \hookrightarrow C^0$ holds for $p > \frac{n}{2}$.

Now, let $K \subseteq M$ be compact and let $u \in \mathcal{G}^{2,p}(P|_K)$. For a smooth test 1-form β with support in K, we define $\tilde{\beta} := u\beta u^{-1}$ and extend it to 0 outside K, such that $\tilde{\beta} \in W^{2,p}(K, \mathrm{ad}(P|_K))$.[11] We have shown that (1.30) holds for such $\tilde{\beta}$, and thus

$$\int_M \langle F_{u^*A},\, \mathrm{d}_{u^*A}\beta\rangle = \int_M \langle u^{-1}F_A u, u^{-1}\,\mathrm{d}_A\tilde{\beta}u\rangle = \int_M \langle F_A,\, \mathrm{d}_A\tilde{\beta}\rangle = 0.$$

This follows from the calculation below, where we use (1.23),

$$\begin{aligned}\left(\mathrm{d}_A(u\beta u^{-1})\right)_\alpha &= \mathrm{d}(\mathrm{Ad}_{u_\alpha}\,\beta_\alpha) + [A_\alpha \wedge u_\alpha\beta_\alpha u_\alpha^{-1}]\\ &= [\,\mathrm{d}u_\alpha \cdot u_\alpha^{-1} \wedge u_\alpha\beta_\alpha u_\alpha^{-1}] + \mathrm{Ad}_{u_\alpha}\,\mathrm{d}\beta_\alpha + u_\alpha^{-1}[u_\alpha A_\alpha u_\alpha^{-1} \wedge \beta_\alpha]u_\alpha\\ &= u_\alpha\,\mathrm{d}\beta_\alpha u_\alpha^{-1} + u_\alpha[u_\alpha^{-1}\,\mathrm{d}u_\alpha \wedge \beta_\alpha]u_\alpha^{-1} + u_\alpha^{-1}[u_\alpha A_\alpha u_\alpha^{-1} \wedge \beta_\alpha]u_\alpha\\ &= \left(u(\,\mathrm{d}_{u^*A}\beta)u^{-1}\right)_\alpha.\end{aligned}$$

□

Next, we show that for sufficient regularity, the weak and strong Yang–Mills equations are equivalent, and this implies the result in the smooth case.

Lemma 1.37. *Let $1 \leq p < \infty$ be such that $p \geq \frac{2n}{n+2}$ and let $k \in \mathbb{N}$. Fix a connection $A \in \mathcal{A}^{1,p}(P)$, and two equivariant forms $\omega \in W^{1,p}(M, \Lambda^k T^*M \otimes \mathrm{ad}(P))$ and $\gamma \in L^p(M, \Lambda^{k-1}T^*M \otimes \mathrm{ad}(P))$. Then the following are equivalent:*

(i) For all smooth $\eta \in \Omega^1(M, \mathrm{ad}(P))$,

$$\int_M \langle \omega,\, \mathrm{d}_A\eta\rangle = \int_M \langle \gamma, \eta\rangle.$$

(ii)

$$\begin{cases} \mathrm{d}_A^*\omega = \gamma,\\ *\omega|_{\partial M} = 0.\end{cases}$$

Proof. Consider the computation below,

$$\begin{aligned}\int_M \langle \omega,\, \mathrm{d}_A\eta\rangle &= \int_M \langle \omega,\, \mathrm{d}\eta\rangle + \int_M \langle \omega, [A \wedge \eta]\rangle\\ &= \int_M \langle \eta,\, \mathrm{d}^*\omega\rangle + \int_M \mathrm{d}\langle \eta \wedge *\omega\rangle \pm \int_M \langle *\omega \wedge [A \wedge \eta]\rangle\end{aligned}$$

[11] Indeed, see lemmata 1.30 and 1.32.

$$
\begin{aligned}
&= \int_M \langle \mathrm{d}^*\omega, \eta\rangle + \int_{\partial M} \langle \eta \wedge *\omega\rangle \pm \int_M \langle [*\omega \wedge A] \wedge \eta\rangle \\
&= \int_M \langle \mathrm{d}^*\omega, \eta\rangle - (-1)^{(n-k)(k-1)} \int_M \langle *[A \wedge *\omega], \eta\rangle + \int_{\partial M} \langle \eta \wedge *\omega\rangle \\
&= \int_M \langle \mathrm{d}_A^*\omega, \eta\rangle + \int_{\partial M} \langle \eta \wedge *\omega\rangle
\end{aligned}
$$

where we have used Stokes's theorem and property (1.20) of the inner product on $\mathfrak{g}$-valued forms. We did not keep track of the sign in front of the $\int\langle\omega, [A \wedge \eta]\rangle$ term when writing down the calculation because it would become too cumbersome, but it can be checked that we obtain the correct sign. Moreover, the hypothesis $p \geq \frac{2n}{n+2}$ guarantees $W^{1,p} \hookrightarrow L^2$, which is needed for the integral of $\langle *\omega \wedge [A \wedge \eta]\rangle$ to be well defined.

We use this identity to prove the lemma. If we assume (ii), then it directly gives (i). Now assume (i): testing with arbitrary η that vanish at the boundary we get $\mathrm{d}_A^*\omega = \gamma$, and this further implies that the boundary term is zero for all η, which proves $*\omega|_{\partial M} = 0$. □

The weak Yang–Mills equation is well behaved under limits, weak and strong.

Lemma 1.38. *Let $1 < p < \infty$ such that $p > \frac{n}{2}$, and in case $n = 2$ let also $p \geq \frac{4}{3}$.*

(i) If a sequence of weak Yang–Mills connections in $\mathcal{A}^{1,p}(P)$ converges strongly in the $W^{1,p}$ topology, the limit is also weak Yang–Mills.

(ii) If in case $n = 2$ there is strict inequality $p > \frac{4}{3}$, for a sequence of weak Yang–Mills connections in $\mathcal{A}^{1,p}(P)$ with L^p-bound on curvature which converges weakly in the $W^{1,p}$ topology, the limit connection is also weak Yang–Mills.

Proof. We will prove (i) and (ii) at the same time, commenting where the extra assumptions for (ii) are needed.[12] Remember that strong convergence implies weak convergence.

Suppose $(A^i) \subseteq \mathcal{A}^{1,p}(P)$ is a sequence which converges weakly to $A \in \mathcal{A}^{1,p}(P)$ such that each A^i is weak Yang–Mills. Note that if the convergence is strong, then there is a uniform bound on $\|F_{A^i}\|_p$ *a priori*; if the convergence is weak, we must assume the uniform bound. Then, using the Cauchy-Schwartz ineqality and Hölder's inequality for $1 = \frac{1}{p} + \frac{1}{p*}$, we prove that A is also weak Yang–Mills, calculating for any $\beta \in \Omega^1(M, \mathrm{ad}(P))$:

$$
\begin{aligned}
\int_M \langle F_A, \mathrm{d}_A\beta\rangle &= \int_M \langle F_A, \mathrm{d}_A\beta\rangle - \int_M \langle F_{A^i}, \mathrm{d}_{A^i}\beta\rangle \\
&= \int_M \langle F_A - F_{A^i}, \mathrm{d}_A\beta\rangle + \int_M \langle F_{A^i}, \mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta\rangle \\
&\leq \int_M \langle F_A - F_{A^i}, \mathrm{d}_A\beta\rangle + \int_M |\langle F_{A^i}, \mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta\rangle| \\
&\leq \int_M \langle F_A - F_{A^i}, \mathrm{d}_A\beta\rangle + \int_M |F_{A^i}| \cdot |\mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta| \\
&\leq \underbrace{\int_M \langle F_A - F_{A^i}, \mathrm{d}_A\beta\rangle}_{\longrightarrow 0} + \underbrace{\|F_{A^i}\|_p}_{\text{bounded}} \underbrace{\|\mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta\|_{p*}}_{\longrightarrow 0} \longrightarrow 0.
\end{aligned}
$$

[12] Some steps would be more straightforward if we only wished to prove the case of strong convergence, however we do not comment on those.

The first limit is the weak L^p-convergence of F_{A^i}, and we check weak convergence of local representatives on all bundle charts U_α, $(F_{A^i})_\alpha = \mathrm{d}A^i_\alpha + [A^i_\alpha \wedge A^i_\alpha]$. For the weak convergence of the $\mathrm{d}A^i_\alpha$ term, test with any $\beta \in \Omega^2(U_\alpha, \mathfrak{g})$ that vanishes on ∂U_α,

$$\int_{U_\alpha} \langle \mathrm{d}A^i_\alpha, \beta\rangle = \int_{U_\alpha} \langle A^i_\alpha, \mathrm{d}^*\beta\rangle \xrightarrow{i\to\infty} \int_{U_\alpha} \langle A_\alpha, \mathrm{d}^*\beta\rangle = \int_{U_\alpha} \langle \mathrm{d}A_\alpha, \beta\rangle$$

then note that the limit holds for all $\beta \in \Omega^2(U_\alpha, \mathfrak{g})$ since these can be L^{p*}-approximated by such test forms which vanish on the boundary.[13] For the $[A^i_\alpha \wedge A^i_\alpha]$ term, use Hölder's inequality for $\frac{1}{p} = \frac{1}{2p} + \frac{1}{2p}$,

$$\left\|[A_\alpha \wedge A_\alpha] - [A^i_\alpha \wedge A^i_\alpha]\right\|_p \le \left\|A_\alpha - A^i_\alpha\right\|_{2p} \|A_\alpha\|_{2p} + \left\|A^i_\alpha\right\|_{2p} \left\|A_\alpha - A^i_\alpha\right\|_{2p} \longrightarrow 0,$$

and this converges strongly because of the compact Sobolev embedding $W^{1,p} \hookrightarrow L^{2p}$ guaranteed by $p > \frac{n}{2}$.

As for the second limit, we again look at local representatives,

$$(\mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta)_\alpha = [(A_\alpha - A^i_\alpha) \wedge \beta_\alpha],$$

and obtain convergence in all bundle charts observing the following:

$$\|(\mathrm{d}_A\beta - \mathrm{d}_{A^i}\beta)_\alpha\|_{p*} \le \left\|A_\alpha - A^i_\alpha\right\|_{p*} \cdot \|\beta_\alpha\|_\infty \longrightarrow 0.$$

This limit follows from the continuous embedding $W^{1,p} \hookrightarrow L^{p*}$ if the original sequence converged strongly in $\mathcal{A}^{1,p}$, however if we only had weak convergence the additional hypothesis $p > \frac{4}{3}$ is needed to ensure that the embedding is compact and therefore there is strong convergence in L^{p*}. □

Finally, we state without proof the main result we will need on the regularity of Yang–Mills connections.

Proposition 1.39. *Let M be a compact n-manifold, and let $\tilde{A} \in \mathcal{A}(P)$ be a smooth reference connection. Let $1 < p < \infty$ and $k \in \mathbb{N}$ be such that either $kp > n$, or if $k = 1$ then $\frac{n}{2} < p < n$, and in either case if $n = 2$, then $p \ge \frac{4}{3}$. Moreover, let $q := p$ in the first case or $q := \frac{np}{2n-p}$ in the second case. Then there exists a constant C with the following significance:*

Let $A = \tilde{A} + \alpha \in \mathcal{A}^{k,p}(P)$ be a connection that satisfies

$$\begin{cases} \mathrm{d}^*_{\tilde{A}}\alpha &= 0, \\ *\alpha|_{\partial M} &= 0, \end{cases}$$

and for all smooth $\beta \in \Omega^1(M, \mathrm{ad}(P))$

$$\int_M \langle F_A, \mathrm{d}_A\beta\rangle = 0.$$

Then $A \in \mathcal{A}^{k+1,q}(P)$ and

$$\|\alpha\|_{k+1,q} \le C\left(1 + \|\alpha\|_{k,p} + \|\alpha\|^3_{k,p}\right).$$

Moreover, the constant C can be chosen such that it depends $W^{k+1,\infty}$-continuously on the metric.

[13] Writing β_i for the test forms vanishing on ∂U_α approximating β, using Hölder's inequality and noting that the compact embedding $W^{1,p} \hookrightarrow L^p$ gives a uniform bound on $\left\|A^i_\alpha\right\|_p$, the boundary term coming from Stokes's theorem will be bounded by $\left\|A^i_\alpha\right\|_p \|\beta - \beta_i\|_{p*} \to 0$.

Proof. See [Weh04, corollary 9.6]. □

If the first pair of equations in this proposition seems arbitrary now, the next chapter should convince you otherwise. Proposition 1.39 along with theorem 2.2 imply the following result on the regularity of Yang–Mills connections. We shall not use it in this work, however we thought it would fit in nicely as the conclusion to this first chapter. This is also essentially corollary 1.4 in Uhlenbeck's original paper [Uhl82].

Theorem 1.40. *Let M be a compact n-manifold, and let $1 < p < \infty$ be such that $p > \frac{n}{2}$, and in case $n = 2$ assume further that $p \geq \frac{4}{3}$. Then for every weak Yang–Mills connection $A \in \mathcal{A}^{1,p}(P)$ there is a gauge transformation $u \in \mathcal{G}^{2,p}(P)$ such that u^*A is smooth.*

Proof. See [Weh04, theorem 9.4]. □

Chapter 2

The gauge fixing lemma

The main theorem in this chapter is very easy to state in simple terms: for connections with bounded curvature, there is a good local gauge. In the first section we motivate the idea behind gauge fixing and show an example of a good gauge in the simple case when G is abelian and M is a closed manifold; next, we give the main definition 2.1 of what a "good" gauge is in our context, and explain how the main theorem 2.2 follows from a similar result in Euclidean space. Then, in the second section, we prove the result on a chart.

The gauge fixing lemma will be essential for the proofs of the compactness theorems in the next chapter. It is theorem 2.1 in [Uhl82], and we mainly follow the expositions in [DK97, chapter 2] and [Weh04, chapter 6].

2.1 Motivation and main result

Flat connections, Coulomb gauge and Hodge theory

In a problem that has a gauge redundancy, solutions come in families (gauge orbits) and it is often necessary, or at least convenient, to choose representatives; this process is called **gauge fixing**. A choice of gauge might be good simply because it simplifies calculations, as in the classical choice of Coulomb gauge $\nabla \cdot A = 0$ for the vector potential in electromagnetism. More abstractly, one might also be after minimizing norms, or making a system of equations elliptic (and therefore more tractable).

If a connection is flat and the base manifold is simply connected, then we may choose a trivialization such that the connection matrix is identically zero, that is, $\nabla^A = \mathrm{d}$ is the canonical (or product) connection on a trivial (product) vector bundle. This may be done with parallel transport, by first choosing a frame over a point p and then extending this frame in each direction on the manifold. Similarly, on a holomorphic vector bundle one may wish to choose local trivializations such that the partial connection $\overline{\partial}_\alpha = \overline{\partial} + \alpha$ has $\alpha = 0$, that is, choose a gauge such that the flat-in-the-(0,1)-direction connection is the canonical one.[1]

Hence, if a connection is somehow flat, one obtains local gauges such that the connection matrix is zero. If instead of flatness there is small curvature, can one find a correspondingly small connection matrix? In the case of the abelian $U(1)$ gauge group, we may use Hodge theory. Assume M is a simply connected closed base manifold for the time being, and let A be a connection on a trivial $U(1)$-bundle over M; recall that the Lie algebra of $U(1)$ is $i\mathbb{R}$.

[1]For the detailed statements and proofs of these results, see [DK97, section 2.2].

The curvature of the connection is simply $F_A = \mathrm{d}A$, and any change $A \mapsto A + i\,\mathrm{d}f$ for a smooth real valued function leaves the curvature unchanged. If $u : M \to U(1)$ is a gauge transformation, then it may be written as $u(x) = \exp(if(x))$ and

$$u^{-1}\,\mathrm{d}u = \exp(-if)\,\mathrm{d}\exp(if) = i\,\mathrm{d}f,$$

so that indeed $A \mapsto A + i\,\mathrm{d}f$ represents the action of a gauge transformation as in lemma 1.10. We want to minimize the L^2-norm of the connection,

$$\int_M |A|^2\,,$$

along this gauge equivalent family. The Euler–Lagrange equation for this functional is

$$\mathrm{d}^*A = 0, \tag{2.1}$$

Indeed,

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle A + it\,\mathrm{d}f, A + it\,\mathrm{d}f\rangle &= 2i\langle\,\mathrm{d}f, A\rangle \\ &= 2i\langle f,\,\mathrm{d}^*A\rangle = 0,\ \ \forall\, f \in C^\infty(M),\end{aligned}$$

implies that $\mathrm{d}^*A = 0$. Equation (2.1) is our gauge fixing condition, and is usually called the **Hodge** or **Coulomb gauge**.[2] Thus we wish to find f such that $\tilde{A} := A + i\,\mathrm{d}f$ is gauge equivalent to A, and

$$\mathrm{d}^*\tilde{A} = \mathrm{d}^*(A + i\,\mathrm{d}f) = 0.$$

This is equivalent to

$$\Delta f = -i\,\mathrm{d}^*A.$$

From the Hodge decomposition theorem,[3] we know that there is a solution f if and only if d^*A is orthogonal to the kernel of the Laplacian, which in this case consists of the constant functions. So what we want is that $\int -1 \cdot i\,\mathrm{d}^*A = 0$, or $\int \mathrm{d}^*A = 0$. But

$$\int_M \mathrm{d}^*A = \pm\int_M *\,\mathrm{d} * A * 1 = \pm\int_M \mathrm{d} * A = 0$$

by Stokes's theorem, since we are assuming that M has no boundary. Therefore we find a solution f. The proof that this critical point of the functional is in fact a minimum is a little more involved and we skip it, as there seems to be no insight to be gained from it in this context.[4] More interesting than that is the fact that $\mathrm{d} + \mathrm{d}^*$ is an elliptic operator, which provides the following estimate for some constant C (since M is simply connected and so $H^1(M) = 0$):[5]

$$\|A\|_{k,p} \le C\left(\|\,\mathrm{d}A\|_{k-1,p} + \|\,\mathrm{d}^*A\|_{k-1,p}\right).$$

[2] Note that we are dealing with the case of classical electromagnetism, where $\nabla \cdot A = \mathrm{d}^*A = 0$ is the Coulomb gauge; see appendix A.

[3] See e.g. [War83, theorem 6.8], or [DK97, theorem (A.7)], which they call "the Fredholm alternative".

[4] For a proof, see e.g. [Jos02, section 2.2].

[5] Alternatively, see theorem 1.26.

When the gauge is fixed such that $\mathrm{d}^* A = 0$, and in the abelian case with $F_A = \mathrm{d}A$, we end up with

$$\|A\|_{k,p} \leq C \|F_A\|_{k-1,p}.$$

In conclusion, we showed that there exists an optimal gauge choice such that it minimizes the L^2-norm of the connection matrix. Beyond that, the gauge fixing condition together with the condition on the curvature yields an elliptic system,

$$\text{curvature } \mathrm{d}A \text{ and gauge fixing condition } \mathrm{d}^* A \longrightarrow \text{ elliptic operator } \mathrm{d} + \mathrm{d}^*,$$

and because of this the $W^{k,p}$-norms of the connection are bounded by the $W^{k-1,p}$-norms of the curvature. Small curvature leads to small connection, as we wanted. As a sanity check, note that if the connection is flat, then the Coulomb gauge indeed forces the connection matrix to be zero.

Uhlenbeck's gauge fixing lemma

In the discussion above we considered the case of an abelian gauge group action on a trivial bundle over a simply connected compact manifold without boundary. Each of these hypotheses were important for this straightforward development: the vanishing bracket lets us write the linear equation $F_A = \mathrm{d}A$; the triviality of the bundle let us choose a global gauge and work with the local representation of connections and gauge transformations; $H^1(M) = 0$ and compactness were needed for the elliptic estimate, and the empty boundary hypothesis was used for Stokes's theorem and to avoid dealing with (elliptic) boundary conditions. What we want now is for a similar result to hold locally on any smooth manifold with a non-abelian gauge group action on a bundle that is not necessarily trivial. Of course, if the result is local we may choose trivializing neighbourhoods and essentially work on a trivial bundle over the unit ball on Euclidean space. For elliptic theory to hold, we need the closed ball for compactness, but now the boundary is not empty and we will need a suitable boundary condition. Finally, for a non-abelian gauge group, the curvature is $F_A = \mathrm{d}A + \frac{1}{2}[A \wedge A]$ which leads to non-linear equations. We assert that the gauge we used above, supplemented with a suitable boundary condition, is still an interesting and profitable gauge choice. Indeed, on the closed ball, the Euler–Lagrange equations for the $\int_B |A|^2$ functional are [6]

$$\begin{cases} \mathrm{d}^* A &= 0, \\ *A|_{\partial U} &= 0. \end{cases}$$

Issues can arise when directly minimizing the norm this way, as the gauges constructed can have singularities; however, in the small curvature regime this is not a problem.[7] As we shall see later, these equations also fit in nicely with the Yang–Mills equation and yield an elliptic system. Finally, we would once again like for bounds on curvature to translate into bounds on the connection matrix. With all of these considerations, the following definition should feel natural:

Definition 2.1 (Uhlenbeck gauge). Let (M, g) be a Riemannian manifold, let G be a compact Lie group, and let $P \to M$ be a principal G-bundle.

[6] See lemma 1.37.
[7] See [FU91], the comment right before Lemma 8.2 on page 119.

Let $U \subseteq M$ be a trivializing neighbourhood of P. We say that a connection $A \in \mathcal{A}^{1,p}(U)$ is in Uhlenbeck gauge with constant $\tilde{C}$ if it satisfies

$$\begin{cases} \mathrm{d}^* A &= 0 \quad \text{on } U, \\ *A|_{\partial U} &= 0 \quad \text{on } \partial U, \end{cases} \tag{2.2}$$

and

$$\|A\|_{1,s} \leq \tilde{C}\, \|F_A\|_s \tag{2.3}$$

for $s = p$ or q, q as in the following theorem.

Our main theorem is then on the local existence of Uhlenbeck gauges.

Theorem 2.2 (Gauge fixing). *Suppose that $1 < q \leq p < \infty$ such that $q \geq \frac{n}{2}$, $p > \frac{n}{2}$, and in case $q < n$, $p \leq \frac{nq}{n-q}$. Then there exist constants $\tilde{C}$ and $\tilde{\varepsilon} > 0$ such that the following holds:*

For every point in M, there is a neighbourhood $U \subseteq M$ such that for every connection $A \in \mathcal{A}^{1,p}(U)$ with $\mathcal{E}_q(A) \leq \tilde{\varepsilon}$ there exists a gauge transformation $u \in \mathcal{G}^{2,p}(U)$ such that $\tilde{A} := u^* A$ is in Uhlenbeck gauge. Note that $|F_{\tilde{A}}| = |F_A|$.

Originally, this theorem was proved by Uhlenbeck in 1982 for $n > p \geq \frac{n}{2}$ and $q = \frac{n}{2}$ on the unit ball; it is theorem 2.1 in [Uhl82]. The condition $p > \frac{n}{2}$ guarantees that the gauge group is indeed a topological group with continuous action. Nevertheless, it is possible to extend the result for $p = \frac{n}{2}$ by a weak-limit argument.

Corollary 2.3. *The theorem also holds for $p = q = \frac{n}{2}$ if $n \geq 3$.*

Proof. See [Weh04, page 105], proof of remark 6.2a). □

It suffices to prove the theorem on a coordinate chart, since the result is local and the Uhlenbeck gauge conditions are invariant under change of coordinates on the base manifold. Thus, we now state the theorem on an open set in Euclidean (half) space, which we will prove in the next section, and then show how this implies theorem 2.2.

Proposition 2.4. *Let G be a compact Lie group and $B \subseteq \mathbb{R}^n$ the open unit ball or the "egg" [8]. Suppose that $1 < q \leq p < \infty$ such that $q \geq \frac{n}{2}$, $p > \frac{n}{2}$, and in case $q < n$, $p \leq \frac{nq}{n-q}$. Then there exist constants $\tilde{C}$, $\tilde{\varepsilon} > 0$ and $\delta > 0$ such that the following holds:*

If B is equipped with a smooth metric g such that $\|g - \mathbb{1}\|_{2,\infty} \leq \delta$ then for every connection $A \in \mathcal{A}^{1,p}(B)$ with $\mathcal{E}(A) \leq \tilde{\varepsilon}$ there exists a gauge transformation $u \in \mathcal{G}^{2,p}(B)$ such that $u^* A$ is in Uhlenbeck gauge with respect to the metric g and with constant $\tilde{C}$.

Proof of Theorem 2.2. Take $\delta > 0$ from Proposition 2.4, and take B as follows:

- For p in the interior of M, $B \subseteq \mathbb{R}^n$ is the unit ball around the origin.
- For $p \in \partial M$, **the "egg"** B is an open subset of the half space $\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \geq 0\}$ that contains a neighbourhood of 0 in $\partial\mathbb{H}^n$, is starshaped relative to 0 and has smooth boundary.[9]

[8] See definition below.

[9] This type of domain is called an "egg squeezed to the boundary" in [Weh04].

For $p \in M$, choose a coordinate chart around p, $\psi : V \to M$ for $V \subseteq \mathbb{R}^n$ or $\mathbb{H}^n$, such that $\psi^* g(0) = \mathbb{1}$.[10] For some small $\sigma \in (0,1]$, $\sigma B \subseteq V$ and we can restrict $\psi : \sigma B \to M$. To get a chart on B, consider $\psi_\sigma := \psi \circ \sigma : B \to M$. The pullback metric is $\psi_\sigma^* g(z) = \sigma^2 \psi^* g(\sigma z)$; indeed,

$$\begin{aligned}\psi_\sigma^* g(v,w)(z) &= (\psi \circ \sigma)^* g(v,w)(z)\\ &= \psi^* g(\,\mathrm{d}_z \sigma v,\, \mathrm{d}_z \sigma w)(\sigma z)\\ &= \sigma^2 \psi^* g(v,w)(\sigma z).\end{aligned}$$

Note that $\psi_\sigma^* g(0) = \sigma^2\, \mathbb{1}$, and so this metric is not close to the identity, but if we simply rescale by σ^{-2}, then $\sigma^{-2}\psi_\sigma^* g(z) = \psi^* g(\sigma z)$ is $W^{2,\infty}$-close to the identity, as the first derivative is $\nabla(\psi^* g \circ \sigma)(z) = \sigma \nabla(\psi^* g)(\sigma z)$, and the second derivative is

$$\begin{aligned}\nabla^2(\sigma^{-2}\psi_\sigma^* g)(z) &= \nabla^2(\psi^* g \circ \sigma)(z)\\ &= \nabla(\sigma(\nabla \psi^* g \circ \sigma))(z)\\ &= \sigma^2 \nabla^2 \psi^* g(\sigma z).\end{aligned}$$

Then, because $\psi^* g$ is smooth on the closure of σB (which is compact), these derivatives are bounded and can be made small by the choice of σ.

Now, having chosen σ such that $\left\|\sigma^{-2}\psi_\sigma^* g - \mathbb{1}\right\|_{2,\infty} \le \delta$, Proposition 2.4 holds on B with metric $g_B := \sigma^{-2}\psi_\sigma^* g$. However, in order to obtain the result on $U := \psi(\sigma B) \subseteq M$ with the intended metric g, we need to show that the result still holds on B with metric $\sigma^2 g_B = \psi_\sigma^* g$.

On the effect of the conformal change of metric on $\mathcal{A}$, $\mathcal{G}$ and condition (2.2), note that over a compact manifold the metrics are equivalent and so the spaces are the same, and as for the equations, the change of metric affects the Hodge star, however only by possible conformal scalings, and the equations still hold.

What is left is to check that the bounds (2.3) are still valid with the same constant $\tilde{C}$. First, we look at the effect of the rescaling on the norm of the curvature:

$$\begin{aligned}\|F_A\|_{\sigma^2 g;q}^q &= \int_B |F_A|_{\sigma^2 g}^q \sqrt{\det(\sigma^2 g)}\, \mathrm{d}^n x\\ &= \int_B (\sigma^{-2} g^{ik} \sigma^{-2} g^{jl} (F_A)_{ij} (F_A)_{kl})^{\frac{q}{2}} \sqrt{\sigma^{2n}\det(g)}\, \mathrm{d}^n x\\ &= \sigma^{n-2q} \|F_A\|_{g;q}^q\,.\end{aligned}$$

For q as in the theorem, this is the L^q-energy $\mathcal{E}_{\sigma^2 g;\,q}(A)$. If it is bounded $\mathcal{E}_{\sigma^2 g;\,q}(A) \le \tilde{\varepsilon}$, then since $q \ge \frac{n}{2}$ and $\sigma \le 1$, $\mathcal{E}_{g;\,q}(A) \le \sigma^{2q-n}\tilde{\varepsilon}$ gives $\mathcal{E}_{g;\,q}(A) \le \tilde{\varepsilon}$. The calculation for the L^p-norm is the same. For the $W^{1,p}$-norm of a connection $A \in \mathcal{A}^{1,p}(B)$, we first have the straightforward calculation

$$\|A\|_{\sigma^2 g;p}^p = \int_B (\sigma^{-2} g^{ij} A_i A_j)^{\frac{p}{2}} \sqrt{\det(\sigma^2 g)}\, \mathrm{d}^n x = \sigma^{n-p} \|A\|_{g;p}^p\,,$$

and then for the covariant derivative note that $(\nabla A)_{ij} = \nabla_i A_j - \Gamma_{ij}^k A_k$ and the Christoffel symbols for $\sigma^2 g$ and g are the same.[11] Then $\nabla^g A = \nabla^{\sigma^2 g} A$ and as for the curvature

$$\left\|\nabla^{\sigma^2 g} A\right\|_{\sigma^2 g;p}^p = \sigma^{n-2p} \|\nabla^g A\|_{g;p}^p\,.$$

[10] Note that this is always possible, as you can simply choose an orthonormal basis on $T_p M$ and then parallel transport it to get a local frame over a contractible domain.

[11] Indeed, Γ_{ij}^k = a sum of multiples (components of g^{-1})×(derivatives of components of g), and so a constant rescaling is cancelled.

Putting both terms together we have

$$\|A\|^p_{\sigma^2 g;\,1,p} = \sigma^{n-p}\,\|A\|^p_{g;\,1,p} + \sigma^{n-2p}\,\|\nabla A\|^p_{g;\,1,p} \le \sigma^{n-2p}\,\|A\|^p_{g;\,1,p}\,.$$

Finally, if A is a connection that has been put in Uhlenbeck gauge and satisfies (2.3) with respect to the metric $\sigma^2 g$, then simply concatenating our inequalities

$$\|A\|_{\sigma^2 g;\,1,p} \le \sigma^{\frac{n-2p}{p}}\,\|A\|_{g;\,1,p} \le \tilde{C}\,\sigma^{\frac{n-2p}{p}}\,\|F_A\|_{g;\,p} = \tilde{C}\,\|F_A\|_{\sigma^2 g;\,p}$$

we see that A also satisfies (2.3) with respect to the metric g and with the same constant $\tilde{C}$. □

2.2 The big bad proof

The proof of proposition 2.4 is the main technical result of this text. It will refer to most of the lemmas from the preceding chapter and appendix B. For its relevance (and length), we give it its own section. The proof will go through via the continuous induction method, and its main interesting features are the use of the implicit function theorem, boundary value spaces, the elliptic estimate for the $\mathrm{d} + \mathrm{d}^*$ operator and results relating to the Neumann problem.

Define the modified energy $\mathcal{E}'_q(A) = \int_B |F_A(x)|^q \,\mathrm{d}x^n$ using the Euclidean metric on B for a connection A, and note that if δ is small enough in $\|g - \mathbb{1}\|_{2,\infty} \le \delta$, then g is sufficiently close to the identity that

$$\frac{1}{2}\mathcal{E}'_q(A) \le \mathcal{E}_q(A) \le 2\mathcal{E}'_q(A).$$

Define

$$\mathcal{A}_\varepsilon := \left\{ A \in \mathcal{A}^{1,p}(B) : \mathcal{E}'_q(A) < \varepsilon \right\}, \tag{2.4}$$

and

$$S_\varepsilon := \{A \in \mathcal{A}_\varepsilon \text{ such that } A \text{ can be put into Uhlenbeck gauge}\}\,. \tag{2.5}$$

We will show that $S_\varepsilon = A_\varepsilon$, and thus every connection with energy at most $\tilde{\varepsilon} := \frac{\varepsilon}{2}$ can be put into Uhlenbeck gauge.

The proof will have three steps, and during each step some care has to be taken with the constants ε, $\tilde{C}$ and δ and their dependence on each other and the metric on B:

(1) For some fixed g and ε, we prove $\mathcal{A}_\varepsilon$ is connected.

(2) For some fixed g, ε and $\tilde{C}$, S_ε is closed. (Proof begins on page 50.)

(3) We find ε, $\tilde{C}$ and δ, and vary the metric with δ, such that S_ε is open. (Proof begins on page 51.)

Throughout, we take the local descriptions explained in subsection 1.2.1. Therefore, a connection in $\mathcal{A}^{1,p}(B)$ as defined in (1.25) is a $\mathfrak{g}$-valued 1-form on B and a gauge transformation in $\mathcal{G}^{2,p}(B)$ as defined in (1.26) is a function $u : B \to G$.

(1) $\mathcal{A}_\varepsilon$ is connected

Proposition 2.5. *Let $\overline{B}$ be equipped with any smooth metric g and let $\varepsilon > 0$; then $\mathcal{A}_\varepsilon$ as defined above is connected.*

Proof. We prove there is a continuous path from each $A \in \mathcal{A}_\varepsilon$ to the canonical flat connection in $\mathcal{A}_\varepsilon$, which is represented by $A \equiv 0 \in \Omega^1(B, \mathfrak{g})$.

Let $A \in \mathcal{A}_\varepsilon$, and define the path $(A_\sigma)_{\sigma\in[0,1]}$ by $A_\sigma(x) = \sigma^* A(x) = \sigma A(\sigma x)$ where we take the pullback under the map $x \mapsto \sigma x$.[12] Clearly, $A_0 = 0$ and $A_1 = A$, and for each $\sigma \in [0,1]$, $A_\sigma \in \mathcal{A}^{1,p}(B)$. The curvature of the connection A_σ is [13]

$$F_{A_\sigma}(x) = \mathrm{d}(A_\sigma) + \frac{1}{2}[A_\sigma \wedge A_\sigma] = \sigma^2\, \mathrm{d}A(\sigma x) + \frac{\sigma^2}{2}[A(\sigma x) \wedge A(\sigma x)] = \sigma^2 F_A(\sigma x),$$

and so we have

$$\mathcal{E}'_q(A_\sigma) = \int_B \sigma^{2q} \left|F_A(\sigma x)\right|^q \,\mathrm{d}x^n = \sigma^{2q-n} \int_{\sigma B} \left|F_A(y)\right|^q \,\mathrm{d}y^n \leq \sigma^{2q-n}\mathcal{E}'(A) \leq \varepsilon,$$

where we changed variables $y = \sigma x$ and used the assumption that $q \geq \frac{n}{2}$ for $\sigma^{2q-n} \leq 1$. Therefore, the whole path is contained in $\mathcal{A}_\varepsilon$.

To show the continuity of the path, we will use the Euclidean metric on $\overline{B}$ instead of the metric g. For continuity at $\sigma = 0$, since $A_0 = 0$, it suffices to show that $\|A_\sigma\|_{\mathbb{1};1,p} \overset{\sigma\to 0}{\longrightarrow} 0$. We look at the two L^p-norms separately. First, we have

$$\begin{aligned}
\|A_\sigma\|^p_{\mathbb{1};p(B)} &= \int_B |A_\sigma|^p \,\mathrm{d}x^n \\
&= \sigma^p \int_B |A(\sigma x)|^p \,\mathrm{d}x^n \\
&= \sigma^{p-n} \int_{\sigma B} |A(y)|^p \,\mathrm{d}y^n \\
&= \sigma^{p-n} \|A\|^p_{\mathbb{1};p(\sigma B)} \\
&\overset{(1)}{\leq} \sigma^{p-n} \|1\|^p_{\mathbb{1};\,2p(\sigma B)} \|A\|^p_{\mathbb{1};\,2p(\sigma B)} \\
&= \sigma^{p-n} Vol(\sigma B)^{\frac{1}{2}} \|A\|^p_{\mathbb{1};\,2p(\sigma B)} \\
&\overset{(2)}{\leq} \sigma^{p-n+\frac{n}{2}} Vol(B)^{\frac{1}{2}} C_W \|A\|^p_{\mathbb{1};\,1,p(\sigma B)} \\
&\leq \underbrace{\sigma^{p-\frac{n}{2}}}_{\to 0} \underbrace{Vol(B)^{\frac{1}{2}} C \|A\|^p_{\mathbb{1};\,1,p(B)}}_{\text{constant}} \longrightarrow 0,
\end{aligned}$$

where in (1) we used the Hölder inequality and (2) is the Sobolev inequality for $W^{1,p} \hookrightarrow L^{2p}$ guaranteed by the hypothesis $p > \frac{n}{2}$, which is also used for $\sigma^{p-\frac{n}{2}} \to 0$. For the derivative term,

$$\begin{aligned}
\|\nabla A_\sigma\|_{\mathbb{1};\,p(B)} &= \int_B |\nabla A_\sigma(x)|^p \,\mathrm{d}x^n \\
&= \int_B \sigma^p |\nabla(A \circ \sigma)(x)|^p \,\mathrm{d}x^n
\end{aligned}$$

[12] This is well defined since B is star-shaped with respect to 0 and therefore the path σx stays within B.

[13] Use $\mathrm{d}_x(A_\sigma) = \mathrm{d}_x(\sigma \circ A \circ \sigma)$, and recall that the differential of a linear map is itself.

$$
\begin{aligned}
&= \sigma^{2p} \int_B |\nabla A(\sigma x)|^p \, \mathrm{d}x^n \\
&= \sigma^{2p-n} \int_{\sigma B} |\nabla A(y)|^p \, \mathrm{d}y^n \\
&= \sigma^{2p-n} \|\nabla A\|^p_{\mathbb{1};\,p(\sigma B)} \longrightarrow 0
\end{aligned}
$$

where we simply use a change of variables and again need the condition $p > \frac{n}{2}$ for convergence.

The continuity at $\sigma_0 \in (0,1]$ is more convoluted, and we will make use of an auxiliary sequence $(A^i)_{i\in\mathbb{N}} \subseteq \mathcal{A}(B)$ of smooth connections which converge to A in $\mathcal{A}^{1,p}(B)$:

$$
\begin{aligned}
|A_\sigma(x) - A_{\sigma_0}(x)| &= |\sigma A(\sigma x) - \sigma_0 A(\sigma_0 x)| \\
&= |\sigma A(\sigma x) - \sigma A^i(\sigma x) + \sigma A^i(\sigma x) \ - \sigma A^i(\sigma_0 x) + \sigma A^i(\sigma_0 x) \\
&\quad - \sigma A(\sigma_0 x) + \sigma A(\sigma_0 x) - \sigma_0 A(\sigma_0 x)| \\
&\leq \sigma \left| A(\sigma x) - A^i(\sigma x) \right| + \sigma \left| A^i(\sigma x) - A^i(\sigma_0 x) \right| \\
&\quad + \sigma \left| A^i(\sigma_0 x) - A(\sigma_0 x) \right| + |\sigma - \sigma_0| \, |A(\sigma_0 x)| \\
&\leq \left| A(\sigma x) - A^i(\sigma x) \right| + |\sigma - \sigma_0| \, C_i \\
&\quad + \left| A^i(\sigma_0 x) - A(\sigma_0 x) \right| + |\sigma - \sigma_0| \, |A(\sigma_0 x)|
\end{aligned}
$$

where we have used $\sigma \leq 1$ and the mean value inequality, proposition B.2, for $\|\nabla A^i(x)\| \leq C_i$, since the A_i and all of their derivatives are bounded on $\overline{B}$. Now we apply the Euclidean norm $\|\cdot\|_{\mathbb{1};\,p(B)}$ to this inequality [14] and change variables to obtain

$$
\begin{aligned}
\|A_\sigma - A_{\sigma_0}\|_{\mathbb{1};\,p} &\leq \sigma^{-\frac{n}{p}} \left\| A - A^i \right\|_{\mathbb{1};\,p(\sigma B)} + |\sigma - \sigma_0| \, C_i Vol(B)^{\frac{1}{p}} \\
&\quad + \sigma_0^{-\frac{n}{p}} \left\| A - A^i \right\|_{\mathbb{1};\,p(\sigma_0 B)} + |\sigma - \sigma_0| \, \sigma_0^{-\frac{n}{p}} \|A\|_{\mathbb{1};\,p(\sigma_0 B)} \\
&\leq \left(\sigma^{-\frac{n}{p}} + \sigma_0^{-\frac{n}{p}} \right) \left\| A - A^i \right\|_{\mathbb{1};\,p(B)} + |\sigma - \sigma_0| \, C_i Vol(B)^{\frac{1}{p}} \\
&\quad + |\sigma - \sigma_0| \, \sigma_0^{-\frac{n}{p}} \|A\|_{\mathbb{1};\,p(B)} \, .
\end{aligned}
$$

It is now necessary to be a bit careful about the mixed terms depending on i and σ, but for fixed σ_0 the right-hand side can be made as small as we want in the following way: first, take σ close to σ_0 such that $\sigma^{-\frac{n}{p}} + \sigma_0^{-\frac{n}{p}}$ is bounded, say, by $2\sigma_0^{-\frac{n}{p}} + c$ for some constant c; then take i such that the first term is small; for this fixed i, C_i is also constant and therefore a further suitable choice of σ even closer to σ_0 makes the second and third terms small.

The computation to check that $\|\nabla A_\sigma - \nabla A_{\sigma_0}\| \overset{\sigma\to\sigma_0}{\longrightarrow} 0$ is completely analogous, using bounds C_i' on the second derivatives of A^i, yielding

$$
\|\nabla A_\sigma - \nabla A_{\sigma_0}\|_{\mathbb{1};\,p(B)} \leq \left(\sigma^{-\frac{n}{p}} + \sigma_0^{-\frac{n}{p}} \right) \left\| \nabla A - \nabla A^i \right\|_{\mathbb{1};\,p(B)} + |\sigma - \sigma_0| \, C_i' Vol(B)^{\frac{1}{p}}
$$

[14] That is, we elevate both sides to the p-th power, integrate over B with the Euclidean volume form and take the $\frac{1}{p}$ power, and then separate the terms, all of which can be done because the integral, $x \mapsto x^p$ and $x \mapsto x^{\frac{1}{p}}$ are subadditive.

$$+ \left|\sigma^2 - \sigma_0^2\right| \sigma_0^{-\frac{n}{p}} \|\nabla A\|_{\mathbb{1};\,p(B)} \longrightarrow 0$$

for suitable choices of i and σ close to σ_0 as before. □

(2) S_ε is closed

Proposition 2.6. *Let B be equipped with any smooth metric g and let $\varepsilon > 0$. Suppose that there is a sequence $(A_i)_{i\in\mathbb{N}} \subseteq \mathcal{A}_\varepsilon$ converging to some $A \in \mathcal{A}_\varepsilon$ such that each A_i can be put into Uhlenbeck gauge with constant $\tilde{C}$ by some gauge transformation $u_i \in \mathcal{G}^{2,p}(B)$. Then there exists $u \in \mathcal{G}^{2,p}(B)$ such that u^*A is also in Uhlenbeck gauge with constant $\tilde{C}$.*

Proof. We will show that there exist the limits $u_i \to u$ and $u_i^* A_i =: \tilde{A}_i \to \tilde{A}$ such that $u^*A = \tilde{A}$ and this is in Uhlenbeck gauge.

We can get a uniform bound on $\left\|\tilde{A}_i\right\|_{1,p}$ by first bounding the curvature, from lemma 1.29,

$$\|F_{A_i}\|_p \le c(\|A_i\|_{1,p} + \|A_i\|_p^2) \le c',$$

with uniform c' due to the $W^{1,p}$-convergence of the A_i, and then applying the Uhlenbeck gauge condition

$$\left\|\tilde{A}_i\right\|_{1,p} \le \tilde{C}\,\|F_{A_i}\|_p \le c'\tilde{C}.$$

With this uniform bound, we may use the Banach–Alaoglu theorem 1.18 to find a subsequence weakly converging to some $\tilde{A} \in \mathcal{A}^{1,p}(B)$, and then from the compact embedding $W^{1,p} \hookrightarrow L^{2p}$, a further subsequence[15] also converges in the L^{2p}-norm to $\tilde{A}$.

Then, since both $\|A_i\|_{1,p}$ and $\left\|\tilde{A}_i\right\|_{1,p}$ are uniformly bounded, lemma 1.33 gives us a subsequence u_i such that it converges in the C^0-topology to some $u \in \mathcal{G}^{2,p}(B)$, and $u_i^{-1}\,\mathrm{d}u_i \longrightarrow u^{-1}\,\mathrm{d}u$ uniformly in the L^{2p}-norm.

Thus we have

$$u^{-1}\,\mathrm{d}u \longleftarrow u_i^{-1}\,\mathrm{d}u_i = \tilde{A}_i - u_i^{-1}A_iu_i \longrightarrow \tilde{A} - u^{-1}Au,$$

where the second limit follows because the two terms converge, and the second term converges because of the continuity of the adjoint action.[16] Equality follows from the uniqueness of the L^{2p}-limit, and so $\tilde{A} = u^*A$.

Now, we check that $\tilde{A}$ is in Uhlenbeck gauge, i.e. check (2.2) and (2.3):

(i) All $\tilde{A}_i$ are already in Uhlenbeck gauge and thus satisfy $\mathrm{d}^*\tilde{A}_i = 0$, and for all $\phi \in C^\infty(B)$ such that $\phi|_{\partial B} = 0$,

$$\langle \phi, \mathrm{d}^*\tilde{A}\rangle = \int_B \phi\,\mathrm{d}^*\tilde{A} * 1 = \int_B \phi * (\,\mathrm{d}^*\tilde{A} - \underbrace{\mathrm{d}^*\tilde{A}_i}_{=0})$$

$$= \underbrace{*^2}_{=\pm 1} \int_B \phi\;\mathrm{d}(*\tilde{A} - *\tilde{A}_i) \overset{(*)}{=} \pm \int_B \mathrm{d}\phi \wedge *(\underbrace{\tilde{A} - \tilde{A}_i}_{\to 0}) \longrightarrow 0,$$

where $(*)$ follows from the following computation (where we write $\alpha \in \Omega^1(B)$), using Stokes's theorem 1.20 and $\phi|_{\partial B} = 0$

$$0 = \int_{\partial B} \phi * \alpha = \int_B \mathrm{d}(\phi * \alpha) = \int_B \mathrm{d}\phi \wedge *\alpha + (-1)^{n-1} \int_B \phi \wedge \,\mathrm{d} * \alpha$$

[15] We keep the same labelling $i \in \mathbb{N}$ for the subsequence.
[16] See calculation in proof of lemma 1.31.

Since $C^\infty_\delta(B)$ is dense in $L^p(B)$, this proves that $\mathrm{d}^* \tilde{A} = 0$.

(ii) Similarly, $*\tilde{A}_i|_{\partial B} = 0$ and this is also preserved under the weak $W^{1,p}$-limit, as we show. For any $\phi \in C^\infty(\partial B)$ we may extend it to some $\Phi \in C^\infty(B)$, and so we have

$$\begin{aligned}
\langle \phi, *\tilde{A}|_{\partial B}\rangle &= \int_{\partial B} \phi * \tilde{A}|_{\partial B} = \int_{\partial B} \Phi * \tilde{A}|_{\partial B} = \int_{\partial B} \Phi * (\tilde{A} - \tilde{A}_i)|_{\partial B} \\
&\overset{(*)}{=} \int_B \mathrm{d}(\Phi * (\tilde{A} - \tilde{A}_i)) = \int_B \mathrm{d}\Phi * (\tilde{A} - \tilde{A}_i) \pm \int_B \Phi \, \mathrm{d} * (\tilde{A} - \tilde{A}_i) \\
&= \int_B \mathrm{d}\Phi \wedge *(\tilde{A} - \tilde{A}_i) \pm \int_B \Phi * (* \mathrm{d} * (\tilde{A} - \tilde{A}_i)) \\
&= \int_B \mathrm{d}\Phi \wedge * \underbrace{(\tilde{A} - \tilde{A}_i)}_{\to 0} \pm \int_B \Phi \underbrace{(\mathrm{d}^*\tilde{A} - \mathrm{d}^*\tilde{A}_i)}_{=0}
\end{aligned}$$

where in $(*)$ again we used Stokes's theorem 1.20. This then shows that $*\tilde{A}|_{\partial B} = 0$ as we wished.

(iii) Let $s = p$ or q. We may write

$$\left\|\tilde{A}\right\|_{1,s} \le \liminf_{i\to\infty} \left\|\tilde{A}_i\right\|_{1,s} \le \tilde{C} \liminf_{i\to\infty} \|F_{A_i}\|_s \le \tilde{C}\, \|F_A\|_s \,,$$

and this gives us our result. The first inequality is true because any norm is sequentially weakly lower semicontinuous; the second follows from $\tilde{A}_i$ being in Uhlenbeck gauge; and the third follows from the continuity of the L^s-energy functional on $\mathcal{A}^{1,s}$ from lemma 1.29, and the convergence $A_i \xrightarrow{1,p} A$, which also implies there is convergence in $W^{1,q}$ because for $p \ge q$ and B of finite volume, $W^{1,q}(B) \hookrightarrow W^{1,p}(B)$.

□

(3) S_ε is open

We would like to show that for any $A \in \mathcal{S}_\varepsilon$ there is a neighbourhood of A in $\mathcal{A}_\varepsilon$ contained in $\mathcal{S}_\varepsilon$. Instead, it is simpler to find a neighbourhood of $A_0 := u^*A$ in $\mathcal{A}^{1,p}(B)$ made up of connections which can be put in Uhlenbeck gauge, pull it back by u^{-1} to a neighbourhood of A in $\mathcal{A}^{1,p}(B)$ and intercept it with $\mathcal{A}_\varepsilon$ to get what we need. This string of operations makes sense because $\mathcal{G}^{2,p}(B)$ acts continuously on $\mathcal{A}^{1,p}(B)$ (see lemma 1.31) and is closed under compositions. Since the energy $\mathcal{E}(A)$ is gauge invariant, we can forget about the original connection A and work with an arbitrary connection in Uhlenbeck gauge.

We will make use of the implicit function theorem for a suitable operator between Banach spaces, which will yield a neighbourhood of connections around A_0 which satisfy (2.2) in the Uhlenbeck gauge definition. We can then ask for a bit more from the solutions so that the connections satisfy condition (2.3), owing to a *a priori* estimates.

Before moving on to the proof, it will be important to establish some inequalities with uniform constants. Since the $W^{1,p}$-norm only depends on the metric, its inverse and first derivatives, control over the $W^{1,\infty}$-norm of the metric is enough to guarantee that if $\|g - \mathbb{1}\|_{1,\infty} \le \delta$ for small enough δ, we get

$$\frac{1}{2}\|\alpha\|_{g;1,p} \le \|\alpha\|_{\mathbb{1};1,p} \le 2\,\|\alpha\|_{g;1,p}\,, \quad \forall \alpha \in W^{1,p}(B, T^*B). \tag{2.6}$$

If we choose δ such that this equivalence is valid for the $W^{1,p}$, $W^{1,q}$ and L^r-norms (with $r = r(n,p,q)$ from lemma B.9) at the same time, we moreover get uniform constants in the estimates between these spaces, by first using the appropriate estimate with the Euclidean metric $\mathbb{1}$, and then using (2.6). This means that throughout the next proofs, the constants C_{rnpq} coming from lemma B.9 and C_W coming from the Sobolev embeddings will not depend on the metric.

Lemma 2.7 (A priori estimates). *There exist positive constants δ, $\tilde{C}$, and Λ such that for every metric g satisfying $\|g - \mathbb{1}\|_{1,\infty} \leq \delta$ the following holds:*

Let $A \in \mathcal{A}^{1,p}(B)$ be such that $\mathrm{d}^* A = 0$ on B and $*A|_{\partial B} = 0$, and $\|A\|_r \leq \Lambda$, where $r = r(n,p,q)$ from lemma B.9. Then

$$\|A\|_{1,p} \leq \tilde{C}\,\|F_A\|_p\,,$$

$$\|A\|_{1,q} \leq \tilde{C}\,\|F_A\|_q\,,$$

that is, A is in Uhlenbeck gauge with constant $\tilde{C}$.

Proof. As we are working over the ball in Euclidean space and $*A|_{\partial B} = 0$, it follows from theorem 1.26 that for all $1 < s < \infty$,

$$\|A\|_{1,s} \leq C_g(\|\,\mathrm{d}A\|_s + \underbrace{\|\,\mathrm{d}^*A\|_s}_{=0}),$$

where C_g depends on the metric g on B; however, we also know that the constant depends $W^{1,\infty}$-continuously on the metric, and so we may choose δ such that there is a uniform constant C for all metrics $\|g - \mathbb{1}\|_{1,\infty} \leq \delta$, and thus

$$\|A\|_{1,s} \leq C\,\|\,\mathrm{d}A\|_s\,.$$

Now, looking at the curvature as $\mathrm{d}A = F_A - \frac{1}{2}[A \wedge A]$,

$$\|A\|_{1,s} \leq C(\|F_A\|_s + \frac{1}{2}\,\|[A \wedge A]\|_s)$$

and we would like to get rid of this $\|[A \wedge A\|$ term somehow. There is a clever trick to make it so that we can absorb this term into the constant and the left hand side. We take $s = p$ or q and use lemma B.9 in the following:

$$\begin{aligned}
\frac{1}{2}\,\|[A \wedge A]\|_s &\leq \|[A \wedge A]\|_{\mathbb{1};s} \\
&\overset{(1)}{\leq} \|2\,|A| \cdot |A|\|_{\mathbb{1};s} \\
&\leq 2C_{rnpq}\,\|A\|_{\mathbb{1};r}\,\||A|_{\mathbb{1}}\|_{\mathbb{1};1,s} \\
&\overset{(2)}{\leq} 8C_{rnpq}\,\|A\|_r\,\|A\|_{1,s} \\
&\leq 8C_{rnpq}\Lambda\,\|A\|_{1,s}\,,
\end{aligned}$$

where in (1) we used that $|[A \wedge A]|_{\mathbb{1}} \leq 2\,|A|^2_{\mathbb{1}}$ (see e.g. (1.28) and proof of lemma 1.29), and in (2) we used that $|\nabla\,|A|| \leq |\nabla A|$ (see lemma B.10). Since we used lemma B.9 for the Euclidean metric and then used the equivalence of the norms (2.6), C_{rnpq} is metric independent and we may absorb it into C,

$$\|A\|_{1,s} \leq C(\|F_A\|_s + \Lambda \|A\|_{1,s}),$$

and now all that is left to do is take $\Lambda = \frac{1}{2C}$ and combine the $\|A\|_{1,s}$ terms, and the result follows for $\tilde{C} = 2C$. □

We will need to keep track of this extra condition $\|A\|_r \leq \Lambda$ while solving the boundary value problem posed by (2.2), and this will appear as a bound λ on X for a gauge transformation $\exp(X)$, so that a transformed connection $\exp(X)^*A$ will satisfy this Λ bound on its L^r-norm.

Lemma 2.8. *There exists $\delta > 0$ such that for every constant $\tilde{C} > 0$ there exists $\varepsilon > 0$ such that for every metric g satisfying $\|g - \mathbb{1}\|_{2,\infty} \leq \delta$ the following is true:*

Let $A_0 \in \mathcal{A}^{1,p}(B)$ be in Uhlenbeck gauge with constant $\tilde{C}$ and energy $\mathcal{E}(A) \leq \varepsilon$; then for all $\lambda > 0$ there exists $R > 0$ such that for every connection $A \in \mathcal{A}^{1,p}(B)$ with $\|A - A_0\|_{1,p} \leq R$ there is a solution $X \in W^{2,p}(B, \mathfrak{g})$ of

$$\begin{cases} \mathrm{d}^*(\exp(X)^*A) & = 0 \text{ on } B \\ *(\exp(X)^*A)|_{\partial B} & = 0 \text{ on } \partial B, \end{cases}$$

with $\|X\|_{2,p} \leq \lambda$.

Proof. As explained, we will use the implicit function theorem B.3. Let us define our Banach spaces:

First, define the space

$$W^{2,p}_m(B, \mathfrak{g}) := \{X \in W^{2,p}(B, \mathfrak{g}) : \int_B X = 0\},$$

which is a closed subspace[17] of a Banach space, and therefore also Banach. Next, define

$$W^{1,p}_\partial(B, \mathfrak{g}) := \frac{W^{1,p}(B, \mathfrak{g})}{W^{1,p}_\delta(B, \mathfrak{g})}$$

where $W^{1,p}_\delta(B, \mathfrak{g})$ is the closure in the $W^{1,p}$-norm of the subspace of sections which vanish at the boundary of B. This is a boundary value space, whose elements are equivalence classes of functions on B which coincide on the boundary ∂B. With the quotient norm

$$\|\phi\|_{W^{1,p}_\partial} (B, \mathfrak{g}) = \inf\{\|\Phi\|_{1,p} : \Phi \in W^{1,p}(B, \mathfrak{g}) \text{ and } \Phi|_{\partial B} = \phi\},$$

$W^{1,p}_\partial(B, \mathfrak{g})$ is a Banach space. Now define

$$\mathcal{Z} = \{(f, \phi) \in L^p(B, \mathfrak{g}) \times W^{1,p}_\partial(B, \mathfrak{g}) : \int_B f + \int_{\partial B} \phi = 0\}.$$

Note that $L^p(B, \mathfrak{g}) \times W^{1,p}_\partial(B, \mathfrak{g})$ with the direct sum norm is a Banach space, and $\mathcal{Z}$ is a closed subspace, and therefore Banach. To see that

$$\int_B f + \int_{\partial B} \phi = 0 \tag{2.7}$$

[17] $\int_B X = 0$ is a (clearly linear and) closed condition, which is easily seen taking limits in each component of $\mathfrak{g}$.

is a closed condition, note that using the trace theorem 1.19 we get

$$\int_{\partial B} |\phi| = \|\phi\|_{1(\partial B)} \leq C\,\|\Phi\|_{1,p\,(B)} \leq C(\|\phi\|_{W^{1,p}_\partial} + \varepsilon)$$

for some choice of Φ such that $\Phi|_{\partial B} = \phi$, which is enough to see that (2.7) is preserved under limits. Finally, $\mathcal{A}^{1,p}(B)$ is a Sobolev space and therefore clearly Banach.

We will use the implicit function theorem on the operator

$$\begin{aligned} F : \mathcal{A}^{1,p}(B) \times W^{2,p}_m(B, \mathfrak{g}) &\longrightarrow \quad \mathcal{Z} \\ (A, X) &\mapsto (\,\mathrm{d}^*(\exp(X)^*A), *(\exp(X)^*A)|_{\partial B}). \end{aligned}$$

That F is a continuous map into $L^p(B, \mathfrak{g}) \times W^{1,p}_\partial(B, \mathfrak{g})$ follows from the facts that $\exp X \in \mathcal{G}^{2,p}(B)$ for $X \in W^{2,p}(B, \mathfrak{g})$ by definition (see lemma 1.21), the map $X \mapsto \exp X$ is continuous, and the gauge action $\mathcal{G}^{2,p} \times \mathcal{A}^{1,p} \to \mathcal{A}^{1,p}$ is continuous (see lemma 1.31). That it maps into $\mathcal{Z}$ can be checked with Stokes's theorem (lemma 1.20): for $\alpha = (\exp X)^*A$,[18]

$$\int_B \mathrm{d}^*\alpha * 1 = -\int_B \mathrm{d} * \alpha = -\int_{\partial B} *\alpha|_{\partial B},$$

so that 2.7 is indeed satisfied by (f, ϕ) in the image of F. Thus, F is a continuous operator between Banach spaces.

Now, knowing that $F(A_0, 0) = 0$, what we want is an open set around A_0 in $\mathcal{A}^{1,p}(B)$ and some continuous map f on this open set to a neighbourhood of $0 \in W^{2,p}_m(B, \mathfrak{g})$ such that $F(A, f(A)) = 0$, and this map will permit us to control the norm of X. Under these circumstances, we need to analyse the derivative of F with respect to the second variable X. The partial derivative $\partial_X F|_{(A,X)}$ is the Fréchet derivative of the map $F(A, \cdot)$, which is a linear map

$$\partial_X F|_{(A,X)} : W^{2,p}_m(B, \mathfrak{g}) \to \mathcal{Z}.$$

We check that this map exists, is continuous and is bijective at $(A_0, 0)$. First, note that since the chain rule is valid for the Fréchet derivative, and d, $*$ and $|_{\partial B}$ are linear maps, we need only look at the Fréchet derivative of the gauge action $X \mapsto (\exp(X))^*A$, which we will call $\mathcal{G}(A, X) : W^{2,p}_m(B, \mathfrak{g}) \to \mathcal{A}^{1,p}(B)$, and then

$$\partial_X F|_{(A,X)} : \xi \mapsto (\,\mathrm{d}^*(\mathcal{G}(A, X)\xi), *(\mathcal{G}(A, X)\xi)|_{\partial B}).$$

We will show that the linearization of the gauge action is

$$\mathcal{G}(A, X)\xi = \mathrm{d}\xi + \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{-X} \exp(-\xi))A,$$

which is to say that

$$\lim_{\|\xi\| \to 0} \frac{\left|\exp(X + \xi)^*A - \exp(X)^*A - \mathrm{d}\xi - \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{-X} \exp(-\xi))A\right|}{\|\xi\|} = 0.$$

We can rewrite this in a more suggestive format,

$$\begin{aligned} &\exp(X + \xi)^*A - \exp(X)^*A - \mathrm{d}\xi - \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{-X} \exp(-\xi))A \\ &\qquad = \operatorname{Ad}_{\exp(X+\xi)^{-1}} A + \exp(X + \xi)^{-1}\,\mathrm{d} \exp(X + \xi) \end{aligned}$$

[18]Use $k = n$ in the formulas $*^2 = (-1)^{k(n-k)}$ and $\mathrm{d}^* = -(-1)^{n(k-1)}$.

$$
\begin{aligned}
&- \operatorname{Ad}_{\exp X^{-1}} A - \exp X^{-1} \mathrm{d} \exp X \\
&- \mathrm{d}\xi - \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{-X} \exp(-\xi))A \\
= &\operatorname{Ad}_{\exp(X+\xi)^{-1}} A - \operatorname{Ad}_{\exp X^{-1}} A - \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{-X} \exp(-\xi))A,
\end{aligned}
$$

and observe that this expression is exactly the expression for the derivative of $\operatorname{Ad}_{\exp Y}$,

$$
\begin{aligned}
\mathrm{d}_{(-X)}(\operatorname{Ad}_{\exp})(-\xi) &= \mathrm{d}_{(-X)}(\operatorname{Ad} \circ \exp)(-\xi) \\
&= \mathrm{d}_{\exp(-X)} \operatorname{Ad}(\mathrm{d}_{(-X)} \exp(-\xi)).
\end{aligned}
$$

Thus the limit is indeed zero. Moreover, $\mathcal{G}(A, X)$ is clearly continuous.

To check bijectivity of $\partial_X F|_{(A_0,0)}$ we simplify the expression

$$
\begin{aligned}
\mathcal{G}(A_0, 0)\xi &= \mathrm{d}\xi + \mathrm{d}_{\mathbb{1}} \operatorname{Ad}(\mathrm{d}_0 \exp(-\xi))A_0 \\
&\overset{(*)}{=} \mathrm{d}\xi + \mathrm{d}_{\mathbb{1}} \operatorname{Ad}(-\xi)A_0 \\
&= \mathrm{d}\xi + \operatorname{ad}(-\xi)A_0 \\
&= \mathrm{d}\xi - [\xi \wedge A_0]
\end{aligned}
$$

where $(*)$ follows because the differential of the exponential map at zero is the identity. Now, we must look at $\mathrm{d}^*\mathcal{G}(A_0, 0)\xi$ and $*\mathcal{G}(A_0, 0)\xi|_{\partial B}$. Note that because $\xi \in W^{2,p}_m(B, \mathfrak{g})$ is just a function, $*[\xi \wedge A_0] = [\xi \wedge *A_0]$, and remember that A_0 is in Uhlenbeck gauge already. Then we calculate

$$
\begin{aligned}
-\mathrm{d}^*[\xi \wedge A_0] &= -(- * \mathrm{d}*)[\xi \wedge A_0] = * \mathrm{d}[\xi \wedge *A_0] \\
&= *[\mathrm{d}\xi \wedge *A_0] + *[\xi \wedge \mathrm{d} * A_0] \\
&= *[\mathrm{d}\xi \wedge *A_0] + [\xi \wedge \underbrace{* \mathrm{d} * A_0}_{-\mathrm{d}^* A_0 = 0}] \\
&= *[\mathrm{d}\xi \wedge *A_0].
\end{aligned}
$$

Moreover, $*A_0|_{\partial B} = 0$ and so

$$
*(\mathrm{d}\xi - [\xi \wedge A_0])|_{\partial B} = * \mathrm{d}\xi|_{\partial B}.
$$

Therefore

$$
\partial_X F\big|_{A_0,0}\xi = (\mathrm{d}^* \mathrm{d}\xi + *[\mathrm{d}\xi \wedge *A_0], * \mathrm{d}\xi|_{\partial B}).
$$

We write $\partial_X F|_{A_0,0} = T + S$, where

$$
T = (\Delta\xi, * \mathrm{d}\xi|_{\partial B}) \quad \text{and} \quad S = (*[\mathrm{d}\xi \wedge *A_0], 0).
$$

Note that T is the operator for the inhomogeneous Neumann problem.[19] From theorem B.6, $\mathcal{Z}$ is exactly the space of functions (f, g) for which the Neumann problem has solution, and so T is surjective onto $\mathcal{Z}$; moreover, the solutions are unique up to additive constant, and so the additional condition on the domain $W^{2,p}_m(B, \mathfrak{g})$ that $\int_B \xi = 0$ makes T injective. Furthermore, by theorem B.7, the inverse of this operator is bounded for every metric $W^{2,\infty}$-close to the identity. Choosing an appropriate δ, we can make $\|T^{-1}\| \leq C_T$ for some

[19] See appendix B, the section on the Neumann problem.

constant C_T which is independent of the metric. Now, if we find a bound for S, lemma B.5 gives the bijectivity of $T+S$. Calculate

$$\begin{aligned}|*[\,\mathrm{d}\xi \wedge *A_0]| &= \left|* \sum [\partial_i \xi, A_j]\,\mathrm{d}x_i \wedge *\,\mathrm{d}x_j\right| = \left|* \sum [\partial_i \xi, A_j]\, g^{ij} * 1\right| \\ &= \left|\sum g^{ij} [\partial_i \xi, A_j]\right| \le \sum \left|g^{ij} [\partial_i \xi, A_j]\right| \\ &\le \sum \left|g^{ij}\right| |\partial_i \xi|\, |A_j| \le \underbrace{\max \left|g^{ij}\right|}_{=:|g^{-1}|} \sum |\partial_i \xi|\, |A_j| \\ &= \left|g^{-1}\right| |\,\mathrm{d}\xi|_{\mathbb{1}}\, |A_0|_{\mathbb{1}}\,.\end{aligned}$$

Then, because $|g^{-1}| \le \|g^{-1}\|_\infty$, we can use lemma B.4 and require $\delta \le \frac{1}{2}$ such that $|g^{-1}| \le (1 - \|g - \mathbb{1}\|_\infty)^{-1} \le 2$, and now we apply this in inequality $(*)$ below:

$$\begin{aligned}\|S\xi\|_{\mathcal{Z}} &= \|*[\,\mathrm{d}\xi \wedge *A_0]\|_{g;p} \\ &\le 2\, \|*[\,\mathrm{d}\xi \wedge *A_0]\|_{\mathbb{1};p} \\ &= 2 \left(\int_B |*_g[\,\mathrm{d}\xi \wedge *_g A_0]|_{\mathbb{1}}^p \right)^{\frac{1}{p}} \\ &\overset{(*)}{\le} 4 \left(\int_B (|\,\mathrm{d}\xi|_{\mathbb{1}}\, |A_0|_{\mathbb{1}})^p \right)^{\frac{1}{p}} \\ &= 4\, \||\,\mathrm{d}\xi|_{\mathbb{1}}\, |A_0|_{\mathbb{1}}\|_{\mathbb{1};p} \\ &\le 4C_{rnpq}\, \||\,\mathrm{d}\xi|_{\mathbb{1}}\|_{\mathbb{1};1,p}\, \||A_0|_{\mathbb{1}}\|_{\mathbb{1};1,q} \\ &\overset{(1)}{\le} 4C_{rnpq}\, \|\xi\|_{\mathbb{1};2,p}\, \|A_0\|_{\mathbb{1};1,q} \\ &\le 16 C_{rnpq}\, \|\xi\|_{2,p}\, \|A_0\|_{1,q}\,,\end{aligned}$$

where the factors of 2 come from the equivalence of the norms (2.6) on B, the constants C_{rnpq} come from lemma B.9 as usual, and in (1) we use lemma B.10 for $|\nabla |A|| \le |\nabla A|$ and that on sections $\xi : B \to \mathfrak{g}$, $\nabla = \mathrm{d}$ so $\nabla\,\mathrm{d}\xi = \nabla^2 \xi$,

Now, let $\tilde{C}$ be given and choose $\varepsilon = (32 C_{rnpq} \tilde{C} C_T)^{-q}$. Since A_0 is in Uhlenbeck gauge with energy $\mathcal{E}_q(A_0) \le \varepsilon$ and constant $\tilde{C}$, we can use $\|A_0\|_{1,q} \le \tilde{C}\, \|F_{A_0}\|_q \le \varepsilon^{\frac{1}{q}} \tilde{C}$ and at last

$$\|S\xi\|_{\mathcal{Z}} \le 16 C_{rnpq} \tilde{C} \varepsilon^{\frac{1}{q}}\, \|\xi\|_{2,p} = \frac{1}{2C_T}\, \|\xi\|_{2,p}\,.$$

This means that S and T satisfy the hypotheses from lemma B.5, namely that T is bijective with bounded inverse and $\|S\|\, \|T^{-1}\| = \frac{1}{2} < 1$, and therefore $\partial_X F|_{A_0,0} = T + S$ is bijective.

Finally, this means that F satisfies all conditions to the implicit function theorem B.3 around $(A_0, 0)$, and therefore there are neighbourhoods U around A_0 and V around 0, and a continuous map $f : U \to V$ such that $F(A, f(A)) = 0$, which implies that $\exp(f(A))^* A$ is in Uhlenbeck gauge for all $A \in U \subseteq \mathcal{A}^{1,p}(B, \mathfrak{g})$. Furthermore, to get the bound $\|X\|_{2,p} \le \lambda$, it suffices to take a ball with sufficiently small radius R within $f^{-1}(B_\lambda(0) \cap V) \subseteq U$, so that $f : B_R(A_0) \to B_\lambda(0)$. □

Lemma 2.7 fixes the constant $\tilde{C}$ we will use for the Uhlenbeck gauge, and this in turn fixes ε, which is defined in terms of $\tilde{C}$ in the proof of lemma 2.8. Then for a connection A_0 in Uhlenbeck gauge with constant $\tilde{C}$ and energy $\mathcal{E}_q \le \varepsilon$, we find R such that for every connection $A \in B_R(A_0) \subseteq \mathcal{A}^{1,p}(B)$, there is a gauge transformation $\exp(X)$ such that $(\exp X)^*A$ satisfies (2.2), the first condition of the Uhlenbeck gauge, and we get a bound $\|X\|_{2,p} \leq \lambda$.

All that is left to do now is check that these solutions (A, X) satisfy $\|\exp(X)^*A\|_r \leq \Lambda$, so that they can be shown to satisfy (2.3) by lemma 2.7 and finally all of $B_R(A_0)$ can be put into Uhlenbeck gauge with constant $\tilde{C}$. Therefore, we show this now: first, we have

$$\|\exp(X)^*A\|_r \leq \left\|\mathrm{Ad}_{\exp(X)} A\right\|_r + \|\exp(-X)\,\mathrm{d}\exp(X)\|_r .$$

Because the inner product on $\mathfrak{g}$ is Ad-invariant, the first term is just $\|A\|_r$ and we have

$$\begin{aligned}\|A\|_r &\leq \|A - A_0\|_r + \|A_0\|_r \\ &\overset{(*)}{\leq} 4C_W(\|A - A_0\|_{1,p} + \|A_0\|_{1,q}) \\ &\leq 4C_W(R + \tilde{C}\varepsilon),\end{aligned}$$

where $(*)$ follows from the equivalence of metrics (2.6) and the Sobolev embeddings $L^r \hookrightarrow W^{1,p}$ and $L^r \hookrightarrow W^{1,q}$. As for the second term, we can write

$$\exp(-X(x))\,\mathrm{d}\exp(X(x)) = \exp(-X(x))\,\mathrm{d}_X \exp \circ\, \mathrm{d}X$$

and note that, since $2p > n$, the embedding $W^{2,p} \hookrightarrow C^0$ provides control over the C^0-norm of X, $\sup_{x\in B} |X(x)| \leq \lambda$. Therefore, for sufficiently small λ, for every point $x \in B$ the map $\exp(-X(x))\,\mathrm{d}_{X(x)} \exp : \mathfrak{g} \to \mathfrak{g}$ is arbitrarily close to $\exp(0)\,\mathrm{d}_0 \exp = \mathrm{d}_0 \exp = \mathbb{1}_{\mathfrak{g}}$, which means that its norm can be bounded by 2. Thus,

$$\begin{aligned}\|\exp(-X)\,\mathrm{d}\exp(X)\|_r &= \|\exp(-X)\,\mathrm{d}_X \exp \circ\, \mathrm{d}X\|_r \\ &\leq 2\,\|\,\mathrm{d}X\|_r \leq 4\,\|\,\mathrm{d}X\|_{\mathbb{1};r} \\ &\leq 4C_W\,\|\,\mathrm{d}X\|_{\mathbb{1};1,p} \\ &\leq 8C_W\,\|\,\mathrm{d}X\|_{1,p} \\ &\leq 8C_W\,\|X\|_{2,p} \leq 8C_W\lambda.\end{aligned}$$

For the first term, $4C_W$ does not depend on the metric, and $\tilde{C}$ has already been fixed; for the second term, $8C_W$ does not depend on the metric, and so putting both terms together we have

$$\|\exp(X)^*A\|_r \leq c(R + \varepsilon + \lambda)$$

for some uniform constant c. Given $\tilde{C}$, we chose ε as a uniform constant which can be made smaller. The constant λ is arbitrary, and so once again can be made even smaller. Finally, while R depends on λ and the metric on B, we can put a uniform upper bound on it which can also be made smaller, and furthermore making λ smaller only makes R smaller as well. Then, making each term small as needed, we may bound $\|\exp(X)^*A\|_r$ by Λ for all metrics $\|g - \mathbb{1}\|_{1,\infty} \leq \delta$.

□

This ends the proof of the gauge fixing lemma on B. Let us recapitulate the choices of constants that were made during the proof of the theorem:

- For proving that S_ε is connected and closed, there was no need to fix anything.
- Pick a δ small enough such that if $\|g - \mathbb{1}\|_{1,\infty} \leq \delta$, there is equivalence of norms (2.6) with different metrics on B for $W^{1,p}$, $W^{1,q}$ and L^r with $r = r(n, p, q)$ from lemma B.9.
- In the implicit function step, pick δ possibly smaller to guarantee bijectivity of the derivative of the operator F defined. Then, for any $\tilde{C}$ given, independent of the metric, choose ε appropriately, which will also be uniform. For each metric and every λ, there is $R(\lambda, g)$ such that $\|X\|_{2,p} \leq \lambda$ for $X = f(A)$ and $\|A - A_0\|_{1,p} \leq R$.
- The Uhlenbeck gauge constant $\tilde{C}$ is fixed with the *a priori* estimates. Along with it, δ is picked once again for an estimate from theorem 1.26 to hold with uniform constant, and we find a bound Λ on the L^r-norm of the connection which guarantees that these *a priori* estimates hold.
- Finally, when checking that the solution to the implicit function problem has L^r-bound Λ, we find that it is bounded by $R(\lambda, g) + \varepsilon + \lambda$, where λ can be chosen arbitrarily, ε was already a uniform constant which can be made smaller, and we may choose a uniform bound R for $R(\lambda, g)$ such that the λ bound holds for each metric and a given λ. Then R, ε and λ can be made small enough for the Λ bound to hold.

Chapter 3

Weak and strong compactness

Compactness is an extremely important and ubiquitous notion. From a first course in real analysis, one learns that a bounded sequence on the real line has a convergent subsequence, and that in Euclidean space it is equivalent for a set to be compact, sequentially compact, or closed and bounded. While a uniform bound is a necessary condition for convergence, in more general contexts is it not usually sufficient, and so many important theorems have emerged to solve this problem. For instance, on infinite dimensional vector spaces, the closed unit ball is no longer compact; weaker topologies come into play, and the Banach–Alaoglu theorem provides compactness. On metric spaces, there is the notion of equicontinuity and the Arzelà–Ascoli theorem. One of the fundamental results in functional analysis is the uniform boundedness principle.

Other than the gauge fixing lemma, the importance of Uhlenbeck's *Connections with L^p bounds on curvature* [Uhl82] comes from a compactness theorem. As is clear from the title, she was looking at sequences of connections with uniform bound on curvature. Therefore, in this case there is not even a uniform bound on the sequence *a priori*. The gauge fixing lemma is the key ingredient used to translate the bound on the curvatures to a uniform bound on the connections; it also deals with the problem of gauge redundancy. Then, the weak compactness theorem 3.5 states that a sequence of connections with uniformly bounded curvatures has a weakly convergent subsequence, modulo gauge transformations. The strong compactness theorem 3.9, on the uniform convergence of sequences of Yang–Mills connections with uniform bounds on curvature, quickly follows.

The goal of this chapter is to prove these two theorems. First, we begin by commenting on compactness for geodesics and flat connections. The case of the geodesics is simple and instructive. The case of the flat connections is the opposite: while these connections do have a uniform bound on curvature (namely, zero), the usual proof of the compactness result for flat connections does not rely on estimates, instead using an identification with a specific class of homomorphisms. We include a sketch of this proof here for completeness, but also to show that not all compactness theorems emerge in the same way. Finally, note that the weak compactness theorem can be seen as a generalization of the result for flat connections.

Geodesics

Let (M, g) be a Riemannian manifold, and consider curves $\gamma : [0, 1] \to M$ with fixed endpoints $\gamma(0), \gamma(1) \in M$. We wish to show the existence of a geodesic between $\gamma(0)$ and

$\gamma(1)$ minimising length, i.e. a critical point of the length functional

$$\mathcal{L}(\gamma) = \int_0^1 |\nabla\gamma| \, \mathrm{d}t.$$

This will follow from a compactness argument. First, observe that a geodesic is also a critical point of the L^2-energy

$$\mathcal{E}(\gamma) = \int_0^1 |\nabla\gamma|^2 \, \mathrm{d}t.$$

For $a, b \in [0,1]$, the distance between two points on the curve is

$$\begin{aligned} \mathrm{d}(\gamma(a), \gamma(b)) = \int_a^b |\nabla\gamma| \, \mathrm{d}t &\overset{(*)}{\leq} \left|\int_a^b 1 \, \mathrm{d}t\right|^{\frac{1}{2}} \left|\int_a^b |\nabla\gamma|^2 \, \mathrm{d}t\right|^{\frac{1}{2}} \\ &\leq |a-b|^{\frac{1}{2}} \, \mathcal{E}(\gamma)^{\frac{1}{2}}, \end{aligned}$$

where $(*)$ is the Cauchy-Schwartz inequality. Using this now it is easy to see how uniform bounds on energy can lead to compactness properties. Let $\{\gamma_i\}_{i\in\mathbb{N}}$ be a minimising sequence of paths $\gamma_i : [0,1] \to M$ with fixed endpoints and $\mathcal{E}(\gamma_i) < C$. Looking at (M, g) as a metric space with this distance as the metric, we have that

$$\mathrm{d}(\gamma_i(a), \gamma_i(b)) \leq C \, |a-b|^{\frac{1}{2}},$$

which makes this sequence equicontinuous[1]. Therefore, by the Arzelà–Ascoli theorem, there is a uniformly convergent subsequence, and the limit is a geodesic.

Flat connections

The previous example had a very clear procedure: use estimates and uniform bounds to enter the setting of a well-known compactness theorem. We now come back to gauge theory and look at flat connections on trivial vector bundles. It is a very classical result that the moduli space of flat G-connections is compact, but the usual way to prove it does not follow the same procedure. Instead, it hinges on the following lemma.

Lemma 3.1 ([Cra15], corollary 1.29)**.** *Let $E \to M$ be a vector bundle with flat connection ∇. Then for any $x, y \in M$, the induced parallel transport from E_x to E_y only depends on the homotopy class of the path from x to y.*

A connection is called G-connection if, for instance, the associated parallel transports induce isomorphisms of the fibres which are in G as a subgroup of $GL(\mathbb{R}^m)$ (this will always be the case for compact G), where m is the rank of the bundle. With this lemma, we can write the following identification.

Proposition 3.2 ([DK97], proposition 2.2.3)**.** *There is a one-to-one correspondence between conjugacy classes of homomorphisms $\rho : \pi_1(M) \to G$ and gauge equivalence classes of flat G-connections on M.*

Sketch of proof. Let ∇ be a flat G-connection, then the associated parallel transport induces a representation $\rho^\nabla : \pi_1(M, x) \to G$ by defining, for each loop γ based at $x \in M$,

$$\rho^\nabla(\lfloor\gamma\rceil) = P_\gamma^\nabla \in G \subseteq GL(E_x).$$

□

[1] For $\varepsilon > 0$, take $\delta = (\varepsilon/C)^2$.

If we prove compactness of the space of homomorphisms $\rho : \pi_1(M) \to G$ when G is compact, then we are done.

Lemma 3.3. *Let G be a compact Lie group, M a smooth manifold. Then $\frac{\mathrm{Hom}(\pi_1(M),G)}{G}$ is compact.*

Sketch of proof. We prove that $\mathrm{Hom}(\pi_1(M), G)$ is closed in $G^{\pi_1(M)}$ with the product topology, which is compact by Tychonoff's theorem and because G is compact. Then $\mathrm{Hom}(\pi_1(M), G)$ is compact, as is quotient.

Let $f : \pi_1(M) \to G$ be a function which is not a homomorphism: there exist $r, s \in \pi_1(M)$ such that $f(r)f(s) \neq f(rs)$. Take $U_{f(r)f(s)}, U_{f(rs)}$ neighbourhoods of $f(r)f(s)$ and $f(rs)$ which do not intersect. Since G is a topological group, the multiplication is continuous and the preimage of $U_{f(r)f(s)}$ gives neighbourhoods of $f(r)$ and $f(s)$, $U_{f(r)}$ and $U_{f(s)}$, respectively. Now, letting $p_r : G^{\pi_1(M)} \to G$ be the continuous projection $p_r(f) = f(r)$, we construct a neighbourhood of f on $G^{\pi_1(M)}$,

$$\mathcal{U} = p_r^{-1}(U_{f(r)}) \cap p_s^{-1}(U_{f(s)}) \cap p_{f(rs)}^{-1}(U_{f(rs)}),$$

such that for any $g \in \mathcal{U}$, g cannot be a homomorphism: $g(rs) \in U_{f(rs)}$ and $g(r)g(s) \in U_{f(r)f(s)}$, and therefore $g(rs) \neq g(r)g(s)$. Thus, the complement of the homomorphisms is open and consequently $\mathrm{Hom}(\pi_1(M), G)$ is closed. □

Finally, the moduli space of flat G-connections is the quotient $\mathcal{A}_{flat}/G$, for $\mathcal{A}_{flat} \subseteq \mathcal{A}$ the subspace containing flat G-connections, and we have the result:

Corollary 3.4. *The moduli space of flat G-connections on M is compact.*

3.1 Weak compactness

Let M be a compact n-manifold with (possibly empty) boundary, and let $P \to M$ be a principal G-bundle. The second main result we wish to prove, theorem 3.6 in [Uhl82], is the following:

Theorem 3.5 (Weak compactness). *Let $\frac{n}{2} < p < \infty$. A sequence of connections in $\mathcal{A}^{1,p}(P)$ with uniform L^p-bound on the sequence of curvatures has a subsequence which is gauge equivalent to a weakly convergent sequence, with gauge transformations in $\mathcal{G}^{2,p}(P)$.*

For the proof of this theorem we are essentially looking for a sequence of global gauges such that the gauged connections are uniformly $W^{1,p}$-bounded and will therefore, by the Banach–Alaoglu theorem, converge weakly. For that we will need the gauge fixing theorem 2.2, and a patching lemma 3.10 which we will prove later, but roughly goes as follows.

Lemma 3.6 (cf. lemma 3.10). *Two sets of transition functions describe isomorphic bundles if they are C^0-close. Furthermore, there are bounds on the gauge transformations relating these transition functions; for a sequence of transition functions and transformations, the bounds are uniform.*

The gauge fixing lemma will give us a bundle atlas of $P \to M$ such that the local connections are in Uhlenbeck gauge, and therefore the uniform bound on the curvature will translate into a uniform bound on the connections. The gauge transformations resulting from the lemma do not necessarily patch up to a global gauge transformation, however they do yield new transition functions for a sequence of bundles. While it will be easy to see that these bundles will all be isomorphic to each other and the original bundle, the obvious changes to the gauge transformations mess with the uniform bounds, and thus the patching lemma will be needed in order to find less straightforward but more suitable isomorphisms between the bundles, which then can be modified in a uniform way to yield globally defined uniformly bounded gauge transformations. To get uniform bounds on the gauge transformations, we will also need the following.

Lemma 3.7 (cf. lemmata 1.33 and 1.34)**.** *Bounds on connection forms give bounds on the gauge transformations relating them, and moreover if two sequences of connections are uniformly bounded, then the sequence of gauge transformations relating them has a C^0-convergent subsequence.*

Proof of weak compactness theorem. Let $(A^i)_{i\in\mathbb{N}} \subseteq \mathcal{A}^{1,p}(P)$ be a sequence such that $\|F_{A^i}\|_p$ is uniformly bounded. Choose $q < p$ such that it satisfies the hypotheses of the gauge fixing lemma 2.2. We can bound the L^q-energy

$$\mathcal{E}_q(A^i|_U) = \|F_{A^i}\|^q_{q(U)} \overset{(1)}{\leq} (\operatorname{Vol} U)^{\frac{p-q}{p}} \|F_{A^i}\|^q_{p(U)} \leq (\operatorname{Vol} U)^{1-\frac{q}{p}} \|F_{A^i}\|^q_p ,$$

where (1) is the Hölder inequality for $\frac{1}{q} = \frac{1}{p} + \frac{p-q}{pq}$. With this expression we can make $\mathcal{E}_q(A^i|_U) \leq \tilde{\varepsilon}$ for U of sufficiently small volume, and this is why it is important that $q < p$ strictly since we need $1 - \frac{q}{p} > 0$ to make the L^q-energy small. Now we are in the setting of 2.2, and all the A^i may be put in Uhlenbeck gauge on open sets which cover M; since M is compact, we can take a finite subcollection $M = \bigcup_{\alpha=1}^{N} U_\alpha$.

These U_α form a bundle atlas for $P \to M$, and on each U_α the connections are represented by connection matrices $A^i_\alpha \in \mathcal{A}^{1,p}(U_\alpha)$. Since $\mathcal{E}_q(A^i_\alpha) = \mathcal{E}_q(A^i|_{U_\alpha}) \leq \tilde{\varepsilon}$, there exist $u^i_\alpha \in \mathcal{G}^{2,p}(U_\alpha)$ such that $u^{i*}_\alpha A^i_\alpha$ is in Uhlenbeck gauge, and in particular $\|u^{i*}_\alpha A^i_\alpha\|_{1,p} \leq \tilde{C} \left\|F_{A^i_\alpha}\right\|_p$ is uniformly bounded. This is sufficient to find weakly convergent subsequences on each U_α, however the u_α do not necessarily define a global gauge transformation. For that to be the case we need

$$u^i_{\alpha\beta} := (u^i_\alpha)^{-1}\phi_{\alpha\beta}u^i_\beta \tag{3.1}$$

to be identical to $\phi_{\alpha\beta}$, the transition functions of the bundle atlas, see (1.18) and the discussion in subsection 1.2.1.[2] Therefore, the next step in the proof is modifying the u^i_α appropriately to achieve this.

In order to use the patching lemma 3.10 we need the transition functions to be C^0-close to each other. To see this, write

$$\begin{aligned} u^{i}_{\alpha\beta}{}^{*}(u^{i*}_\alpha A^i_\alpha) &= (u^i_\alpha u^i_{\alpha\beta})^* A^i_\alpha = (\phi_{\alpha\beta}u^i_\beta)^* A^i_\alpha = u^{i*}_\beta(\phi^*_{\alpha\beta}A^i_\alpha) \\ &= u^{i*}_\beta A^i_\beta, \end{aligned} \tag{3.2}$$

and since all $u^{i*}_\alpha A^i_\alpha$ are uniformly bounded for any $\alpha \in A$, lemma 1.33 tells us that the gauge transformations relating these connection forms, $u^i_{\alpha\beta}$, are also uniformly bounded, and furthermore there is a subsequence of the $u^i_{\alpha\beta}$ (also labelled $i \in \mathbb{N}$) that converges in C^0. Thus, for a further subsequence, and for each $\alpha, \beta = 1, \ldots, N$, all the $u^i_{\alpha\beta}$ can be made

[2]Note that if we look at the gauge transformations u^i_α as local changes of trivialization, then the $u^i_{\alpha\beta}$ are new transition functions for an isomorphic bundle for each $i \leq N$; even so, this is not enough.

to be within a geodesic δ-ball of each other for any $\delta > 0$; in particular we can single out the first element of this subsequence $g_{\alpha\beta} := u^1_{\alpha\beta}$ and denote also $g_\alpha := u^1_\alpha$, and we have [3] [4]

$$\mathrm{d}(u^i_{\alpha\beta}, g_{\alpha\beta}) \leq \delta,$$

Picking $\delta = \Delta_{\exp}$ the radius of a convex geodesic ball on G, the patching lemma 3.10 gives us a refinement $V_\alpha \subseteq U_\alpha$ of the original cover and new gauge transformations $h^i_\alpha : V_\alpha \to G$ with uniform bounds and such that

$$(h^i_\alpha)^{-1} u^i_{\alpha\beta} h^i_\beta = g_{\alpha\beta}.$$

With these new gauge transformations we can now modify the original u^i_α. Define $\tilde{u}^i_\alpha := u^i_\alpha h^i_\alpha g_\alpha$ on V_α, and note that this defines a global gauge transformation, as $M = \bigcup_{\alpha=1}^N V_\alpha$ and on $V_\alpha \cap V_\beta$

$$(\tilde{u}^i_\alpha)^{-1} \phi_{\alpha\beta} \tilde{u}^i_\beta = g_\alpha (h^i_\alpha)^{-1} \underbrace{(u^i_\alpha)^{-1} \phi_{\alpha\beta} u^i_\beta}_{u^i_{\alpha\beta}} h^i_\beta g_\beta^{-1} = g_\alpha g_{\alpha\beta} g_\beta^{-1} = \phi_{\alpha\beta}. \tag{3.3}$$

Moreover, $\tilde{u}^i$ as defined by the local $\tilde{u}^i_\alpha$ is in $\mathcal{G}^{2,p}(P)$, as lemma 3.10 yields $h^i_\alpha \in \mathcal{G}^{2,p}(V_\alpha)$, we had $u^i_\alpha \in \mathcal{G}^{2,p}(V_\alpha)$ from the start and $\mathcal{G}^{k,p}$ is closed under group multiplication for $kp > n$, which is the case since $p > \frac{n}{2}$.

It remains to prove that $\tilde{u}^{i*}_\alpha A^i_\alpha$ is uniformly bounded in $\mathcal{A}^{1,p}(V_\alpha)$ for all $\alpha = 1, \ldots, N$. This follows easily from lemma 1.31, which in this case states that for $A^i \in \mathcal{A}^{1,p}(V)$ and $u^i \in \mathcal{G}^{2,p}(V)$ for some trivializing neighbourhood V, the following holds [5]

$$\left\| u^{i*} A^i \right\|_{1,p} \leq \left\| (u^i)^{-1} \,\mathrm{d} u^i \right\|_{1,p} + c \left\| A^i \right\| \left(1 + C_W \left\| (u^i)^{-1} \,\mathrm{d} u^i \right\|_{1,p} \right), \tag{3.4}$$

where c, C_W are constants. Then writing

$$\tilde{u}^{i*}_\alpha A^i_\alpha = (g_\alpha^{-1})^* h^{i*}_\alpha u^{i*}_\alpha A^i_\alpha,$$

we first note that $h^{i*}_\alpha (u^{i*}_\alpha A^i_\alpha)$ is uniformly bounded: $\|u^{i*}_\alpha A^i_\alpha\|_{1,p}$ is bounded uniformly by the uniform bound on the curvature (because of the Uhlenbeck gauge), and $\|(h^i_\alpha)\,\mathrm{d}h^i_\alpha\|_{1,p}$ is uniformly bounded by the patching lemma and the fact that $\left\|(u^i_{\alpha\beta})^{-1}\,\mathrm{d}u^i_{\alpha\beta}\right\|_{1,p}$ has a uniform bound (lemma 1.33). Using (3.4) again, the uniform bound on $h^{i*}_\alpha u^{i*}_\alpha A^i_\alpha$ and the fact that g_α^{-1} is independent of $i \in \mathbb{N}$, we get a uniform $W^{1,p}$-bound on $\tilde{u}^{i*}_\alpha A^i_\alpha$.

Finally, we can use the Banach–Alaoglu theorem 1.18 to guarantee that for every $\alpha = 1, \ldots, N$, the sequence $\tilde{u}^{i*}_\alpha A^i_\alpha$ has a $W^{1,p}$-weakly convergent subsequence, and because the $\tilde{u}^i$ are global gauges, we can choose the same [6] subsequence for all (finite) α, which finally gives us a weakly convergent subsequence of $\tilde{u}^{i*} A^i$ in $\mathcal{A}^{1,p}(P)$. □

[3]See subsection 1.2.2 for the definitions of a geodesic convex ball and this metric.

[4]Also note that the first element $g_{\alpha\beta}$ can only be fixed after the choice of δ.

[5]We used the Sobolev embedding $W^{1,p} \hookrightarrow L^{2p}$, note that $1 - \frac{n}{p} \geq -\frac{n}{2p}$.

[6]For $\alpha = 1$, there is a convergent subsequence; this subsequence is also uniformly bounded on $\alpha = 2$, and so some further subsequence converges. By repeating this process until you find a convergent subsequence for $\alpha = N$, this same subsequence can be used for all α. If there were countably many α, a similar argument would work by taking the diagonal of the subsequences.

Remark 3.8. Note that the choice of $\tilde{u}^i_\alpha$ makes sense. In finding a global gauge, the simplest thing to do would be to take $\tilde{u}^i_\alpha = \mathbb{1}_G$, and in order to keep the original u^i_α around for the Uhlenbeck gauge, it would have sufficed to have $\tilde{u}^i_\alpha = \mathbb{1}_G = u^i_\alpha (u^i_\alpha)^{-1}$; however, there is no uniform bound on $(u^i_\alpha)^{-1}$. In the patching lemma, it is already clear that $h^i_\alpha = (u^i_\alpha)^{-1} g_\alpha$ would have sufficed for the isomorphism, and if this could be guaranteed to have a uniform bound in this form then one could naturally write

$$\tilde{u}^i_\alpha = \mathbb{1}_G = u^i_\alpha \underbrace{(u^i_\alpha)^{-1} g_\alpha}_{h^i_\alpha} g_\alpha^{-1}.$$

Thus it makes sense to simply substitute the more complicated h^i_α found in the proof of the patching lemma in the expression above when choosing a uniformly bounded globally defined gauge transformation.

3.2 Strong compactness

While the strong compactness theorem is attributed to Uhlenbeck, it is not stated in her 1982 papers. Nevertheless, it is readily adapted from the proof of the weak compactness theorem. The only additional ingredient is a result on regularity of Yang–Mills connections in Coulomb gauge, a consequence of the gauge fixing lemma which she also stated in her paper [Uhl82] as corollary 1.4. We show the adaptation of the proof, sketching the points which are already explained in detail in the proof of theorem 3.5.[7]

Theorem 3.9 (Strong compactness). *Let M be a compact Riemannian n-manifold with (possibly empty) boundary, and let $1 < p < \infty$ such that $p > \frac{n}{2}$ and in case $n = 2$, $p \geq \frac{4}{3}$. Suppose a sequence of connections $(A^i)_{i\in\mathbb{N}} \subseteq \mathcal{A}^{1,p}(P)$ is such that the A^i are weak Yang–Mills connections and $\|F_{A^i}\|_p$ is uniformly bounded. Then there exists a subsequence (with same label $i \in \mathbb{N}$) and a sequence of gauge transformations $(u^i)_{i\in\mathbb{N}} \subseteq \mathcal{G}^{2,p}(P)$ such that $u^{i*}A^i$ converges strongly with all derivatives to a smooth Yang–Mills connection.*

Note that the assumptions on p in this theorem are stricter than in the weak compactness theorem; the reason for that lies in the definition of the weak Yang–Mills connections, see (1.30).

Proof. Let $A^i \in \mathcal{A}^{1,p}(P)$ be as in the statement of the theorem. Choose q to satisfy the hypotheses of the gauge fixing lemma, then there is a finite cover $M = \bigcup_{\alpha=1}^N U_\alpha$ such that $\mathcal{E}_q(A^i_\alpha) \leq \tilde{\varepsilon}$, and therefore there exist sequences of gauge transformations $(u^i_\alpha)_{i\in\mathbb{N}} \subseteq \mathcal{G}^{2,p}(U_\alpha)$ on each U_α such that from the uniform bound on $\|F_{A^i}\|_p$ we get a uniform bound on $\|u^{i*}_\alpha A^i_\alpha\|_{1,p}$ for each α. Since the A^i are weak Yang–Mills connections, then so are the $u^{i*}_\alpha A^i_\alpha$ (lemma 1.36), and therefore from the regularity theorem 1.39 we find uniform bounds on $\|u^{i*}_\alpha A^i_\alpha\|_{k,p}$ for all $k \in \mathbb{N}$.

Once again we look at the transition functions $u^i_{\alpha\beta}$ as defined in (3.1), where the $\phi_{\alpha\beta}$ are the transition functions for our bundle. From (3.2) and lemma 1.34 we see $u^i_{\alpha\beta} \in \mathcal{G}^{k,p}$ for all k, and then from lemma 1.33 we find uniform bounds on $\left\|(u^i_{\alpha\beta})^{-1}\, \mathrm{d}u^i_{\alpha\beta}\right\|_{k,p}$ for all k, and a subsequence of the $u^i_{\alpha\beta}$ that converges in C^0 which can be taken the same subsequence for

[7] Throughout this proof we once again keep taking subsequences and relabelling them the same as the original sequence, with $i \in \mathbb{N}$.

all $\alpha, \beta = 1, \dots, N$. Therefore there is some i such that all the $u^i_{\alpha\beta}$ are within a Δ_{exp} sized C^0-ball, for Δ_{exp} the radius of a convex geodesic ball on G, and we take this i to be the first element of the sequence. Now, instead of fixing u^1_α as we did for the proof of 3.5, we will need to take smooth $g_\alpha \in \mathcal{G}(U_\alpha)$ that are C^0-close to u^1_α for each α, which will imply that $g_{\alpha\beta} := g_\alpha^{-1}\phi_{\alpha\beta}g_\beta$ will also be C^0-close to the $u^i_{\alpha\beta}$.

We apply the patching lemma 3.10 to find a refinement $V_\alpha \subseteq U_\alpha$ of the original cover and $h^i_\alpha \in \mathcal{G}^{k,p}(V_\alpha)$ such that

$$(h^i_\alpha)^{-1}u^i_{\alpha\beta}h^i_\beta = g_{\alpha\beta}$$

on $V_\alpha \cap V_\beta$ and such that there are uniform bounds on $\|(h^i_\alpha)^{-1}\,\mathrm{d}h^i_\alpha\|_{k,p(V_\alpha)}$ for all $k \in \mathbb{N}$. Then $\tilde{u}^i_\alpha := u^i_\alpha h^i_\alpha g_\alpha^{-1}$ will patch to global gauge transformations as in (3.3), and the inequality in lemma 1.31 will take the form [8]

$$\left\|u^{i\,*}A^i\right\|_{k,p} \leq \left\|(u^i)^{-1}\,\mathrm{d}u^i\right\|_{k,p} + c\left\|A^i\right\|\left(1 + C_W\left\|(u^i)^{-1}\,\mathrm{d}u^i\right\|_{k,p}\right)^k,$$

slightly more complicated than (3.4), which we use to find uniform $W^{k,p}$-bounds on $\tilde{u}^{i\,*}_\alpha A^i_\alpha = (g_\alpha^{-1})^* h^{i*}_\alpha u^{i\,*}_\alpha A^i_\alpha$ for all $k \in \mathbb{N}$, on each V_α. Here it is important that the g^i_α are smooth in order to preserve the bounds for all k.

Finally, having obtained uniform bounds on the $W^{k,p}$-norms of the local gauged connections for all $k \in \mathbb{N}$, we may use the Arzelà–Ascoli theorem to find subsequences which converge uniformly with all derivatives, and then take the same subsequence on all V_α such that the $\tilde{u}^i_\alpha$ patch to a global $\tilde{u}^i$ and $\tilde{u}^{i\,*}A^i$ converges uniformly with all derivatives to some smooth connection $\tilde{A} \in \mathcal{A}(P)$. Then, from lemmata 1.36 and 1.38, $\tilde{u}^i A^i$ is weak Yang–Mills for each $i \in \mathbb{N}$ and the limit connection $\tilde{A}$ will also be weak Yang–Mills; since it is smooth, it is a Yang–Mills connection. □

In the book [Weh04], another approach is used to prove the strong compactness theorem, due to Dietmar Salamon. Rather than adapt the proof of the weak compactness theorem, it relies on a local slice theorem and then applies theorem 3.5 directly. Observe, however, that it needs strict inequality $p > \frac{4}{3}$ in case $n = 2$, as there will only be weak convergence of the connections, see lemma 1.38.(ii).

3.3 Patching

Finally, we prove the patching lemma used in the proofs of the weak and strong compactness theorems.

Lemma 3.10 (Patching lemma)**.** *Let M be an n-manifold, $p > \frac{n}{2}$ and let $M = \bigcup_{\alpha\in\mathbb{N}} U_\alpha$ be a locally finite open cover by precompact [9] sets. Then there is a refinement $V_\alpha \subseteq U_\alpha$ such that the following holds:*

(i) Let $k \in \mathbb{N}$ and let $g_{\alpha\beta}$, $h_{\alpha\beta} \in \mathcal{G}^{k+1,p}(U_\alpha \cap U_\beta)$ be two sets of transition functions for some principal G-bundle over M such that

$$\mathrm{d}(g_{\alpha\beta}, h_{\alpha\beta}) \leq \Delta_{\text{exp}}, \quad \forall \alpha, \beta \in \mathbb{N}. \tag{3.5}$$

[8] $W^{k-1,2p} \hookrightarrow W^{k,p}$ since $p > \frac{n}{2}$.

[9] Closure is compact.

Then there exist gauge transformations $h_\alpha \in \mathcal{G}^{k+1,p}(V_\alpha)$ for all $\alpha \in \mathbb{N}$ such that on all intersections $V_\alpha \cap V_\beta$

$$h_\alpha^{-1} h_{\alpha\beta} h_\beta = g_{\alpha\beta}. \tag{3.6}$$

(ii) Let $K \geq 2$ be an integer or $K = \infty$. If the $h_{\alpha\beta}$ in (i) run through a sequence $h^i_{\alpha\beta}$ of sets of transition functions such that $h^i_{\alpha\beta}$, $g_{\alpha\beta} \in \mathcal{G}^{k+1,p}(U_\alpha \cap U_\beta)$ for all $k < K$, and moreover for all $\alpha, \beta \in \mathbb{N}$ and $k < K$ there is a uniform bound on

$$\left\|(h^i_{\alpha\beta})^{-1}\, \mathrm{d}h^i_{\alpha\beta}\right\|_{k,p(U_\alpha \cap U_\beta)}.$$

Then the gauge transformations h^i_α in (i) satisfy, for all $\alpha \in \mathbb{N}$ and $k < K$, $h^i_\alpha \in \mathcal{G}^{k+1}(V_\alpha)$ and

$$\sup_{i\in\mathbb{N}} \left\|(h^i_\alpha)^{-1}\, \mathrm{d}h^i_\alpha\right\|_{k,p(V_\alpha)} < \infty, \tag{3.7}$$

which is to say that these norms are uniformly bounded for each α and k.

Proof. Note that we assume the cover is countable. We will first prove (i) by **induction on the cover**, and then for a sequence of the constructed h^i_α constructed in (i), regularity will follow directly and the uniform bounds will be proved by another induction on the cover.

For (i), on each step $j \in \mathbb{N}$, we will construct h_j on $V_j := U_j$, while changing some (or none) of the previous V_α already constructed for $\alpha \leq j-1$ and keeping h_α the same, albeit with possibly smaller domain. For each V_α, the process will end in finitely many steps, because as will be seen during the construction, on the j-th step a certain V_α will only be modified if $V_\alpha \cap U_j \neq \varnothing$, and the cover is locally finite. Moreover, the V_α will not depend on $k \in \mathbb{N}$ (from the Sobolev exponent) or the transition functions $h_{\alpha\beta}, g_{\alpha\beta}$. For the h_j constructed, we will want three conditions to be satisfied at each step j:

(1) the one we are trying to achieve, that is, condition (3.6),

$$h_\alpha^{-1} h_{\alpha\beta} h_\beta = g_{\alpha\beta},$$

on $V_\alpha \cap V_\beta$ for all $\alpha, \beta \leq j$,

(2) a technical condition which will be important within the construction,

$$\mathrm{d}(h_{i\alpha} h_\alpha g_{\alpha i}, \mathbb{1}) \leq \Delta_{\mathrm{exp}}, \tag{3.8}$$

on $V_\alpha \cap U_i$ for all $\alpha \leq j$ and $i \geq j$,

(3) and regularity, $h_\alpha \in \mathcal{G}^{k+1}(V_\alpha)$ for all $\alpha \leq j$.

For the first step, $j = 1$, we take $V_1 := U_1$ and $h_1 := \mathbb{1}$. Conditions (1) and (3) are trivially satisfied [10] and for all $i \geq 1$, $\mathrm{d}(h_{i1} h_1 g_{1i}, \mathbb{1}) = \mathrm{d}(g_{1i}, h_{1i}) \leq \Delta_{\mathrm{exp}}$ on $V_1 \cap U_i$ by assumption, therefore condition (2) is also satisfied.

At an arbitrary j-th step, for the induction hypothesis we will assume that for all $\alpha \leq j-1$ we have constructed $V_\alpha \subseteq U_\alpha$ and h_α such that there is still an open cover of

[10] Remember $h_{\alpha\alpha} = g_{\alpha\alpha} = \mathbb{1}$.

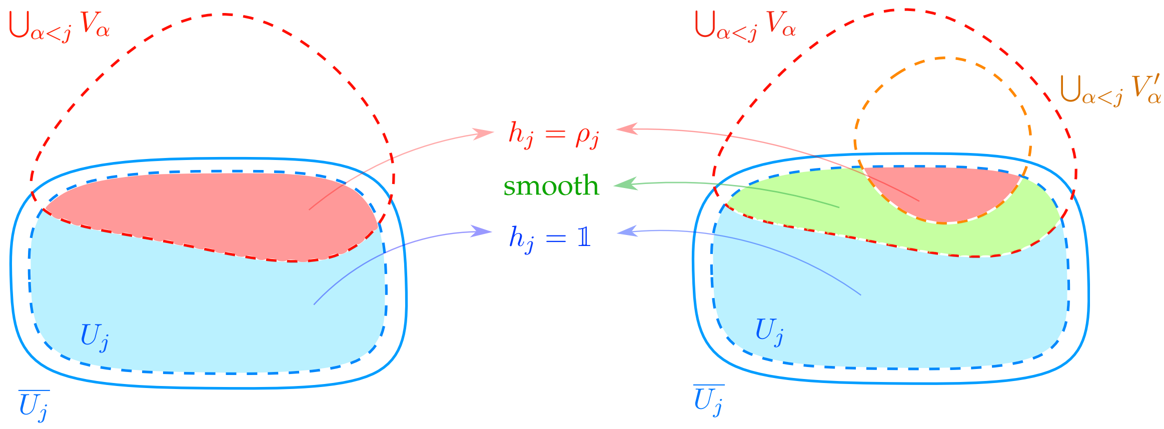

Figure 3.1: A sketch of the relevant sets on the j-th step, illustrating the reason to make the V_α into smaller V'_α : so that the h_j will be appropriately smooth.

M and the h_α respect conditions (1)–(3). That is, $M = \bigcup_{\alpha<j} V_\alpha \cup \bigcup_{\alpha\geq j} U_\alpha$, and conditions (1)–(3) were met on the $(j-1)$-th step.

Then, for $\alpha = j$, we will take $V_j := U_j$ and construct h_j such that (1)–(3) are satisfied. We will see that there is a natural choice of h_j on the intersection with the V_α which will automatically yield condition (1). We could then take $h_j = \mathbb{1}$ on $U_j \setminus \bigcup_{\alpha<j} V_\alpha$ and be done, but then h_j might not even be continuous, let alone in $\mathcal{G}^{k+1}(U_j)$. The idea, then, is to write $h_j := \exp(\psi_j \xi_j)$ for some $\xi_j \in \mathfrak{g}$ and a cutoff function ψ_j. Condition (2) is what is needed to have a well defined ξ_j, and the V_α will be modified so that there can be two disjoint compact sets where ψ_j will take the values 0 and 1.

Consider $\rho_j : U_j \cap \bigcup_{\alpha<j} V_\alpha \to G$ given by $\rho_j := h_{j\alpha} h_\alpha g_{\alpha j}$ on $U_j \cap V_\alpha$. This is well defined, since (3.6) is satisfied for $\alpha, \beta < j$ by assumption (condition (1)) and so on intersections $U_j \cap V_\alpha \cap V_\beta$

$$h_{j\beta} h_\beta g_{\beta j} = (h_{j\alpha} h_{\alpha\beta}) h_\beta (g_{\beta\alpha} g_{\alpha j}) = h_{j\alpha} h_\alpha g_{\alpha j}.$$

This is exactly what we need h_j to be on the intersections $U_j \cap V_\alpha$, as

$$h_\alpha^{-1} h_{\alpha j} (h_{j\alpha} h_\alpha g_{\alpha j}) = g_{\alpha j}. \tag{3.9}$$

As a product of $\mathcal{G}^{k+1,p}$ maps,[11] $\rho \in \mathcal{G}^{k+1,p}(U_j \cap \bigcup_{\alpha<j} V_\alpha)$, and moreover condition (2) for $\alpha \leq j-1$ guarantees that $\mathrm{d}(\rho, \mathbb{1}) \leq \Delta_{\exp}$, and therefore ρ takes values in the convex geodesic ball $B_{\Delta_{\exp}}(\mathbb{1})$. Therefore, there exists $\xi_j : W^{k+1,p}(U_j \cap \bigcup_{\alpha<j} V_\alpha) \to \mathfrak{g}$ such that $\rho_j = \exp(\xi_j)$.

Next, for $\alpha < j$, we replace the V_α by possibly smaller $V'_\alpha \subseteq V_\alpha$, so that we can take

$$h_j = \begin{cases} \rho_j = \exp(\xi_j) & \text{on } U_j \cap \bigcup_{\alpha<j} V'_\alpha, \\ \mathbb{1} = \exp(0) & \text{on } U_j \setminus \bigcup_{\alpha<j} V_\alpha \end{cases}$$

in a $W^{k+1,p}$-smooth way. Changing the domains V_α will not interfere with the induction hypothesis, as conditions (1)–(3) remain valid when diminishing the domain of the h_α, so long as we are still left with an open cover of M.

[11] By assumption for the transition functions and by the induction hypothesis for the h_α.

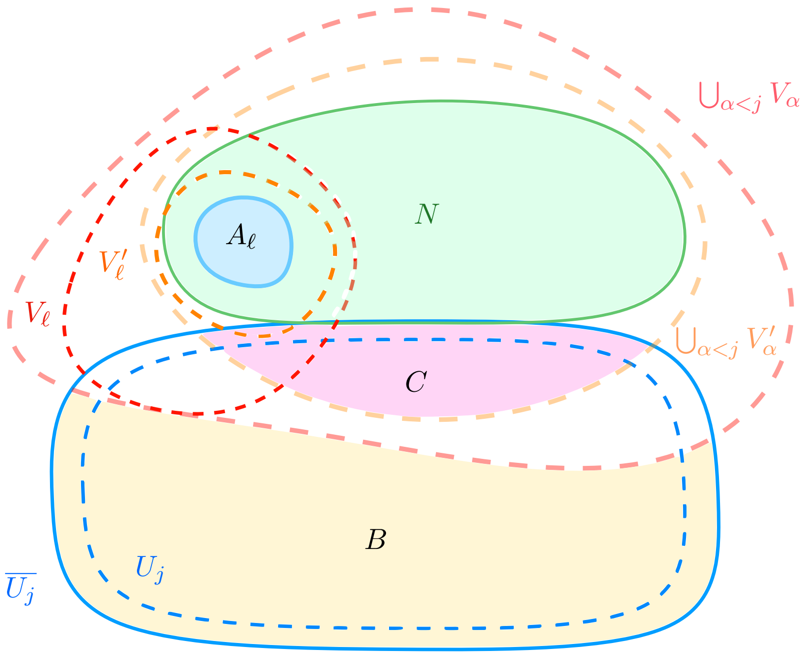

Figure 3.2: A sketch of the construction of the sets B and C, that we need to make disjoint.

Define the region where we will take $h_j = \mathbb{1}$ to be the compact set $B := \overline{U_j} \setminus \bigcup_{\alpha<j} V_\alpha$. We need to make it disjoint from $C := \overline{U_j \cap \bigcup_{\alpha<j} V'_\alpha}$, where we will let $h_j = \rho_j$. Let $N = M \setminus \bigcup_{\alpha \geq j} U_\alpha \subseteq \bigcup_{\alpha<j} U_\alpha$. Note that it is closed and covered by finitely many precompact sets, and therefore it is compact. For $\ell = 1, \dots, j-1$, we will change one V_ℓ at a time. Define

$$A_\ell := N \setminus \left(\bigcup_{\alpha<\ell} V'_\alpha \cup \bigcup_{\ell<\alpha<j} V_\alpha \right).$$

The A_ℓ are compact and

$$A_\ell \subseteq V_\ell \subseteq B^c.$$

Then there exists an open set $V'_\ell \subseteq V_\ell$ such that

$$A_\ell \subseteq V'_\ell \subseteq \overline{V'_\ell} \subseteq B^c.$$

The cover is preserved at each step, since $A_\ell \subseteq V'_\ell$ and the A_ℓ are chosen such that they cover whatever parts of N the other V_α and V'_α do not:

$$M = N \cup \bigcup_{\alpha>j} U_\alpha \text{ and } N \subseteq \bigcup_{\alpha<\ell} V'_\alpha \cup A_\ell \cup \bigcup_{\ell<\alpha<j} V_\alpha.$$

Also, $C \subseteq \bigcup_{\alpha<j} \overline{V'_\alpha} \subseteq B^c$, and therefore $B \cap C = \varnothing$. Note that if $V_\ell \cap U_j = \varnothing$, it follows that $V_\ell \cap \overline{U_j} = \varnothing$ and so V_ℓ makes no difference for the definition of C and does not need to be changed. We can let $V'_\ell := V_\ell$, making the process finite for each $\alpha \in \mathbb{N}$.

We can now say that there exists a cutoff function $\psi_j : \overline{U_j} \to [0,1]$ such that $\psi(B) = 0$ and $\psi(C) = 1$.[12] Then let

$$h_j := \begin{cases} \exp(\psi_j \xi_j) & \text{on } V_j \cap \bigcup_{\alpha<j} V_\alpha, \\ \mathbb{1} & \text{on } V_j \setminus \bigcup_{\alpha<j} V_\alpha. \end{cases}$$

Note that, for $x \in V_j \cap V_\alpha$, $h_j(x) = \exp(a\xi) = \phi_\xi^{t=a}(\mathbb{1})$ for some $a \in [0,1]$ and $\xi \in \mathfrak{g}$, and thus $h_j(x)$ is part of some geodesic between $\mathbb{1}$ and $\exp(\xi) = \exp(\xi_j(x)) = \rho_j(x)$. Then

$$h_j(x) \begin{cases} = h_{j\alpha} h_\alpha g_{\alpha j}(x) & \text{on } V_j \cap V'_\alpha, \text{ for } \alpha \le j-1, \\ \in \gamma(\mathbb{1}, (h_{j\alpha} h_\alpha g_{\alpha j})(x)) & \text{on } V_j \cap V_\alpha, \text{ for } \alpha \le j-1 \\ \equiv \mathbb{1} & \text{on } V_j \setminus \bigcup_{\alpha<j} V_\alpha. \end{cases}$$

Now, we check that h_j satisfies conditions (1)–(3). The first equality shows that condition (1) holds now for $\alpha, \beta \le j$ when replacing V_α by V'_α, as we had already seen with (3.9). Condition (2) remains valid for $\alpha \le j-1$ from the induction hypothesis with $V'_\alpha \subseteq V_\alpha$, and we now check that for $i \ge j+1$ it is valid on $V_j \cap U_i$. On $V_j \cap U_i \setminus \bigcup_{\alpha<j} V_\alpha$, we have $h_j \equiv \mathbb{1}$, and so

$$\mathrm{d}(h_{ij} h_j g_{ji}, \mathbb{1}) = \mathrm{d}(h_{ij} g_{ji}, \mathbb{1}) = \mathrm{d}(h_{ij}, g_{ij}) \le \Delta_{\exp},$$

from the original assumption on the transition functions. On $V_j \cap U_i \cap V_\alpha$ for some $\alpha \le j-1$, we show that h_j lies on the convex geodesic ball $B_{\Delta_{\exp}}(h_{ji} g_{ij})$, and therefore

$$\mathrm{d}(h_{ij} h_j g_{ji}, \mathbb{1}) = \mathrm{d}(h_j, h_{ji} g_{ij}) \le \Delta_{\exp},$$

as we wish. First, note that we have shown that h_j lies on the unique minimal geodesic from $\mathbb{1}$ to $h_{j\alpha} h_\alpha g_{\alpha j}$. Now,

$$\mathrm{d}(\mathbb{1}, h_{ji} g_{ij}) = \mathrm{d}(h_{ij}, g_{ij}) \le \Delta_{\exp}$$

by assumption on the transition functions, and

$$\mathrm{d}(h_{j\alpha} h_\alpha g_{\alpha j}, h_{ji} g_{ij}) = \mathrm{d}(h_\alpha, h_{\alpha j} h_{ji} g_{ij} g_{j\alpha}) = \mathrm{d}(h_\alpha, h_{\alpha i} g_{i\alpha}) = \mathrm{d}(h_{i\alpha} h_\alpha g_{\alpha i}, \mathbb{1}) \le \Delta_{\exp}$$

by the induction hypothesis for $\alpha \le j-1$. Both endpoints of $\gamma(\mathbb{1}, h_{j\alpha} h_\alpha g_{\alpha j})$ lie in $B_{\Delta_{\exp}}(h_{ji} g_{ij})$, and therefore the geodesic is entirely contained in the ball. Finally, condition (3) is still met by h_α for $\alpha \le j-1$ because $h_\alpha \in \mathcal{G}^{k+1,p}(V_\alpha)$ and restricting to a smaller domain preserves the regularity; for $\alpha = j$, we know that ψ_j is smooth and $\xi_j \in W^{k+1,p}(V_j, \mathfrak{g})$, and therefore $h_j = \exp(\psi_j \xi_j) \in \mathcal{G}^{k+1,p}(V_j)$, by lemma 1.21.

We are done with the proof of (i).

For (ii), we have instead of $h_{\alpha\beta}$ a sequence of $h^i_{\alpha\beta}$, each of which is close to $g_{\alpha\beta}$, the transition functions $g_{\alpha\beta}, h^i_{\alpha\beta}$ are said to be in the appropriate Sobolev gauge group $\mathcal{G}^{k+1,p}$ for all $k < K$, and moreover there are uniform bounds on $\left\|(h^i_{\alpha\beta})^{-1}\,\mathrm{d}h^i_{\alpha\beta}\right\|_{k,p\,(U_\alpha \cap U_\beta)}$ for each α, β. We need to check the regularity of the h^i_α for all $k < K$, and that there are uniform bounds on $\|(h^i_\alpha)^{-1}\,\mathrm{d}h^i_\alpha\|_{k,p(V_\alpha)}$ for each α.

[12]Take a partition of unity subordinate to the cover B^c and C^c, for instance.

Since the construction of each h^i_α in (i) is independent of the Sobolev exponent k and regularity followed from the original regularity of the transition functions, we have $h^i_\alpha \in \mathcal{G}^{k+1,p}(V_\alpha)$ for all $k < K$.

For each $k < K$, the uniform bounds will follow by another induction on the cover. The first step is trivial, since the construction in (i) begins with $h^i_1 \equiv \mathbb{1}$ and so $\mathrm{d}h^i_1 = 0$ and the norm is just zero. For the induction hypothesis we will assume that for all $\alpha \leq j-1$ there are uniform bounds on $\|(h^i_\alpha)^{-1}\,\mathrm{d}h^i_\alpha\|_{k,p\,(U_\alpha)}$. Note that $h^i_j = \exp(\psi_j \xi^i_j)$ on U_j and ξ^i_j is defined by $\rho^i_j = \exp(\xi^i_j) = h^i_{j\alpha} h^i_\alpha g_{\alpha j}$ on $U_j \cap V_\alpha$ for $\alpha \leq j-1$. We will show that there are uniform bounds on ρ^i_j on the intersections of U_j with each V_α, which will then give uniform bounds on ξ^i_j on $U_j \cap \bigcup_{\alpha<j} V_\alpha$, and in turn this will give the uniform bounds on h^i_j.

For the uniform bound for ρ^i_j, first note the estimate for $u, v \in \mathcal{G}^{k+1,p}(U)$ over some precompact set U,

$$\begin{aligned}
\left\|(uv)^{-1}\,\mathrm{d}(uv)\right\|_{k,p} &= \left\|v^{-1}u^{-1}(u\,\mathrm{d}v + \mathrm{d}u \cdot v)\right\|_{k,p} \\
&\leq \left\|v^{-1}\,\mathrm{d}v\right\|_{k,p} + \left\|v^{-1}(u^{-1}\,\mathrm{d}u)v\right\|_{k,p} \\
&\overset{(1)}{\leq} \left\|v^{-1}\,\mathrm{d}v\right\|_{k,p} + C\left\|u^{-1}\,\mathrm{d}u\right\|_{k,p}\left(1 + \left\|v^{-1}\,\mathrm{d}v\right\|_{k-1,2p}\right)^k \\
&\overset{(2)}{\leq} \left\|v^{-1}\,\mathrm{d}v\right\|_{k,p} + C\left\|u^{-1}\,\mathrm{d}u\right\|_{k,p}\left(1 + \left\|v^{-1}\,\mathrm{d}v\right\|_{k,p}\right)^k
\end{aligned}$$

where in (1) we used lemma 1.32, and in (2) we used the Sobolev estimate for $W^{k-1,2p} \hookrightarrow W^{k,p}$, which is valid over U even if M is not compact. Since we have uniform bounds over $U_j \cap V_\alpha$ on $\left\|(h^i_{j\alpha})^{-1}\,\mathrm{d}h^i_{j\alpha}\right\|_{k,p}$ by assumption, on $\|(h^i_\alpha)^{-1}\,\mathrm{d}h^i_\alpha\|_{k,p}$ by induction hypothesis and on $\|(g_{\alpha j})^{-1}\,\mathrm{d}g_{\alpha j}\|_{k,p}$ because it is independent of i, we can apply this estimate to $(u,v) = (h^i_{j\alpha}, h^i_\alpha)$ and then to $(u,v) = (h^i_{j\alpha}h^i_\alpha, g_{\alpha j})$ to get a uniform bound on $\left\|(\rho^i_j)^{-1}\,\mathrm{d}\rho^i_j\right\|_{k,p(U_j\cap V_\alpha)}$. Because there are finitely many intersections $U_j \cap V_\alpha$, this gives an overall uniform bound for $\left\|(\rho^i_j)^{-1}\,\mathrm{d}\rho^i_j\right\|_{k,p}$ on its whole domain of definition, $U_j \cap \bigcup_{\alpha<j} V_\alpha$.

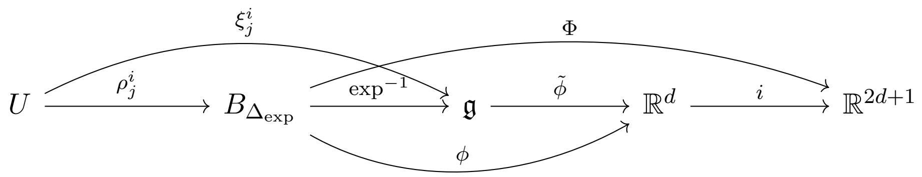

Figure 3.3: Embedding Φ and chart ϕ of $B_{\Delta_{\exp}} \subseteq G$. We have dropped the subscript j and defined $U := U_j \cap \bigcup_{\alpha<j} V_\alpha$.

Finally, we can use this to get a uniform bound for $\left\|\xi^i_j\right\|_{k,p}$ on $U_j \cap \bigcup_{\alpha<j} V_\alpha$ in a rather roundabout way, which we only sketch. First, fix an embedding $\Phi : B_{\Delta_{\exp}}(\mathbb{1}) \subseteq G \to \mathbb{R}^{2d+1}$, which contains the image of every ρ^i_j, and use the bound on $(\rho^i_j)^{-1}\,\mathrm{d}\rho^i_j$ to achieve bounds on the embedding; this will then give bounds on a global chart $\phi : B_{\Delta_{\exp}}(\mathbb{1}) \to \mathbb{R}^d$, and subsequently on ξ^i_j. To see this more clearly, make the following definitions: let

$U := U_j \cap \bigcup_{\alpha<j} V_\alpha$, then let $\tilde{\phi} : \mathfrak{g} \to \mathbb{R}^d$ be an isomorphism, and use it to define a chart $\phi := \tilde{\phi} \circ \exp^{-1}$ and an embedding $\Phi := (\phi, 0) = \iota \circ \phi$ for $\iota : \mathbb{R}^d \hookrightarrow \mathbb{R}^{2d+1}$. To get from bounds on ρ to bounds on the embedding, use these definitions and see the calculations on page 189 of [Weh04]; then the calculations on page 187 give estimates on the chart, which gives estimates on ξ^i_j via the isomorphism, and finally see the calculations on page 188 to achieve estimates on the h^i_j.

□

Appendix A

From the point of view of physics

In this appendix, we seek to motivate the study of gauge theory from another point of view. We tell the history of the concept of gauge invariance, with its origins in the classical theory of electromagnetism, through the revolution of quantum mechanics, and finally emerging as a central concept in contemporary particle physics. For this, we mainly follow the survey [JO01]. At the same time, we recast Maxwell's equations in terms of differential forms, and comment on the relation between physics, and the theory of fibre bundles and Yang–Mills connections that were studied in this work. Even though we do not necessarily follow them here, two interesting mathematical textbooks on differential geometry and gauge theory motivated by physics are [BM94] and [Nab11].

Although it was not called by this name yet, the idea of gauge invariance was noticed as a feature of classical electromagnetism during the time this theory was being formulated, around the second half of the 19th century. After several partial developments, the first clear statement of the arbitrariness of the potentials appeared in a 1909 book by H. A. Lorentz. The history of how the equations of electromagnetism came to be, from empirical observations to subsequent refinements of the formulations, is a fascinating subject that we will not get into; it can be found in any introductory textbook on the subject. Instead, we will skip ahead to the differential formulation of Maxwell's equations, using modern notation and nomenclature. For simplicity, we will work with the vacuum equations (where charge and current densities vanish).[1]

Define two vector fields $E(x, y, z, t)$ and $B(x, y, z, t)$ on $\mathbb{R}^3 \times \mathbb{R}$, called the electric and magnetic fields, respectively. With the ∇ operator acting only on the spatial part, i.e., $\mathbb{R}^3$, Maxwell's equations are written

$$\begin{aligned} \nabla \cdot B &= 0, & \nabla \cdot E &= 0, \\ \nabla \times E &= -\frac{\partial B}{\partial t}, & \nabla \times B &= \frac{1}{c^2}\frac{\partial E}{\partial t}. \end{aligned} \tag{A.1}$$

One important feature of electromagnetism is that we can work with what are called scalar and vector potentials, ϕ and A, which satisfy

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

[1]We will also use the international system of units, with the constants c for the speed of light, $\hbar$ for the reduced Planck's constant and e for the elementary charge.

Using these potentials, the system of equations (A.1) remains unchanged under the transformations

$$\begin{aligned} \phi &\mapsto \phi - \frac{\partial f}{\partial t}, \\ A &\mapsto A + \nabla f, \end{aligned} \tag{A.2}$$

for some real valued function f, since E and B themselves remain unchanged. This can be checked with a simple computation, recalling that $\nabla \times \nabla \equiv 0$. Nowadays, these are called *gauge transformations*, and a choice of A and ϕ is referred to as a *choice of gauge*. Specific gauges can be chosen for practical purposes to simplify calculations in different situations. Some famous choices are named, like the Coulomb gauge in $\mathbb{R}^3$ or the Lorenz gauge in Minkowski space,

$$\nabla \cdot A = 0, \qquad \partial_\mu A^\mu = 0.$$

Using an identification of vector fields with forms, we write E as a 1-form and B as a 2-form on $\mathbb{R}^3$ with time-dependent components, and so we can rewrite Maxwell's equations with differential forms as

$$\begin{aligned} \mathrm{d}B &= 0, & \mathrm{d} * E &= 0, \\ \mathrm{d}E + \frac{\partial B}{\partial t} &= 0, & \mathrm{d} * B &= \frac{1}{c^2}\frac{\partial * E}{\partial t}. \end{aligned}$$

Let us further rewrite the equations, this time as equations on the Minkowski space[2] $\mathbb{R}^{3,1}$. We combine $E \in \Omega^1(\mathbb{R}^3)$ and $B \in \Omega^2(\mathbb{R}^3)$ into a 2-form

$$F = B + E \wedge \mathrm{d}t \in \Omega^2(\mathbb{R}^{3,1})$$

known as the electromagnetic field. Thus we have Maxwell's equations in Minkowski space

$$\mathrm{d}F = 0, \tag{A.3}$$

$$\mathrm{d} * F = 0. \tag{A.4}$$

Note that F is a closed form, and because we are working over $\mathbb{R}^4$ it is also exact, thus we may define a potential $A \in \Omega^1(\mathbb{R}^{3,1})$ such that

$$F = \mathrm{d}A.$$

Now, a transformation

$$A \mapsto A + \mathrm{d}f \tag{A.5}$$

clearly leaves F unchanged; if $A = -\phi\,\mathrm{d}t + A_1\,\mathrm{d}x_1 + A_2\,\mathrm{d}x_2 + A_3\,\mathrm{d}x_3$, for the potentials $A = (A_1, A_2, A_3)$ and ϕ above, then it is clear that (A.5) corresponds to a gauge transformation (A.2). It is also straightforward to show that the Coulomb or Lorentz gauge fixing condition can be written as

$$\mathrm{d}^* A = - * \,\mathrm{d} * A = 0$$

[2] Aside from it being the usual setting for dealing with special relativity, there is nothing special about the Minkowski space for this formulation of Maxwell's equations. We can take spacetime to be any manifold and define an electromagnetic field there as a 2-form F. However, in order to define separate electric and magnetic fields, we need a splitting $\mathbb{R} \times S$ for some manifold S we call space.

on $\mathbb{R}^3$ or $\mathbb{R}^{3,1}$.

If we make a slight modification, and let the fields and functions take pure imaginary values, all that preceded can be reinterpreted in terms of the mathematical gauge theory discussed in this work: A is a connection 1-form on a trivial bundle over $\mathbb{R}^{3,1}$, taking values in the Lie algebra $i\mathbb{R}$; letting the gauge (or structure) group G be abelian, $F = \mathrm{d}A$ is the curvature of A, and Maxwell's equations are the Bianchi identity (A.3) and the Yang–Mills equation (A.4). There are then two possibilities for the Lie group with $\mathfrak{g} = i\mathbb{R}$: either $U(1)$ or $i\mathbb{R}$ itself. How do we know which one it is?

One way of finding out is by looking at the quantum mechanical picture. In quantum mechanics, one works with a Hilbert space of physical states, which are then represented by unitary vectors $\psi \in \mathcal{H}$. Two vectors which differ by a phase, say

$$\psi' = \psi \exp(i\theta),$$

still represent the same state. In 1926, V. Fock had been looking at the relativistic wave equation for spinless particles, now known as the Klein–Gordon equation,

$$\left\{ \left(\partial^\mu + \frac{ie}{\hbar c} A^\mu \right)^2 + \left(\frac{mc}{\hbar} \right)^2 \right\} \psi = 0.$$

For such an equation to maintain its form under a gauge transformation (A.2), Fock proposed to incorporate the following change,

$$\psi \mapsto \psi \exp(i\theta(x)), \tag{A.6}$$

with no physical meaning. As $\exp(i\theta) \in U(1)$, this shows that $U(1)$ is the correct group.

Because of their arbitrariness, it is clear that local potentials are not physical observables. Still, that is not to say they are completely devoid of physical meaning. In the ideal picture of the Aharonov-Bohm effect, an infinite solenoid produces a magnetic field inside but no field outside when the current is on. Does switching the current on or off alter the trajectory of a charged particle passing near the solenoid? It turns out that the answer is yes! The physical explanation is that the wave function of the particle experiences a phase shift when the current is on, which causes interference in the calculation of the path integral. This effect has been shown experimentally: in a double slit experiment, a solenoid placed between the two slits changes the interference pattern that appears. Thus, local fields cannot account for all of the physics we observe, and we must also use potentials.

Mathematically, the solenoid represents a hole in the base manifold (spacetime), which is now no longer simply connected; the connection is flat since the curvature is zero everywhere outside the solenoid, but there is no global choice of gauge that makes the connection matrix $A = 0$ everywhere. Then the Aharonov-Bohm effect is a manifestation of holonomy. In more general gauge theories, physicists quantify holonomy using observables called Wilson loops.

Remark A.1. Not long after Fock's proposal of a better principle of "gradient invariance" (as he called it), it was noticed that this was very similar to an idea of H. Weyl. In 1919, Weyl had been attempting to reconcile electromagnetism with general relativity by looking at scale changes in the metric, which he called called *eichinvarianz*, or scale invariance. When discussing the transformations (A.2) and (A.6), Weyl called it the "principle of gauge invariance" in analogy to his own previous work. This was the first use the word "gauge" in a physics paper in English, and the name stuck.

Thus, we have showed that electromagnetism fits into the theory of Yang–Mills connections that was discussed in chapter 1. However, far from being a mere abstract generalization of Maxwell's equations, Yang–Mills theory underpins contemporary particle physics.

In 1954, the physicists C.-N. Yang and R. Mills were examining symmetries of something called isospin, a property of particles used, for instance, to differentiate between protons and neutrons, and attempting to write equations of motion which would remain invariant under some local transformation. In direct analogy with electromagnetism (where gauge invariance was already well known), in their paper [YM54] they proposed a formula for a covariant derivative adding a term to the usual derivative, and then defined a field strength that transforms in a very simple way under (isospin) gauge transformations. From a mathematician's point of view, Yang and Mills managed to guess the formula for the curvature of a connection, without knowing what a connection was.

The gauge invariant field equations they obtained from this formulation came to be known as Yang–Mills equations, and their idea of writing field equations by using principles of gauge invariance became fundamental. In the years that followed, physicists came up with the Standard Model of particle physics, which classifies all elementary particles and explains three of the four fundamental forces of the universe: the strong force is explained by Quantum Chromodynamics, a Yang–Mills gauge theory with symmetry group $SU(3)$; and the electromagnetic and weak forces were joined in electroweak theory, which is a Yang–Mills gauge theory with symmetry group $U(1) \times SU(2)$. Gravity, of course, is explained by general relativity, which is also a gauge theory but not of Yang–Mills type.

The Standard Model is an incredibly successful theory, in that its predictions have been tested to unprecedented levels of accuracy in experiments such as those at the Large Hadron Collider. However, it is still a very incomplete model, as it leaves many phenomena unexplained and fails to incorporate gravitation. While there seems to be no clear direction for physicists to go beyond the Standard Model, many of its proposed extensions continue bearing fruit for mathematicians.

Appendix B

Background material

We will need some results on calculus of Banach spaces. A good source is [AP95].

Definition B.1. Let X and Y be Banach spaces, and $U \subseteq X$ an open subset, and consider a map $F : U \to Y$. We say that F is Fréchet differentiable at $u \in U$ if there exists a bounded operator $A : X \to Y$ such that

$$\lim_{\|h\| \to 0} \frac{\|F(u+h) - F(u) - A(h)\|}{\|h\|} = 0.$$

Such an A is uniquely determined as is called the (Fréchet) differential of F at u.

The following result is the analogous to the usual mean value theorem of calculus.

Proposition B.2 (Mean value inequality). *Let $F : U \to Y$ be differentiable. For $u, v \in U$ such that the line segment $[u, v]$ is contained in U,*

$$\|F(u) - F(v)\| \leq \sup_{w \in [u,v]} \|\,\mathrm{d}F(w)\| \,.$$

And this is the implicit function theorem for Banach spaces.

Theorem B.3 (Implicit function theorem). *Let $T : X \times Y \to Z$ be a continuous map between Banach spaces that is differentiable with respect to Y, and suppose there is a point $(\alpha, \beta) \in X \times Y$ such that $T(\alpha, \beta) = 0$ and $\partial_Y T|_{(\alpha,\beta)}$ is bijective.*

Then there exist neighbourhoods $U \subseteq X$ and $V \subseteq Y$ of α and β, respectively, and a continuous map $f : U \to V$ such that for all $x \in U$, $T(x, f(x)) = 0$.

Lemma B.4. *If T is an operator on a Banach space such that $\|T - I\| < 1$, then*

$$\left\|T^{-1}\right\| \leq \frac{1}{1 - \|T - I\|}.$$

Lemma B.5 ([Weh04], lemma E.4). *Let $T, S : X \to Z$ be bounded linear operators between Banach spaces, and suppose that T is bijective and $\|T^{-1}\| \, \|S\| < 1$. Then the perturbed operator $T + S$ is also bijective with bounded inverse.*

The Neumann problem

The first four chapters of [Weh04] are fully devoted to giving a good and thorough introduction to the Neumann problem, beginning with the L^2 theory for the homogeneous problem, generalization to L^p-spaces and sections of vector bundles, and the inhomogeneous problem. We quote here only a few results that will be needed in chapter 2. For M a compact manifold with boundary, the Neumann problem is as follows:

$$\begin{cases} \Delta u = f & \text{on } M, \\ \frac{\partial u}{\partial \nu} = g & \text{on } \partial M, \end{cases} \tag{B.1}$$

where ν is the exterior normal direction. The problem is said homogeneous when $g = 0$. From now on let $1 < p < \infty$ and $k \in \mathbb{N}_0$. If $f \in W^{k,p}(M)$, then the natural space for the boundary values g is

$$W_\partial^{k+1,p}(M) := \frac{W^{1,p}(M)}{W_\delta^{k+1,p}(M)},$$

where $W_\delta^{k+1,p}(M)$ is defined as the closure in $W^{k+1,p}(M)$ of the smooth functions vanishing on the boundary. The norm on this space is

$$\|g\|_{W_\partial^{k+1,p}} = \inf \left\{ \|G\|_{k+1,p} : G \in W^{k+1,p}(M) \text{ and } g = G|_{\partial M} \right\}.$$

Theorem B.6 ([Weh04], theorem 3.1)**.** *Let $f \in L^p(M)$ and $g \in W_\partial^{1,p}(M)$. Then there exists a solution $u \in W^{2,p}(M)$ the Neumann problem if and only if*

$$\int_M f + \int_{\partial M} g = 0.$$

The solution is unique up to an additive constant.

Proposition B.7 ([Weh04], theorem 3.2)**.** *There exist constants C, C' such that for all $u \in W^{k+2,p}(M)$,*

$$\|u\|_{k+2,p} \le C' \left(\|\Delta u\|_{k,p} + \left\| \frac{\partial u}{\partial \nu} \right\|_{W_\partial^{k+1,p}} + \|u\|_{k+1,p} \right),$$
$$\|u\|_{k+2,p} \le C \left(\|\Delta u\|_{k,p} + \left\| \frac{\partial u}{\partial \nu} \right\|_{W_\partial^{k+1,p}} \right), \qquad \text{if } \int_M u = 0.$$

Moreover, for each $k \in \mathbb{N}_0$, these constants depend continuously on the metric on M. For C the dependence is with respect to the $W^{k+1,\infty}$-topology, and for C' the dependence is with respect to the $W^{k+2,\infty}$-topology on the space of metrics.

When restricted to functions such that $\int_M u = 0$, the second estimate implies that the operator associated to the Neumann problem has bounded inverse.

Orphaned lemmas and estimates

Finally, we leave in this section a couple of estimates on Sobolev spaces which did not add to the presentation of the theory in chapter 1, but are nonetheless useful.

Lemma B.8 ([Weh04], lemma B.3). *Let M be a compact Riemannian n-manifold, and let $k \in \mathbb{N}_0$ and $1 \leq p, r, s < \infty$ be such that either*

$$r, s \geq p \quad \text{and} \quad \frac{1}{r} + \frac{1}{s} < \frac{k}{n} + \frac{1}{p},$$

or

$$r, s > p \quad \text{and} \quad \frac{1}{r} + \frac{1}{s} \leq \frac{k}{n} + \frac{1}{p}.$$

Then there is a constant C such that for all $\alpha \in W^{k,r}(M)$ and $\beta \in W^{k,s}(M)$ the product lies in $W^{k,p}(M)$ and satisfies

$$\|\alpha \cdot \beta\|_{k,p} \leq C \|\alpha\|_{k,r} \|\beta\|_{k,s}.$$

A particular important case is $k \leq 1$, $r = s = p$ and $kp > n$.

Lemma B.9 ([Weh04], lemma 6.5). *Let M be a compact Riemannian n-manifold and $1 \leq q \leq p < \infty$ such that $q \geq \frac{n}{2}$. In case $q < n$, assume further that $p \leq \frac{nq}{n-q}$. Then there exists a constant C_{rnpq} such that, for all $f, g \in W^{1,p}(M)$*

$$\|f \cdot g\|_q \leq C_{rnpq} \|f\|_r \|g\|_{1,q},$$
$$\|f \cdot g\|_p \leq C_{rnpq} \|f\|_r \|g\|_{1,p},$$
$$\|f \cdot g\|_p \leq C_{rnpq} \|f\|_{1,q} \|g\|_{1,p},$$

where we have:

(i) For $q < n$, $r = \frac{nq}{n-q} \geq p$,

(ii) for $q = n$, $r = 2p$,

(iii) and for $q > n$, $r = \infty$.

Whenever this lemma is used the constant will be called C_{rnpq}. The following lemma is sometimes called Kato's inequality, and we prove it in a particular case.

Lemma B.10. *Let $E \to B$ be a vector bundle over $B \subseteq \mathbb{R}^n$ with the Euclidean metric $g_{ij} = \delta_{ij}$, and choose a metric and a compatible connection for E. Then for a section A of $T^*M \otimes E$,*

$$|\nabla |A|| \leq |\nabla A|.$$

Proof. Write $A = A_i \, \mathrm{d}x_i$, then

$$(\nabla A)_{ij} = \nabla_i A(\partial_j) - A(\nabla_i \partial_j) = \nabla_i A_j,$$

because $\nabla_i \partial_j = \Gamma_{ij}^k \partial_k$ and the Christoffel symbols are zero for the Euclidean metric. Then simply write

$$\nabla |A| = \mathrm{d} |A| = \frac{1}{2|A|} \mathrm{d}\langle A, A\rangle = \frac{1}{2|A|} \sum \partial_i \langle A, A\rangle \, \mathrm{d}x_i = \frac{1}{|A|} \sum \langle \nabla_i A, A\rangle \, \mathrm{d}x_i,$$

and use the Cauchy-Schwartz inequality in the following:

$$|\nabla |A||^2 = \sum \frac{\langle \nabla_i A, A\rangle^2}{|A|^2} \leq \sum \frac{|\nabla_i A|^2 |A|^2}{|A|^2} = \sum |\nabla_i A|^2 = |\nabla A|^2.$$

□

Bibliography

[ADHM78] M. F. Atiyah, V. G. Drinfel'd, N. J. Hitchin, and Yu. I. Manin. Construction of instantons. *Phys. Lett. A*, 65(3):185–187, 1978. 2

[AHS78] M. F. Atiyah, N. J. Hitchin, and I. M. Singer. Self-duality in four-dimensional Riemannian geometry. *Proc. Roy. Soc. London Ser. A*, 362(1711):425–461, 1978. 2

[AP95] A. Ambrosetti and G. Prodi. *A primer of nonlinear analysis*, volume 34 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. 76

[Ati79] M. F. Atiyah. *Geometry of Yang-Mills fields*. Lezione Fermiani. Scuola Normale Superiore, 1979. 2

[Aub82] T. Aubin. *Nonlinear analysis on manifolds. Monge-Ampère equations*, volume 252 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. 11, 25

[AW77] M. F. Atiyah and R. S. Ward. Instantons and algebraic geometry. *Comm. Math. Phys.*, 55, 1977. 2

[BM94] J. Baez and J. P. Muniain. *Gauge Fields, Knots and Gravity*. Series on Knots and Everything. World Scientific, 1994. 2, 72

[BPST75] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Y. S. Tyupkin. Pseudoparticle solutions of the yang-mills equations. *Physics Letters B*, 59(1):85–87, 1975. 2

[Cra15] M. Crainic. Lecture notes for Mastermath course Differential Geometry 2015/2016. `https://webspace.science.uu.nl/~crain101/DG-2015/main10.pdf`, 2015. [Online; accessed 26-May-2021]. 6, 10, 13, 60

[DK97] S. K. Donaldson and P. B. Kronheimer. *The Geometry of Four-Manifolds*. Oxford Mathematical Monographs. Clarendon Press, 1997. 30, 42, 43, 60

[Don83] S. K. Donaldson. An application of gauge theory to four-dimensional topology. *J. Differential Geom.*, 18(2), 1983. 3

[Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010. 25

[Flo88] A. Floer. An instanton-invariant for 3-manifolds. *Comm. Math. Phys.*, 118(2), 1988. 3

[FU91] D. S. Freed and K. K. Uhlenbeck. *Instantons and Four-Manifolds*. MSRI Publications. Springer-Verlag, 2 edition, 1991. 44

[GHL90] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, second edition, 1990. 24

[Hö7] L. Hörmander. *The analysis of linear partial differential operators. III*. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition. 31

[JO01] J. D. Jackson and L. B. Okun. Historical roots of gauge invariance. *Rev. Mod. Phys.*, 73:663–680, Sep 2001. 72

[Jos02] J. Jost. *Riemannian geometry and geometric analysis*. Universitext. Springer-Verlag, Berlin, third edition, 2002. 43

[Kna96] A. W. Knapp. *Lie Groups Beyond an Introduction*. Birkhäuser, 1996. 23

[KW07] A. Kapustin and E. Witten. Electric-Magnetic Duality And The Geometric Langlands Program. *Commun. Num. Theor. Phys.*, 1:1–236, 2007. 3

[Nab11] G. L. Naber. *Topology, geometry, and gauge fields*, volume 25 of *Texts in Applied Mathematics*. Springer, New York, second edition, 2011. Foundations. 72

[Nic07] L. I. Nicolaescu. *Lectures on the geometry of manifolds*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition, 2007. 25, 31

[Sch95] G. Schwarz. *Hodge decomposition—a method for solving boundary value problems*, volume 1607 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1995. 31

[Ste51] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton Mathematical Series, vol. 14. Princeton University Press, Princeton, N. J., 1951. 2

[Tau07] C. H. Taubes. The Seiberg-Witten equations and the Weinstein conjecture. *Geom. Topol.*, 11, 2007. 3

[Uhl82] K. K. Uhlenbeck. Connections with L^p bounds on curvature. *Comm. Math. Phys.*, 83(1):31–42, 1982. 3, 30, 41, 42, 45, 59, 61, 64

[War83] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. 43

[Weh04] K. Wehrheim. *Uhlenbeck Compactness*. Series of Lectures in Mathematics. EMS, 2004. 3, 23, 25, 27, 29, 30, 36, 41, 42, 45, 65, 71, 76, 77, 78

[Wel80] R. O. Wells. *Differential analysis on complex manifolds*. Graduate Texts in Mathematics. Springer-Verlag, 1980. 31

[WY75] T. T. Wu and C. N. Yang. Concept of Nonintegrable Phase Factors and Global Formulation of Gauge Fields. *Phys. Rev. D*, 12:3845–3857, 1975. 2

[YM54] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 1954. 2, 75

Index