Gauge Theory

José Figueroa-O'Farrill*

http://www.maths.ed.ac.uk/empg/Activities/GT

Version of 6th October 2019

These are the notes accompanying the first few lectures on **Gauge Theory**, a PG course taught in Edinburgh in the Spring of 2006.
The only requirement is a working familiarity with differential geometry and Lie groups; although scholia on the necessary definitions will be scattered throughout the notes.
Any statement which is not proved to your satisfaction is to be thought of as an exercise, even if not explicitly labelled as such!
These notes are still in a state of flux and I am happy to receive comments and suggestions either by email or in person.

* ✉ j.m.figueroa(at)ed.ac.uk

Contents

Lecture 1: Connections on principal fibre bundles

The beauty and profundity of the geometry of fibre bundles were to a large extent brought forth by the (early) work of Chern. I must admit, however, that the appreciation of this beauty came to physicists only in recent years.

— CN Yang, 1979

In this lecture we introduce the notion of a principal fibre bundle and of a connection on it, but we start with some motivation.

1.1 Motivation: the Dirac monopole

It is only fitting to start this course, which takes place at the JCMB, with a solution of Maxwell's equations. The magnetic field $\mathbf{B}$ of a magnetic monopole sitting at the origin in $\mathbb{R}^3$ is given by

$$\mathbf{B}(\boldsymbol{x}) = \frac{\boldsymbol{x}}{4\pi r^3} \,,$$

where $r = |\boldsymbol{x}|$. This satisfies $\operatorname{div}\mathbf{B} = 0$ and hence is a solution of Maxwell's equations in $\mathbb{R}^3 \setminus \{\mathbf{0}\}$.

> Maxwell's equations are linear and hence this solution requires a source of magnetic field, namely the monopole which sits at the origin. We will see later in this course that there are other monopoles which do not require sources and which extend smoothly to the 'origin'.

In modern language, the vector field $\mathbf{B}$ is understood as the 2-form $\mathrm{F} \in \Omega^2(\mathbb{R}^3 \setminus \{\mathbf{0}\})$ given by

$$\mathrm{F} = \frac{1}{2\pi r^3}\left(x_1\, dx_2 \wedge dx_3 + \text{cyclic}\right)$$

in the cartesian coordinates of $\mathbb{R}^3$. Maxwell's equations say that $d\mathrm{F} = 0$. This is perhaps more evident in spherical coordinates ($x_1 = r\sin\theta\cos\phi, \dots$), where

$$\mathrm{F} = \frac{1}{4\pi}\sin\theta\, d\theta \wedge d\phi \,.$$

Since $d\mathrm{F} = 0$ we may hope to find a one-form A such that $\mathrm{F} = d\mathrm{A}$. For example,

$$\mathrm{A} = -\frac{1}{4\pi}\cos\theta\, d\phi \,.$$

This is not regular over all of $\mathbb{R}^3 \setminus \{\mathbf{0}\}$, however. This should not come as a surprise, since after all spherical coordinates are singular on the x_3-axis. Rewriting A in cartesian coordinates

$$\mathrm{A} = \frac{x_3}{4\pi r}\,\frac{x_2\, dx_1 - x_1\, dx_2}{x_1^2 + x_2^2} \,,$$

clearly displays the singularity at $x_1 = x_2 = 0$.

> In the old Physics literature this singularity is known as the "Dirac string," a language we shall distance ourselves from in this course.

This singularity will afflict any A we come up with. Indeed, notice that F restricts to the (normalised) area form on the unit sphere $\mathrm{S} \subset \mathbb{R}^3$, whence

$$\int_{\mathrm{S}} \mathrm{F} = 1 \,.$$

If $\mathrm{F} = d\mathrm{A}$ for a smooth one-form A, then Stokes's theorem would have forced $\int_{\mathrm{S}} \mathrm{F} = 0$.

The principal aim of the first couple of lectures is to develop the geometric framework to which F (and A) belong: the theory of connections on principal fibre bundles, to which we now turn.

1.2 Principal fibre bundles

A **principal fibre bundle** consists of the following data:

- a manifold P, called the **total space**;
- a Lie group G acting freely on P on the right:

$$\begin{aligned} \mathrm{P}\times\mathrm{G} &\to \mathrm{P} \\ (p,g) &\mapsto pg \quad \text{(or sometimes } \mathrm{R}_g p) \end{aligned}$$

  where by a free action we mean that the stabiliser of every point is trivial, or paraphrasing, that every element G (except the identity) moves every point in P. We will also assume that the space of orbits $\mathrm{M} = \mathrm{P}/\mathrm{G}$ is a manifold (called the **base**) and the natural map $\pi : \mathrm{P} \to \mathrm{M}$ taking a point to its orbit is a smooth surjection. For every $m \in \mathrm{M}$, the submanifold $\pi^{-1}(m) \subset \mathrm{P}$ is called the **fibre over** m.

Further, this data will be subject to the condition of **local triviality**: that M admits an open cover $\{\mathrm{U}_\alpha\}$ and G-equivariant diffeomorphisms $\psi_\alpha : \pi^{-1}\mathrm{U}_\alpha \to \mathrm{U}_\alpha \times \mathrm{G}$ such that the following diagram commutes

$$\begin{array}{ccc} \pi^{-1}\mathrm{U}_\alpha & \xrightarrow{\ \psi_\alpha\ } & \mathrm{U}_\alpha \times \mathrm{G} \\ & {\scriptstyle \pi}\searrow \quad \swarrow{\scriptstyle \mathrm{pr}_1} & \\ & \mathrm{U}_\alpha & \end{array}$$

This means that $\psi_\alpha(p) = (\pi(p), g_\alpha(p))$, for some G-equivariant map $g_\alpha : \pi^{-1}\mathrm{U}_\alpha \to \mathrm{G}$ which is a fibrewise diffeomorphism. Equivariance means that

$$g_\alpha(pg) = g_\alpha(p)g\ .$$

One often abbreviates the above data by saying that

$$\begin{array}{c} \mathrm{P} \\ \downarrow{\scriptstyle \pi} \\ \mathrm{M} \end{array} \quad \text{or} \quad \begin{array}{ccc} \mathrm{G} & \longrightarrow & \mathrm{P} \\ & & \downarrow{\scriptstyle \pi} \\ & & \mathrm{M} \end{array} \quad \text{is a principal G-bundle,}$$

but be aware that abuses of language are rife in this topic. Don't be surprised by statements such as "Let P be a principal bundle..."

We say that the bundle is **trivial** if there exists a diffeomorphism $\Psi : \mathrm{P} \to \mathrm{M} \times \mathrm{G}$ such that $\Psi(p) = (\pi(p), \chi(p))$ and such that $\chi(pg) = \chi(p)g$. This last condition is simply the G-equivariance of Ψ.

A **section** is a (smooth) map $s : \mathrm{M} \to \mathrm{P}$ such that $\pi \circ s = \mathrm{id}$. In other words, it is a smooth assignment to each point in the base of a point in the fibre over it. Sections are rare. Indeed, one has

Done? ❑ **Exercise 1.1.** Show that a principal fibre bundle admits a section if and only if it is trivial. (This is in sharp contrast with, say, vector bundles, which always admit sections.)

Solution. If $s : \mathrm{M} \to \mathrm{P}$ is a section, define $\Psi : \mathrm{P} \to \mathrm{M} \times \mathrm{G}$ by sending $\Psi(p) = (\pi(p), \chi(p))$, where $\chi(p) \in \mathrm{G}$ is uniquely specified by $p = s(\pi(p))\chi(p)$. Conversely, given $\Psi : \mathrm{P} \to \mathrm{M} \times \mathrm{G}$, define a section by $s(m) = \Psi^{-1}(m, e)$. ◆

Nevertheless, since P is locally trivial, local sections do exist. In fact, there are local sections $s_\alpha : \mathrm{U}_\alpha \to \pi^{-1}\mathrm{U}_\alpha$ canonically associated to the trivialisation, defined so that for every $m \in \mathrm{U}_\alpha$, $\psi_\alpha(s_\alpha(m)) = (m, e)$, where $e \in \mathrm{G}$ is the identity element. In other words, $g_\alpha \circ s_\alpha : \mathrm{U}_\alpha \to \mathrm{G}$ is the constant function sending every point to the identity. Conversely, a local section s_α allows us to identify the fibre over m with G. Indeed, given any $p \in \pi^{-1}(m)$, there is a unique group element $g_\alpha(p) \in \mathrm{G}$ such that $p = s_\alpha(m)g_\alpha(p)$.

On nonempty overlaps $U_{\alpha\beta} := U_\alpha \cap U_\beta$, we have two ways of trivialising the bundle:

$$\begin{array}{ccccc} U_{\alpha\beta}\times G & \xleftarrow{\psi_\beta} & \pi^{-1}U_{\alpha\beta} & \xrightarrow{\psi_\alpha} & U_{\alpha\beta}\times G \\ & \searrow^{\mathrm{pr}_1} & \downarrow \pi & \swarrow^{\mathrm{pr}_1} & \\ & & U_{\alpha\beta} & & \end{array}$$

For $m \in U_{\alpha\beta}$ and $p \in \pi^{-1}(m)$, we have $\psi_\alpha(p) = (m, g_\alpha(p))$ and $\psi_\beta(p) = (m, g_\beta(p))$, whence there is $\bar{g}_{\alpha\beta}(p) \in G$ such that $g_\alpha(p) = \bar{g}_{\alpha\beta}(p) g_\beta(p)$. In other words,

$$\bar{g}_{\alpha\beta}(p) = g_\alpha(p) g_\beta(p)^{-1} . \tag{1}$$

In fact, $\bar{g}_{\alpha\beta}(p)$ is constant on each fibre:

$$\begin{aligned} \bar{g}_{\alpha\beta}(pg) &= g_\alpha(pg) g_\beta(pg)^{-1} \\ &= g_\alpha(p) g g^{-1} g_\beta(p) \qquad \text{(by equivariance of } g_\alpha, g_\beta) \\ &= \bar{g}_{\alpha\beta}(p) . \end{aligned}$$

In other words, $\bar{g}_{\alpha\beta}(p) = g_{\alpha\beta}(\pi(p))$ for some function

$$g_{\alpha\beta} : U_{\alpha\beta} \to G .$$

From equation (1), it follows that these **transition functions** obey the following **cocycle conditions**:

$$g_{\alpha\beta}(m) g_{\beta\alpha}(m) = e \qquad \text{for every } m \in U_{\alpha\beta}, \text{ and} \tag{2}$$

$$g_{\alpha\beta}(m) g_{\beta\gamma}(m) g_{\gamma\alpha}(m) = e \qquad \text{for every } m \in U_{\alpha\beta\gamma}, \tag{3}$$

where $U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma$.

Done? ❑ **Exercise 1.2.** Show that on double overlaps, the canonical sections s_α are related by

$$s_\beta(m) = s_\alpha(m) g_{\alpha\beta}(m) \qquad \text{for every } m \in U_{\alpha\beta}.$$

Solution. Simply observe that $g_\alpha \circ s_\beta = g_\alpha \circ g_\beta^{-1} \circ g_\beta \circ s_\beta = g_{\alpha\beta}$. ◆

One can reconstruct the bundle from an open cover $\{U_\alpha\}$ and transition functions $\{g_{\alpha\beta}\}$ obeying the cocycle conditions (2) as follows:

$$P = \bigsqcup_\alpha (U_\alpha \times G) \Big/ \sim ,$$

where $(m, g) \sim (m, g_{\alpha\beta}(m) g)$ for all $m \in U_{\alpha\beta}$ and $g \in G$. Notice that π is induced by the projection onto the first factor and the action of G on P is induced by right multiplication on G, both of which are preserved by the equivalence relation, which uses *left* multiplication by the transition functions. (Associativity of group multiplication guarantees that right and left multiplications commute.)

Example 1.1 (Möbius band)**.** The boundary of the Möbius band is an example of a nontrivial principal $\mathbb{Z}_2$-bundle. This can be described as follows. Let $S^1 \subset \mathbb{C}$ denote the complex numbers of unit modulus and let $\pi : S^1 \to S^1$ be the map defined by $z \mapsto z^2$. Then the fibre $\pi^{-1}(z^2) = \{\pm z\}$ consists of two points. A global section would correspond to choosing a square-root function smoothly on the unit circle. This does not exist, however, since any definition of $z^{1/2}$ always has a branch cut from the origin out to the point at infinity. Therefore the bundle is not trivial. In fact, if the bundle were trivial, the total space would be disconnected, being two disjoint copies of the circle. However building a paper model of the Möbius band one quickly sees that its boundary is connected.

We can understand this bundle in terms of the local data as follows. Cover the circle by two overlapping open sets: U_1 and U_2. Their intersection is the disjoint union of two intervals in the circle: $V_1 \sqcup V_2$. Let $g_i : V_i \to \mathbb{Z}_2$ denote the transition functions, which are actually constant since V_i are connected and $\mathbb{Z}_2$ is discrete, so we can think of $g_i \in \mathbb{Z}_2$. There are no triple overlaps, so the cocycle condition is vacuously satisfied. It is an easy exercise to check that the resulting bundle is trivial if and only if $g_1 = g_2$ and nontrivial otherwise.

1.3 Connections

The push-forward and the pull-back

Let $f : M \to N$ be a smooth map between manifolds. The **push-forward**

$$f_* : TM \to TN \qquad \text{(also written } Tf \text{ when in a categorical mood)}$$

is the collection of fibre-wise linear maps $f_* : T_m M \to T_{f(m)} N$ defined as follows. Let $v \in T_m M$ be represented as the velocity of a curve $t \mapsto \gamma(t)$ through m; that is, $\gamma(0) = m$ and $\gamma'(0) = v$. Then $f_*(v) \in T_{f(m)} N$ is the velocity at $f(m)$ of the curve $t \mapsto f(\gamma(t))$; that is, $f_* \gamma'(0) = (f \circ \gamma)'(0)$. If $g : N \to Q$ is another smooth map between manifolds, then so is their composition $g \circ f : M \to Q$. The **chain rule** is then simply the "functoriality of the push-forward": $(g \circ f)_* = g_* \circ f_*$. Dual to the push-forward, there is the **pull-back** $f^* : T^*_{f(m)} N \to T^*_m M$, defined pointwise and only on the image of f. If α is a one-form on N, then $(f^*\alpha)(v) = \alpha(f_* v)$ is a one-form on M. The pull-back is also functorial, but now reversing the order $(g \circ f)^* = f^* \circ g^*$.

Let $\pi : P \to M$ be a principal G-bundle and let $m \in M$ and $p \in \pi^{-1}(m)$. The **vertical** subspace $V_p \subset T_p P$ consists of those vectors tangent to the fibre at p; in other words, $V_p = \ker \pi_* : T_p P \to T_m M$. A vector field $v \in \mathscr{X}(P)$ is **vertical** if $v(p) \in V_p$ for all p. The Lie bracket of two vertical vector fields is again vertical. The vertical subspaces define a G-invariant distribution $V \subset TP$: indeed, since $\pi \circ R_g = \pi$, we have that $(R_g)_* V_p = V_{pg}$. In the absence of any extra structure, there is no natural complement to V_p in $T_p P$. This is in a sense what a connection provides.

1.3.1 Connections as horizontal distributions

A **connection** on P is a smooth choice of **horizontal** subspaces $H_p \subset T_p P$ complementary to V_p:

$$T_p P = V_p \oplus H_p$$

and such that $(R_g)_* H_p = H_{pg}$. In other words, a connection is a G-invariant distribution $H \subset TP$ complementary to V.

For example, a G-invariant riemannian metric on P gives rise to a connection, simply by defining $H_p = V_p^\perp$. This simple observation underlies the Kałuża–Klein programme relating gravity on P to gauge theory on M. It also underlies many geometric constructions, since it is often the case that 'nice' metrics will give rise to 'nice' connections and vice versa.

We will give two more characterisations of connections on P, but first, a little revision.

Some Lie group technology

A Lie group is a manifold with two smooth operations: a multiplication $G \times G \to G$, and an inverse $G \to G$ obeying the group axioms. If $g \in G$ we define diffeomorphisms

$$\begin{aligned} L_g : G &\to G \\ x &\mapsto gx \end{aligned} \qquad \text{and} \qquad \begin{aligned} R_g : G &\to G \\ x &\mapsto xg \end{aligned}$$

called **left** and **right** multiplications by g, respectively.
A vector field $v \in \mathscr{X}(G)$ is **left-invariant** if $(L_g)_* v = v$ for all $g \in G$. In other words, $v(g) = (L_g)_* v(e)$ for all $g \in G$, where e is the identity. The Lie bracket of two left-invariant vector fields is left-invariant. The vector space of left-invariant vector fields defines the **Lie algebra** $\mathfrak{g}$ of G. A left-invariant vector field is uniquely determined by its value at the identity, whence $\mathfrak{g} \cong T_e G$.
The **(left-invariant) Maurer–Cartan form** is the $\mathfrak{g}$-valued 1-form θ on G defined by

$$\theta_g = (L_{g^{-1}})_* : T_g G \to T_e G = \mathfrak{g} .$$

If v is a left-invariant vector field, then $\theta(v) = v(e)$, whence θ_e is the natural identification between $T_e G$ and $\mathfrak{g}$. For a matrix group, $\theta_g = g^{-1} dg$, from where it follows that θ is left-invariant and satisfies the **structure equation**:

$$d\theta = -\tfrac{1}{2}[\theta, \theta] ,$$

where the bracket in the RHS denotes both the Lie bracket in $\mathfrak{g}$ and the wedge product of 1-forms.
Every $g \in G$ defines a smooth map $\mathrm{Ad}_g : G \to G$ by $\mathrm{Ad}_g = L_g \circ R_g^{-1}$; that is,

$$\mathrm{Ad}_g h = g h g^{-1} .$$

This map preserves the identity, whence its derivative there defines a linear representation of the group on the Lie algebra known as the **adjoint representation** $\mathrm{ad}_g := (\mathrm{A}d_g)_* : \mathfrak{g} \to \mathfrak{g}$, defined explicitly by

$$\mathrm{ad}_g X = \frac{d}{dt} \left(g e^{tX} g^{-1} \right) \Big|_{t=0} .$$

For G a matrix group, $\mathrm{ad}_g(X) = gXg^{-1}$. Finally, notice that $R_g^* \theta = \mathrm{ad}_{g^{-1}} \circ \theta$.

The action of G on P defines a map $\sigma : \mathfrak{g} \to \mathscr{X}(P)$ assigning to every $X \in \mathfrak{g}$, the **fundamental vector field** $\sigma(X)$ whose value at p is given by

$$\sigma_p(X) = \frac{d}{dt} \left(p e^{tX} \right) \Big|_{t=0} .$$

Notice that

$$\pi_* \sigma_p(X) = \frac{d}{dt} \pi \left(p e^{tX} \right) \Big|_{t=0} = \frac{d}{dt} \pi(p) \Big|_{t=0} = 0 ,$$

whence $\sigma(X)$ is a vertical vector field. In fact, since G acts freely, the map $X \mapsto \sigma_p(X)$ is an isomorphism $\sigma_p : \mathfrak{g} \xrightarrow{\cong} V_p$ for every p.

Lemma 1.1.

$$(R_g)_* \sigma(X) = \sigma\left(\mathrm{ad}_{g^{-1}} X \right) .$$

Proof. By definition, at $p \in P$, we have

$$\begin{aligned}
(R_g)_* \sigma_p(X) &= \frac{d}{dt} R_g \left(pe^{tX}\right)\Big|_{t=0} \\
&= \frac{d}{dt}\left(pe^{tX}g\right)\Big|_{t=0} \\
&= \frac{d}{dt}\left(pgg^{-1}e^{tX}g\right)\Big|_{t=0} \\
&= \frac{d}{dt}\left(pge^{t\,\mathrm{ad}_{g^{-1}} X}\right)\Big|_{t=0} \\
&= \sigma_{pg}\left(\mathrm{ad}_{g^{-1}} X\right) .
\end{aligned}$$

□

Done? ❑ **Exercise 1.3.** Let $g_\alpha : \pi^{-1}U_\alpha \to G$ be the maps defined by the local trivialisation. Show that $(g_\alpha)_* \sigma_p(X) = (L_{g_\alpha(p)})_* X$.

Solution. By definition,

$$\begin{aligned}
(g_\alpha)_* \sigma_p(X) &= \frac{d}{dt} g_\alpha(pe^{tX})\Big|_{t=0} \\
\text{(by equivariance of } g_\alpha) \quad &= \frac{d}{dt}\left(g_\alpha(p)e^{tX}\right)\Big|_{t=0} \\
&= \frac{d}{dt}\left(L_{g_\alpha(p)} e^{tX}\right)\Big|_{t=0} \\
&= (L_{g_\alpha(p)})_* X .
\end{aligned}$$

◆

1.3.2 The connection one-form

The horizontal subspace $H_p \subset T_pP$, being a linear subspace, is cut out by $k = \dim G$ linear equations $T_pP \to \mathbb{R}$. In other words, H_p is the kernel of k one-forms at p, the components of a one-form ω at p with values in a k-dimensional vector space. There is a natural such vector space, namely the Lie algebra $\mathfrak{g}$ of G, and since ω annihilates horizontal vectors it is defined by what it does to the vertical vectors, and we do have a natural map $V_p \to \mathfrak{g}$ given by the inverse of σ_p. This prompts the following definition.

The **connection one-form** of a connection $H \subset TP$ is the $\mathfrak{g}$-valued one-form $\omega \in \Omega^1(P;\mathfrak{g})$ defined by

$$\omega(v) = \begin{cases} X & \text{if } v = \sigma(X) \\ 0 & \text{if } v \text{ is horizontal.} \end{cases}$$

Proposition 1.2. *The connection one-form obeys*

$$R_g^* \omega = \mathrm{ad}_{g^{-1}} \circ \omega .$$

Proof. Let $v \in H_p$, so that $\omega(v) = 0$. By the G-invariance of H, $(R_g)_* v \in H_{pg}$, whence $R_g^*\omega$ also annihilates v and the identity is trivially satisfied. Now let $v = \sigma_p(X)$ for some $X \in \mathfrak{g}$. Then, using Lemma 1.1,

$$R_g^*\omega(\sigma(X)) = \omega\left((R_g)_*\sigma(X)\right) = \omega\left(\sigma\left(\mathrm{ad}_{g^{-1}} X\right)\right) = \mathrm{ad}_{g^{-1}} X .$$

□

Conversely, given a one-form $\omega \in \Omega^1(P;\mathfrak{g})$ satisfying the identity in Proposition 1.2 and such that $\omega(\sigma(X)) = X$, the distribution $H = \ker\omega$ defines a connection on P.

We say that a form on P is **horizontal** if it annihilates the vertical vectors. Notice that if ω and ω' are connection one-forms for two connections H and H′ on P, their difference $\omega - \omega' \in \Omega^1(P;\mathfrak{g})$ is horizontal. We will see later that this means that it defines a section through a bundle on M associated to P.

1.3.3 Gauge fields

Finally, as advertised, we make contact with the more familiar notion of gauge fields as used in Physics, which live on M instead of P.

Recall that we have local sections $s_\alpha : U_\alpha \to \pi^{-1}U_\alpha$ associated canonically to the trivialisation of the bundle, along which we can pull-back the connection one-form ω, defining in the process the following $\mathfrak{g}$-valued one-forms on U_α:

$$A_\alpha := s_\alpha^* \omega \in \Omega^1(U_\alpha; \mathfrak{g}) \ .$$

Proposition 1.3. *The restriction of the connection one-form ω to $\pi^{-1}U_\alpha$ agrees with*

$$\omega_\alpha = \mathrm{ad}_{g_\alpha^{-1}} \circ \pi^* A_\alpha + g_\alpha^* \theta \ ,$$

where θ is the Maurer–Cartan one-form.

Proof. We will prove this result in two steps.

1. First we show that ω_α and ω agree on the image of s_α. Indeed, let $m \in U_\alpha$ and $p = s_\alpha(m)$. We have a direct sum decomposition
$$T_p P = \mathrm{im}(s_\alpha \circ \pi)_* \oplus V_p \ ,$$
so that every $v \in T_p P$ can be written uniquely as $v = (s_\alpha)_* \pi_*(v) + \bar{v}$, for a unique vertical vector $\bar{v}$. Applying ω_α on v, we obtain (since $g_\alpha(s_\alpha(m)) = e$)
$$\begin{aligned} \omega_\alpha(v) &= \left(\pi^* s_\alpha^* \omega\right)(v) + \left(g_\alpha^* \theta_e\right)(v) \\ &= \omega\left((s_\alpha)_* \pi_* v\right) + \theta_e\left((g_\alpha)_* v\right) \\ \text{(since } (g_\alpha \circ s_\alpha)_* = 0\text{)} \quad &= \omega\left((s_\alpha)_* \pi_* v\right) + \theta_e\left((g_\alpha)_* \bar{v}\right) \\ &= \omega\left((s_\alpha)_* \pi_* v\right) + \omega(\bar{v}) \\ &= \omega(v) \ . \end{aligned}$$

2. Next we show that they transform in the same way under the right action of G:
$$\begin{aligned} R_g^* (\omega_\alpha)_{pg} &= \mathrm{ad}_{g_\alpha(pg)^{-1}} \circ R_g^* \pi^* s_\alpha^* \omega + R_g^* g_\alpha^* \theta \\ \text{(equivariance of } g_\alpha\text{)} \quad &= \mathrm{ad}_{(g_\alpha(p)g)^{-1}} \circ R_g^* \pi^* s_\alpha^* \omega + g_\alpha^* R_g^* \theta \\ \text{(since } \pi \circ R_g = \pi\text{)} \quad &= \mathrm{ad}_{g^{-1} g_\alpha(p)^{-1}} \circ \pi^* s_\alpha^* \omega + g_\alpha^* (\mathrm{ad}_{g^{-1}} \circ \theta) \\ &= \mathrm{ad}_{g^{-1}} \circ \left(\mathrm{ad}_{g_\alpha(p)^{-1}} \circ \pi^* s_\alpha^* \omega + g_\alpha^* \theta\right) \\ &= \mathrm{ad}_{g^{-1}} \circ (\omega_\alpha)_p \ . \end{aligned}$$

Therefore they agree everywhere on $\pi^{-1}U_\alpha$. □

Now since ω is defined globally, we have that $\omega_\alpha = \omega_\beta$ on $\pi^{-1}U_{\alpha\beta}$. This allows us to relate A_α and A_β on $U_{\alpha\beta}$. Indeed, on $U_{\alpha\beta}$,

$$\begin{aligned} A_\alpha = s_\alpha^* \omega_\alpha &= s_\alpha^* \omega_\beta \\ &= s_\alpha^* \left(\mathrm{ad}_{g_\beta(s_\alpha)^{-1}} \circ \pi^* A_\beta + g_\beta^* \theta\right) \\ \text{(using } g_\beta \circ s_\alpha = g_{\beta\alpha}\text{)} \quad &= \mathrm{ad}_{g_{\alpha\beta}} \circ A_\beta + g_{\beta\alpha}^* \theta \ . \end{aligned}$$

In summary,

$$A_\alpha = \mathrm{ad}_{g_{\alpha\beta}} \circ A_\beta + g_{\beta\alpha}^* \theta \ , \tag{4}$$

or, equivalently,

$$A_\alpha = \mathrm{ad}_{g_{\alpha\beta}} \circ \left(A_\beta - g_{\alpha\beta}^* \theta\right) \ ,$$

where we have used the result of the following

Done? ❑ **Exercise 1.4.** Show that
$$\mathrm{ad}_{g_{\alpha\beta}} \circ g_{\alpha\beta}^{*}\theta = -g_{\beta\alpha}^{*}\theta\ .$$

Solution. This is easiest to do for matrix groups. Let us simplify notation so that $g := g_{\alpha\beta}$ and hence $g_{\beta\alpha} = g^{-1}$. In this notation, the exercise asks us to show that

$$\mathrm{ad}_g \circ g^{*}\theta = -(g^{-1})^{*}\theta.$$

For matrix groups, $\mathrm{ad}_g X = gXg^{-1}$ and $g^{*}\theta = g^{-1}dg$, so that

$$\mathrm{ad}_g \circ g^{*}\theta = g\left(g^{-1}dg\right)g^{-1} = dgg^{-1}. \tag{5}$$

But now notice that $0 = d(gg^{-1}) = dgg^{-1} + gdg^{-1}$, so that

$$dgg^{-1} = -gdg^{-1} = -(g^{-1})^{*}\theta,$$

as desired. ◆

For matrix groups this becomes the more familiar

$$\boxed{A_{\alpha} = g_{\alpha\beta}A_{\beta}g_{\alpha\beta}^{-1} - dg_{\alpha\beta}g_{\alpha\beta}^{-1}\ .} \tag{6}$$

Conversely, given a family of one-forms $A_{\alpha} \in \Omega^{1}(U_{\alpha};\mathfrak{g})$ satisfying equation (4) on overlaps $U_{\alpha\beta}$, we can construct a globally defined $\omega \in \Omega^{1}(P;\mathfrak{g})$ by the formula in Proposition 1.3. Then ω is the connection one-form of a connection on P.

In summary, we have three equivalent descriptions of a connection on P:

1. a G-invariant horizontal distribution $H \subset TP$,
2. a one-form $\omega \in \Omega^{1}(P;\mathfrak{g})$ satisfying $\omega(\sigma(X)) = X$ and the identity in Proposition 1.2, and
3. a family of one-forms $A_{\alpha} \in \Omega^{1}(U_{\alpha};\mathfrak{g})$ satisfying equation (4) on overlaps.

Each description has its virtue and we're lucky to have all three!

1.4 The space of connections

Connections exist! This is a fact which we are not going to prove in this course. The proof can be found in [KN63, § II.2]. What we will prove is that the space of connections is an (infinite-dimensional) affine space. In fact, we have already seen this. Indeed, we saw that if ω and ω' are the connection one-forms of two connections H and H′, their difference $\tau = \omega - \omega'$ is a horizontal $\mathfrak{g}$-valued one-form on P satisfying the equivariance condition $R_g^{*}\tau = \mathrm{ad}_{g^{-1}} \circ \tau$. Let us see what this means on M. Let $\tau_{\alpha} \in \Omega^{1}(U_{\alpha};\mathfrak{g})$ be the pull-back of τ by the local sections:

$$\tau_{\alpha} = s_{\alpha}^{*}\tau = s_{\alpha}^{*}(\omega - \omega') = A_{\alpha} - A_{\alpha}'\ .$$

Then equation (4) on $U_{\alpha\beta}$, says that

$$\tau_{\alpha} = g_{\alpha\beta}\tau_{\beta}g_{\alpha\beta}^{-1} = \mathrm{ad}_{g_{\alpha\beta}} \circ \tau_{\beta}\ . \tag{7}$$

We claim that the $\{\tau_{\alpha}\}$ define a section of a vector bundle associated to P.

Associated fibre bundles

Let G act on a space F via automorphisms and let $\varrho : G \to \mathrm{Aut}(F)$ be the corresponding representation. For example, F could be a vector space and $\mathrm{Aut}(F) = \mathrm{GL}(F)$, or F could be a manifold and $\mathrm{Aut}(F) = \mathrm{Diff}(F)$.
The data defining the principal fibre bundle $P \to M$ allows to define a fibre bundle over M as follows. Consider the quotient

$$P \times_G F := (P \times F)\big/G$$

by the G-action $(p, f)g = (pg, \varrho(g^{-1})f)$. Since G acts freely on P, it acts freely on $P \times F$ and since P/G is a smooth manifold, so is $P \times_G F$. Moreover the projection $\pi : P \to M$ induces a projection $\pi_F : P \times_G F \to M$, by $\pi_F(p, f) = \pi(p)$, which is well-defined because $\pi(pg) = \pi(p)$.
The data $\pi_F : P \times_G F \to M$ defines a fibre bundle **associated to** P **via** ϱ. For example, taking ϱ to be the adjoint representations $\mathrm{ad} : G \to \mathrm{GL}(\mathfrak{g})$ and $\mathrm{Ad} : G \to \mathrm{Diff}(G)$ in turn, we arrive at the associated vector bundle $\mathrm{ad}\,P := P \times_G \mathfrak{g}$ and the associated fibre bundle $\mathrm{Ad}\,P := P \times_G G$.
The associated bundle $P \times_G F$ can also be constructed locally from the local data defining P, namely the open cover $\{U_\alpha\}$ and the transition functions $\{g_{\alpha\beta}\}$ on double overlaps. Indeed, we have that

$$P \times_G F = \bigsqcup_\alpha (U_\alpha \times F)\Big/ \sim \,,$$

where $(m, f) \sim (m, \varrho(g_{\alpha\beta}(m))f)$ for all $m \in U_{\alpha\beta}$ and $f \in F$.
Sections of $P \times_G F$ are represented by functions $f : P \to F$ with the equivariance condition:

$$R_g^* f = \varrho(g^{-1}) \circ f \,,$$

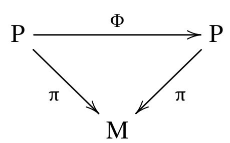
or, equivalently, by a family of functions $f_\alpha : U_\alpha \to F$ such that

$$f_\alpha(m) = \varrho(g_{\alpha\beta}(m)) f_\beta(m) \qquad \text{for all } m \in U_{\alpha\beta}.$$

We therefore interpret equation (6) as saying that the family of one-forms $\tau_\alpha \in \Omega^1(U_\alpha; \mathfrak{g})$ defines a one-form with values in the adjoint bundle ad P. The space $\Omega^1(M; \mathrm{ad}\,P)$ of such one-forms is an (infinite-dimensional) vector space, whence the space $\mathscr{A}$ of connections on P is an infinite-dimensional affine space modelled on $\Omega^1(M; \mathrm{ad}\,P)$. It follows that $\mathscr{A}$ is contractible. In particular, the tangent space $T_A\mathscr{A}$ to $\mathscr{A}$ at a connection A is naturally identified with $\Omega^1(M; \mathrm{ad}\,P)$.

1.5 Gauge transformations

Every geometrical object has a natural notion of automorphism and principal fibre bundles are no exception. A **gauge transformation** of a principal fibre bundle $\pi : P \to M$ is a G-equivariant diffeomorphism $\Phi : P \to P$ making the following diagram commute



In particular, Φ maps fibres to themselves and equivariance means that $\Phi(pg) = \Phi(p)g$. Composition makes gauge transformations into a group, which we will denote $\mathscr{G}$.

We can describe $\mathscr{G}$ in terms of a trivialisation. Since it maps fibres to themselves, a gauge transformation Φ restricts to a gauge transformation of the trivial bundle $\pi^{-1}U_\alpha$ over U_α. Applying the trivialisation map $\psi_\alpha(\Phi(p)) = (\pi(p), g_\alpha(\Phi(p)))$, which lets us define $\bar\phi_\alpha : \pi^{-1}U_\alpha \to G$ by

$$\bar\phi_\alpha(p) = g_\alpha(\Phi(p)) g_\alpha(p)^{-1} \,.$$

Equivariance of g_α and of Φ means that

$$\bar{\phi}_\alpha(pg) = \bar{\phi}_\alpha(p) ,$$

whence $\bar{\phi}_\alpha(p) = \phi_\alpha(\pi(p))$ for some function

$$\phi_\alpha : U_\alpha \to G .$$

For $m \in U_{\alpha\beta}$, and letting $p \in \pi^{-1}(m)$, we have

$$\begin{aligned}
\phi_\alpha(m) &= g_\alpha(\Phi(p)) g_\alpha(p)^{-1} \\
&= g_\alpha(\Phi(p)) g_\beta(\Phi(p))^{-1} g_\beta(\Phi(p)) g_\beta(p)^{-1} g_\beta(p) g_\alpha(p)^{-1} \\
\text{(since } \pi(\Phi(p)) = m\text{)} \quad &= g_{\alpha\beta}(m) \phi_\beta(m) g_{\alpha\beta}(m)^{-1} \\
&= \mathrm{Ad}_{g_{\alpha\beta}(m)} \phi_\beta(m) ,
\end{aligned}$$

whence the $\{\phi_\alpha\}$ define a section of the associated fibre bundle $\mathrm{Ad}\,P$. Since the $\{\phi_\alpha\}$ determine Φ uniquely (and vice versa), we see that $\mathscr{G} = C^\infty(M; \mathrm{Ad}\,P)$.

1.6 The action of $\mathscr{G}$ on $\mathscr{A}$

The group $\mathscr{G}$ of gauge transformations acts naturally on the space $\mathscr{A}$ of connections. We can see this in several different ways.

Let $H \subset TP$ be a connection and let $\Phi : P \to P$ be a gauge transformation. Define $H^\Phi := \Phi_* H$. This is also a connection on P. Indeed, the equivariance of Φ makes $H^\Phi \subset TP$ into a G-invariant distribution:

$$\begin{aligned}
(R_g)_* H^\Phi_{\Phi(p)} &= (R_g)_* \Phi_* H_p \\
\text{(equivariance of } \Phi\text{)} \quad &= \Phi_* (R_g)_* H_p \\
\text{(invariance of H)} \quad &= \Phi_* H_{pg} \\
\text{(definition of } H^\Phi\text{)} \quad &= H^\Phi_{\Phi(pg)} \\
\text{(equivariance of } \Phi\text{)} \quad &= H^\Phi_{\Phi(p)g} .
\end{aligned}$$

Moreover, H^Φ is still complementary to V because Φ_* is an isomorphism which preserves the vertical subspace.

Done? ❑ **Exercise 1.5.** Show that the fundamental vector fields $\sigma(X)$ of the G-action are **gauge invariant**; that is, $\Phi_* \sigma(X) = \sigma(X)$ for every $\Phi \in \mathscr{G}$. Deduce that if ω is the connection one-form for a connection H then $\omega^\Phi := (\Phi^*)^{-1}\omega$ is the connection one-form for the gauge-transformed connection H^Φ.

Solution. To prove the gauge-invariance of the fundamental vector fields, we simply calculate:

$$\begin{aligned}
\Phi_* \sigma_p(X) &= \frac{d}{dt} \Phi(pe^{tX}) \Big|_{t=0} \\
\text{(equivariance of } \Phi\text{)} \quad &= \frac{d}{dt} (\Phi(p) e^{tX}) \Big|_{t=0} \\
&= \sigma_{\Phi(p)}(X) .
\end{aligned}$$

To show that ω^Φ is the connection one-form for H^Φ we must show that $\omega^\Phi(\sigma(X)) = X$ for every $X \in \mathfrak{g}$ and that $\ker \omega^\Phi = H^\Phi$. The first follows from the gauge-invariance of the fundamental vector fields:

$$\omega^\Phi(\sigma(X)) = \left((\Phi^{-1})^* \omega\right)(\sigma(X)) = \omega\left(\Phi_*^{-1} \sigma(X)\right) = \omega(\sigma(X)) = X ,$$

whereas for the second,

$$\ker \omega^\Phi = \Phi_* \ker \omega = \Phi_* H = H^\Phi .$$

◆

Finally we work out the effect of gauge transformations on a gauge field. Let $m \in U_\alpha$ and $p \in \pi^{-1}(m)$. Let A_α and A_α^Φ be the gauge fields on U_α corresponding to the connections H and H^Φ. By Proposition 1.3, the connection one-forms ω and ω^Φ are given at p by

$$\begin{aligned} \omega_p &= \operatorname{ad}_{g_\alpha(p)^{-1}} \circ \pi^* A_\alpha + g_\alpha^* \theta \\ \omega_p^\Phi &= \operatorname{ad}_{g_\alpha(p)^{-1}} \circ \pi^* A_\alpha^\Phi + g_\alpha^* \theta \,. \end{aligned} \tag{8}$$

On the other hand, $\omega^\Phi = (\Phi^{-1})^* \omega$, from where we can obtain a relation between A_α and A_α^Φ. Indeed, letting $q = \Phi^{-1}(p)$, we have

$$\begin{aligned} \omega_p^\Phi = (\Phi^{-1})^* \omega_q &= \operatorname{ad}_{g_\alpha(q)^{-1}} \circ (\Phi^{-1})^* \pi^* A_\alpha + (\Phi^{-1})^* g_\alpha^* \theta \\ \text{(functoriality of pull-back)} \quad &= \operatorname{ad}_{g_\alpha(q)^{-1}} \circ (\pi \circ \Phi^{-1})^* A_\alpha + (g_\alpha \circ \Phi^{-1})^* \theta \\ \text{(since } \pi \circ \Phi^{-1} = \pi) \quad &= \operatorname{ad}_{g_\alpha(q)^{-1}} \circ \pi^* A_\alpha + (g_\alpha \circ \Phi^{-1})^* \theta \\ \text{(since } g_\alpha(p) = \bar\phi_\alpha(p) g_\alpha(q)) \quad &= \operatorname{ad}_{g_\alpha(p)^{-1} \bar\phi_\alpha(p)} \circ \pi^* A_\alpha + (g_\alpha \circ \Phi^{-1})^* \theta \,. \end{aligned}$$

Now, $(g_\alpha \circ \Phi^{-1})(p) = g_\alpha(q) = \bar\phi_\alpha(p)^{-1} g_\alpha(p)$, whence

$$(g_\alpha \circ \Phi^{-1})^* \theta = g_\alpha^* \theta - \operatorname{ad}_{g_\alpha(p)^{-1} \bar\phi_\alpha(p)} \bar\phi_\alpha^* \theta \,.$$

This identity is easier to prove for matrix groups, since

$$(g_\alpha \circ \Phi^{-1})^* \theta = g_\alpha(p)^{-1} \bar\phi(p) d(\bar\phi(p)^{-1} g_\alpha(p)) \,.$$

Now we put everything together using that $\bar\phi_\alpha = \phi_\alpha \circ \pi$ to arrive at

$$\omega_p^\Phi = \operatorname{ad}_{g_\alpha(p)^{-1} \phi_\alpha(m)} \circ \pi^* \left(A_\alpha - \phi_\alpha^* \theta\right) + g_\alpha^* \theta \,,$$

whence comparing with the second equation in (7), we conclude that

$$A_\alpha^\Phi = \operatorname{ad}_{\phi_\alpha} \circ \left(A_\alpha - \phi_\alpha^* \theta\right) \,,$$

or for matrix groups,

$$\boxed{A_\alpha^\Phi = \phi_\alpha A_\alpha \phi_\alpha^{-1} - d\phi_\alpha \, \phi_\alpha^{-1} \,.} \tag{9}$$

Comparing with equation (4), we see that in overlaps gauge fields change by a local gauge transformation defined on the overlap. This means that any gauge-invariant object which is constructed out of the gauge fields will be well-defined globally on M.

Lecture 2: Curvature

In this lecture we will define the curvature of a connection on a principal fibre bundle and interpret it geometrically in several different ways. Along the way we define the covariant derivative of sections of associated vector bundles. Throughout this lecture, $\pi : P \to M$ will denote a principal G-bundle.

2.1 The curvature of a connection

2.1.1 The horizontal projection

Given a connection $H \subset TP$, we define the **horizontal projection** $h : TP \to TP$ to be the projection onto the horizontal distribution along the vertical distribution. It is a collection of linear maps $h_p : T_pP \to T_pP$, for every $p \in P$, defined by

$$h_p(v) = \begin{cases} v & \text{if } v \in H_p, \text{ and} \\ 0 & \text{if } v \in V_p. \end{cases}$$

In other words, $\operatorname{im} h = H$ and $\ker h = V$. Since both H and V are invariant under the the action of G, the horizontal projection is equivariant:

$$h \circ (R_g)_* = (R_g)_* \circ h\,.$$

We will let $h_p^* : T_p^*P \to T_p^*P$ denote the collection of dual maps, whence if, say, $\alpha \in \Omega^1(P)$ is a one-form, $h^*\alpha = \alpha \circ h$. More generally if $\beta \in \Omega^k(P)$, then $(h^*\beta)(v_1, \ldots, v_k) = \beta(hv_1, \ldots, hv_k)$. However...

> Despite the notation, h^* is *not* the pull-back by a smooth map! In particular, h^* will *not* commute with the exterior derivative d!

2.1.2 The curvature 2-form

Let $\omega \in \Omega^1(P;\mathfrak{g})$ be the connection one-form for a connection $H \subset TP$. The 2-form $\Omega := h^* d\omega \in \Omega^2(P;\mathfrak{g})$ is called the **curvature (2-form)** of the connection. We will derive more explicit formulae for Ω later on, but first let us interpret the curvature geometrically.

By definition,

$$\begin{aligned} \Omega(u,v) &= d\omega(hu, hv) \\ &= (hu)\omega(hv) - (hv)\omega(hu) - \omega([hu,hv]) \\ \text{(since } h^*\omega = 0) \qquad &= -\omega([hu,hv])\,; \end{aligned}$$

whence $\Omega(u,v) = 0$ if and only if $[hu,hv]$ is horizontal. In other words, the curvature of the connection measures the failure of integrability of the horizontal distribution $H \subset TP$.

> **Frobenius integrability**
>
> A distribution $D \subset TP$ is said to be integrable if the Lie bracket of any two sections of D lies again in D. The theorem of Frobenius states that a distribution is integrable if every $p \in P$ lies in a unique submanifold of P whose tangent space at p agrees with the subspace $D_p \subset T_pP$. These submanifolds are said to *foliate* P. As we have just seen, a connection $H \subset TP$ is integrable if and only if its curvature 2-form vanishes.
> In contrast, the vertical distribution $V \subset TP$ is always integrable, since the Lie bracket of two vertical vector fields is again vertical, and Frobenius's theorem guarantees that P is foliated by submanifolds whose tangent spaces are the vertical subspaces. These submanifolds are of course the fibres of $\pi : P \to M$.

The integrability of a distribution has a dual formulation in terms of differential forms. A horizontal distribution $H = \ker\omega$ is integrable if and only if (the components of) ω generate a differential ideal, so that $d\omega = \Theta \wedge \omega$, for some $\Theta \in \Omega^1(P;\mathrm{End}(\mathfrak{g}))$. Since Ω measures the failure of integrability of H, the following formula should not come as a surprise.

Proposition 2.1 (Structure equation)**.**

$$\boxed{\Omega = d\omega + \tfrac{1}{2}[\omega,\omega]\ ,}$$

where, as before, $[-,-]$ *is the* symmetric *bilinear product consisting of the Lie bracket on* $\mathfrak{g}$ *and the wedge product of one-forms.*

Proof. We need to show that

$$d\omega(hu,hv) = d\omega(u,v) + [\omega(u),\omega(v)] \tag{10}$$

for all vector fields $u,v \in \mathscr{X}(P)$. We can treat this case by case.

- Let u,v be horizontal. In this case there is nothing to show, since $\omega(u) = \omega(v) = 0$ and $hu = u$ and $hv = v$.
- Let u,v be vertical. Without loss of generality we can take $u = \sigma(X)$ and $v = \sigma(Y)$, for some $X,Y \in \mathfrak{g}$. Then equation (9) becomes

$$\begin{aligned} 0 &\overset{?}{=} d\omega(\sigma(X),\sigma(Y)) + [\omega(\sigma(X)),\omega(\sigma(Y))] \\ &= \sigma(X)Y - \sigma(Y)X - \omega([\sigma(X),\sigma(Y)]) + [X,Y] && (\omega(\sigma(X)) = X, \text{ etc}) \\ &= -\omega([\sigma(X),\sigma(Y)]) + [X,Y] \\ &= -\omega(\sigma([X,Y])) + [X,Y]\ , && ([\sigma(X),\sigma(Y)] = \sigma([X,Y])) \end{aligned}$$

  which is clearly true.
- Finally, let u be horizontal and $v = \sigma(X)$ be vertical, whence equation (9) becomes

$$d\omega(u,\sigma(X)) = 0\ ,$$

  which in turn reduces to

$$\omega([u,\sigma(X)]) = 0\ .$$

  In other words, we have to show that the Lie bracket of a vertical and a horizontal vector field is again horizontal. But this is simply the infinitesimal version of the G-invariance of H.

□

An immediate consequence of this formula is the

Proposition 2.2 (Bianchi identity)**.**

$$h^* d\Omega = 0\ .$$

Proof. This is simply a calculation using the structure equation:

$$\begin{aligned} h^* d\Omega &= h^* d\left(d\omega + \tfrac{1}{2}[\omega,\omega]\right) \\ &= h^*\left(\tfrac{1}{2}[d\omega,\omega] - \tfrac{1}{2}[\omega,d\omega]\right) \\ &= h^*[d\omega,\omega] \\ &= [h^* d\omega, h^*\omega] \\ &= 0\ . \end{aligned}$$

□

Under a gauge transformation $\Phi : P \to P$, the connection one-form changes by $\omega \mapsto \omega^\Phi = (\Phi^{-1})^*\omega$. The curvature also transforms in this way.

Done? ❑ **Exercise 2.1.** Show that under a gauge transformation $\Phi : P \to P$, the horizontal projections h, h^Φ of H and H^Φ are related by

$$h^\Phi = \Phi_* h \Phi_*^{-1} .$$

Deduce that the curvature 2-form transforms as

$$\Omega \mapsto \Omega^\Phi = (\Phi^{-1})^*\Omega .$$

(This can also be shown directly from the structure equation.)

Solution. Let $u \in T_pP$. Relative to the connection H we can decompose it as $u = u_V + hu$ with $u_V \in V_p$. Similarly, relative to the gauge-transformed connection H^Φ, $\Phi_* u = (\Phi_* u)_V + h^\Phi(\Phi_* u)$. Since V is gauge-invariant, $(\Phi_* u)_V = \Phi_* u_V$, whereas since $\Phi_* H = H^\Phi$, we have that $\Phi_* hu = h^\Phi(\Phi_* u)$. In other words, $\Phi_* h = h^\Phi \Phi_*$.

To relate the curvature 2-forms, we simply calculate:

$$\begin{aligned}\Omega^\Phi &= (h^\Phi)^* d\omega^\Phi\\ &= (\Phi^*)^{-1} h^* \Phi^* d\left((\Phi^*)^{-1}\omega\right)\\ &= (\Phi^*)^{-1} h^* d\omega \qquad \text{(since } d \text{ commutes with pull-backs)}\\ &= (\Phi^*)^{-1}\Omega .\end{aligned}$$

◆

2.1.3 Gauge field-strengths

Pulling back Ω via the canonical sections $s_\alpha : U_\alpha \to P$ yields the **gauge field-strength** $F_\alpha := s_\alpha^*\Omega \in \Omega^2(U_\alpha; \mathfrak{g})$. It follows from the structure equation that

$$F_\alpha = dA_\alpha + \tfrac{1}{2}[A_\alpha, A_\alpha] . \tag{11}$$

As usual, the natural question to ask is how do F_α and F_β differ on $U_{\alpha\beta}$. From equation (4), using the Maurer–Cartan structure equation $d\theta = -\frac{1}{2}[\theta, \theta]$ and simplifying, we find

$$F_\alpha = \mathrm{ad}_{g_{\alpha\beta}} \circ F_\beta \tag{12}$$

or, for matrix groups,

$$F_\alpha = g_{\alpha\beta} F_\beta g_{\alpha\beta}^{-1} .$$

In other words, the $\{F_\alpha\}$ define a global 2-form $F \in \Omega^2(M; \mathrm{ad}\,P)$ with values in ad P. We may sometimes write F_A if we want to make the dependence on the gauge fields manifest.

Done? ❑ **Exercise 2.2.** Show that the gauge-transformed field-strength is given by

$$F_\alpha^\Phi = \mathrm{ad}_{\phi_\alpha} \circ F_\alpha .$$

Solution. This is a simple calculation using equations (10) and (8). ◆

2.2 The covariant derivative

A connection allows us to define a "covariant" derivative on sections of associated vector bundles to $P \to M$, but first we need to understand better the relation between forms on P and forms on M.

2.2.1 Basic forms

A k-form $\alpha \in \Omega^k(P)$ is **horizontal** if $h^*\alpha = \alpha$. A horizontal form which in addition is G-invariant is called **basic**. It is a basic fact (no pun intended) that α is basic if and only if $\alpha = \pi^*\bar{\alpha}$ for some k-form $\bar{\alpha}$ on M (hence the name). This story extends to forms on P taking values in a vector space V admitting a representation $\varrho : G \to GL(V)$ of G. Let α be such a form. Then α is **horizontal** if $h^*\alpha = \alpha$ and it is **invariant** if for all $g \in G$,

$$R_g^*\alpha = \varrho(g^{-1}) \circ \alpha \,.$$

If α is both horizontal and invariant, it is said to be **basic**. Basic forms are in one-to-one correspondence with forms on M with values in the associated bundle $P \times_G V$. Indeed, let

$$\Omega_G^k(P;V) = \left\{ \bar{\zeta} \in \Omega^k(P;V) \middle| h^*\bar{\zeta} = \bar{\zeta} \quad \text{and} \quad R_g^*\bar{\zeta} = \varrho(g^{-1}) \circ \bar{\zeta} \right\} \tag{13}$$

denote the basic forms on P with values in V. The k-forms on M with values in the associated bundle $P \times_G V$ are best described relative to a trivialisation of P as a family $\zeta_\alpha \in \Omega^k(U_\alpha;V)$ subject to the gluing condition

$$\zeta_\alpha = \varrho(g_{\alpha\beta}) \circ \zeta_\beta \tag{14}$$

on nonempty overlaps $U_{\alpha\beta}$. Let $\Omega^k(M; P \times_G V)$ denote the space of such bundle-valued forms. We will now construct isomorphisms

$$\Omega_G^k(P;V) \rightleftarrows \Omega^k(M; P \times_G V)$$

as follows in terms of local data.

Let $\bar{\zeta} \in \Omega_G^k(P;V)$ and define $\zeta_\alpha = s_\alpha^*\bar{\zeta} \in \Omega^k(U_\alpha;V)$.

Done? ❑ **Exercise 2.3.** Show that the $\{\zeta_\alpha\}$ define a form in $\Omega^k(M; P \times_G V)$, by showing that equation (13) is satisfied on nonempty overlaps.

Solution. Let $m \in U_{\alpha\beta}$, then $s_\beta(m) = s_\alpha(m) g_{\alpha\beta}(m)$, as shown in Exercise 1.2. In other words, $s_\beta = R_{g_{\alpha\beta}} \circ s_\alpha$, whence

$$\begin{aligned}
\zeta_\beta &= s_\beta^*\bar{\zeta} \\
&= (R_{g_{\alpha\beta}} \circ s_\alpha)^*\bar{\zeta} \\
&= s_\alpha^* R_{g_{\alpha\beta}}^* \bar{\zeta} \\
&= s_\alpha^* \left(\varrho(g_{\alpha\beta}^{-1}) \circ \bar{\zeta} \right) \\
&= \varrho(g_{\alpha\beta}^{-1}) \circ \zeta_\alpha \,,
\end{aligned}$$

which is equivalent to equation (13). ◆

Conversely, if $\zeta_\alpha \in \Omega^k(U_\alpha;V)$ define a form in $\Omega^k(M; P \times_G V)$, then define

$$\bar{\zeta}_\alpha := \varrho(g_\alpha^{-1}) \circ \pi^*\zeta_\alpha \in \Omega^k(\pi^{-1}U_\alpha;V) \,.$$

Done? ❑ **Exercise 2.4.** Show that $\bar{\zeta}_\alpha$ is the restriction to $\pi^{-1}U_\alpha$ of a basic form $\bar{\zeta} \in \Omega_G^k(P;V)$.

Solution. We must show that $\bar{\zeta}_\alpha = \bar{\zeta}_\beta$ on $\pi^{-1}U_{\alpha\beta}$ and that the resulting form $\bar{\zeta} \in \Omega^k(P;V)$ is both horizontal and invariant. Let $m \in U_{\alpha\beta}$ and let $p \in \pi^{-1}(m)$. Then

$$\begin{aligned}
\bar{\zeta}_\alpha(p) &= \varrho(g_\alpha(p)^{-1}) \circ \pi^*\zeta_\alpha(m) \\
&= \varrho(g_\alpha(p)^{-1}) \circ \pi^* \left(\varrho(g_{\alpha\beta}(m)) \circ \zeta_\beta(m) \right) \\
&= \varrho(g_\alpha(p)^{-1} g_{\alpha\beta}(m)) \circ \pi^*\zeta_\beta(m) \\
&= \varrho(g_\beta(p)^{-1}) \circ \pi^*\zeta_\beta(m) \\
&= \bar{\zeta}_\beta(p) \,,
\end{aligned}$$

whence it glues to a global form. We now show that $\bar\zeta$ is horizontal. Let $u = hu + u_V \in \mathscr{X}(P)$ be a vector field. Then restricted to $\pi^{-1}U_\alpha$,

$$\begin{aligned}\bar\zeta_\alpha(u) &= \bar\zeta_\alpha(hu + u_V)\\ &= \varrho(g_\alpha)^{-1}\circ\zeta_\alpha(\pi_* hu) \qquad \text{(since } \pi_* u_V = 0\text{)}\\ &= \bar\zeta_\alpha(hu)\\ &= (h^*\bar\zeta_\alpha)(u)\,.\end{aligned}$$

Finally, we show that it is invariant:

$$\begin{aligned}R_g^*\bar\zeta(pg) &= R_g^*\varrho(g_\alpha(pg)^{-1})\circ\pi^*\zeta_\alpha(m)\\ &= \varrho(g_\alpha(p)g)^{-1}\circ R_g^*\pi^*\zeta_\alpha(m) \qquad \text{(using } g_\alpha(pg) = g_\alpha(p)g\text{)}\\ &= \varrho(g^{-1})\circ\varrho(g_\alpha(p)^{-1})\circ\pi^*\zeta_\alpha(m)\\ &= \varrho(g^{-1})\circ\bar\zeta(p)\,.\end{aligned}$$

In summary, $\bar\zeta \in \Omega^k_G(P;V)$. ◆

Finally we observe that these two constructions are mutual inverses, hence they define the desired isomorphism. This isomorphism is very useful: it allows us to work with bundle-valued forms on M either locally in terms of a trivilisation or globally on P subject to an equivariance condition.

2.2.2 The covariant derivative

The exterior derivative $d : \Omega^k(P;V) \to \Omega^{k+1}(P;V)$ obeys $d^2 = 0$ and defines a complex: the V-**valued de Rham complex**. The invariant forms do form a subcomplex, but the basic forms do not, since $d\alpha$ need not be horizontal even if α is. Projecting onto the horizontal forms defines the **exterior covariant derivative**

$$d^H : \Omega^k_G(P;V) \to \Omega^{k+1}_G(P;V) \qquad \text{by} \qquad d^H\alpha = h^* d\alpha\,.$$

The price we pay is that $(d^H)^2 \neq 0$ in general, so we no longer have a complex. Indeed, the failure of d^H defining a complex is again measured by the curvature of the connection.

Let us start by deriving a more explicit formula for the exterior covariant derivative on sections of $P\times_G V$. Every section $\zeta \in \Omega^0(M; P\times_G V)$ defines an equivariant function $\bar\zeta \in \Omega^0_G(P;V)$ obeying $R_g^*\bar\zeta = \varrho(g^{-1})\circ\bar\zeta$ and whose exterior covariant derivative is given by $d^H\bar\zeta = h^* d\bar\zeta$. Applying this to a vector field $u = u_V + hu \in \mathscr{X}(P)$,

$$(d^H\bar\zeta)(u) = d\bar\zeta(hu) = d\bar\zeta(u - u_V) = d\bar\zeta(u) - u_V(\bar\zeta)\,.$$

The derivative $u_V\bar\zeta$ at a point p only depends on the value of u_V at that point, whence we can take $u_V = \sigma(\omega(u))$, so that

$$u_V\bar\zeta = \sigma(\omega(u))\bar\zeta = \frac{d}{dt}\Big|_{t=0} R^*_{g(t)}\bar\zeta \qquad \text{for } g(t) = e^{t\omega(u)}.$$

By equivariance,

$$u_V\bar\zeta = \frac{d}{dt}\Big|_{t=0}\varrho(g(t)^{-1})\circ\bar\zeta = -\varrho(\omega(u))\circ\bar\zeta\,,$$

where we also denote by $\varrho : \mathfrak{g} \to \mathrm{End}(V)$ the representation of the Lie algebra. In summary,

$$(d^H\bar\zeta)(u) = d\bar\zeta(u) + \varrho(\omega)(u)\circ\bar\zeta$$

or, abstracting u,

$$d^H\bar\zeta = d\bar\zeta + \varrho(\omega)\circ\bar\zeta\,. \tag{15}$$

This form is clearly horizontal by construction, and it is also invariant:

$$
\begin{aligned}
R_g^* d^H \bar\zeta &= R_g^* h^* d\bar\zeta \\
\text{(since H is invariant)} \quad &= h^* R_g^* d\bar\zeta \\
\text{(since } d \text{ commutes with pull-backs)} \quad &= h^* d R_g^* \bar\zeta \\
\text{(equivariance of } \bar\zeta) \quad &= h^* d\left(\varrho(g^{-1}) \circ \bar\zeta\right) \\
&= \varrho(g^{-1}) \circ h^* d\bar\zeta \\
&= \varrho(g^{-1}) \circ d^H \bar\zeta \,.
\end{aligned}
$$

As a result, it is a basic form and hence comes from a 1-form $d^H\zeta \in \Omega^1(M; P \times_G V)$. In this way, we have defined a covariant exterior derivative

$$d^H : \Omega^0(M; P \times_G V) \to \Omega^1(M; P \times_G V) \,.$$

Contrary to the exterior derivative, $(d^H)^2\bar\zeta \neq 0$ in general. Instead,

$$
\begin{aligned}
(d^H)^2\bar\zeta &= h^* d h^* d\bar\zeta \\
&= h^* d\left(d\bar\zeta + \varrho(\omega) \circ \bar\zeta\right) \\
&= h^* \left(\varrho(d\omega) \circ \bar\zeta - \varrho(\omega) \wedge d\bar\zeta\right) \\
\text{(since } h^*\omega = 0) \quad &= \varrho(h^* d\omega) \circ \bar\zeta \\
&= \varrho(\Omega) \circ \bar\zeta \,.
\end{aligned}
$$

In other words, the curvature measures the obstruction of the exterior covariant derivative to define a de-Rham-type complex.

This story extends to k-forms in the obvious way. Let $\alpha \in \Omega^k(M; P \times_G V)$ and represent it by a basic form $\bar\alpha \in \Omega^k_G(P; V)$. Define $d^H\bar\alpha = h^* d\bar\alpha$.

Done? ❑ **Exercise 2.5.** Show that

$$d^H\bar\alpha = d\bar\alpha + \varrho(\omega) \wedge \bar\alpha \in \Omega^{k+1}_G(P; V) \,,$$

where $\wedge$ denotes both the wedge product of forms and the composition of the components of $\varrho(\omega)$ with $\bar\alpha$, whence it defines an element $d^H\alpha \in \Omega^{k+1}(M; P \times_G V)$. Furthermore, show that

$$(d^H)^2\bar\alpha = \varrho(\Omega) \wedge \bar\alpha \,.$$

Solution. FIXME: Later. ◆

Let us derive a formula for the covariant derivative of a section $\zeta \in \Omega^k(M; P \times_G V)$ defined locally by a family of forms $\zeta_\alpha \in \Omega^k(U_\alpha; V)$, such that on every nonempty overlap $U_{\alpha\beta}$,

$$\zeta_\alpha = \varrho(g_{\alpha\beta}) \circ \zeta_\beta \,.$$

As seen before, $\zeta_\alpha = s_\alpha^* \bar\zeta$ for $\bar\zeta \in \Omega^k(P; V)$. We define the covariant derivative $d^H\zeta_\alpha$ by pulling back $d^H\bar\zeta$ via the canonical section s_α:

$$
\begin{aligned}
d^H\zeta_\alpha := s_\alpha^* d^H\bar\zeta &= s_\alpha^*\left(d\bar\zeta + \varrho(\omega) \wedge \bar\zeta\right) \\
&= d s_\alpha^*\bar\zeta + \varrho(s_\alpha^*\omega) \wedge s_\alpha^*\bar\zeta \\
&= d\zeta_\alpha + \varrho(A_\alpha) \wedge \zeta_\alpha \,.
\end{aligned}
$$

It is not hard to see, using the transformation properties of A_α and ζ_α on overlaps that on $U_{\alpha\beta}$,

$$d^H\zeta_\alpha = \varrho(g_{\alpha\beta}) \circ d^H\zeta_\beta \,.$$

This result justifies the name "covariant derivative" as used in the Physics literature.

Notation

We will change notation and write the exterior covariant derivative on basic forms as

$$d^{\omega} : \Omega^k_G(P;V) \to \Omega^{k+1}_G(P;V) ,$$

to make manifest the dependence on the connection one-form, and the one on bundle-valued forms on M by

$$d_A : \Omega^k(M; P \times_G V) \to \Omega^{k+1}(M; P \times_G V) ,$$

to make manifest the dependence on the gauge field. For example, in this notation, the Bianchi identity for the curvature can be rewritten as

$$d_A F_A = 0 . \tag{16}$$

Lecture 3: The Yang–Mills equations

In this lecture we will introduce the Yang–Mills action functional on the space of connections and the corresponding Yang–Mills equations. The strategy will be to work locally with the gauge fields and ensure that the objects we construct are gauge-invariant.

Throughout this lecture $P \to M$ will denote a principal G-bundle and $H \subset TP$ a connection with connection one-form ω and curvature two-form Ω. We will let $s_\alpha : U_\alpha \to P$ denote the canonical sections associated to a trivialisation. We will let $A_\alpha = s_\alpha^*$ and $F_\alpha = s_\alpha^*\Omega$ denote the corresponding gauge field and field-strength. On overlaps, the field-strengths are related as in equation (11).

3.1 Some geometry

Until now we have imposed no conditions on M or on G, but this will now change. From now on M will be an oriented pseudo-riemannian n-dimensional manifold with metric g. The orientation on M is given by a nowhere-vanishing n-form, which we will take to be the volume form of the metric.

3.1.1 The volume form

By passing to a refinement, if necessary, we will assume that our trivialising cover $\{U_\alpha\}$ is such that on each U_α the tangent bundle too is trivial. This represents no loss of generality. Then on each U_α we can find one-forms $\theta_i \in \Omega^1(U_\alpha)$ such that the metric takes the form

$$g = \sum_{i=1}^{n} \varepsilon_i \theta_i^2 \,,$$

for some signs ε_i. Let there be s positive and t negative signs. On overlaps, the θ_i will transform by local (special, since M is orientable) orthogonal transformations, but the numbers s and t will not change (Sylvester's law of inertia). We say that M has signature (s,t). Let us define an n-form

$$\theta_1 \wedge \theta_2 \wedge \cdots \wedge \theta_n \in \Omega^n(U_\alpha) \,.$$

on each U_α. The orientability of M implies that these forms agree on overlaps and hence define an n-form $\mathrm{dvol} \in \Omega^n(M)$ called the **volume form** of the metric g. We will assume that dvol gives M its orientation. The volume form allows us to integrate (e.g., compactly supported) functions on M: $\int_M f\,\mathrm{dvol}$ invariantly.

3.1.2 The Hodge $\star$ operator

The metric g defines an inner product $\langle -,- \rangle$ on one-forms by declaring the θ_i to be orthonormal:

$$\langle \theta_i, \theta_j \rangle = \begin{cases} \varepsilon_i \,, & \text{if } i = j, \\ 0 \,, & \text{otherwise,} \end{cases}$$

and extending bilinearly to arbitrary one-forms on U_α. Since on overlaps the θ_i transform by (special) orthogonal transformations, the inner product is well-defined on one-forms on M. Similarly, the metric defines an inner product on k-forms, but to define it, we need to introduce some notation.

A sequence $I = (i_1, \ldots, i_k)$, where $1 \le i_1 < i_2 < \cdots < i_k \le n$, is called a **multi-index of length** $|I| = k$. Let us define $\theta_I := \theta_{i_1} \wedge \theta_{i_2} \wedge \cdots \wedge \theta_{i_k}$. Then every k-form on $\Omega^k(U_\alpha)$ can be written as a linear combination of the $(\theta_I)_{|I|=k}$ with coefficients which are functions on U_α. The inner product on $\Omega^k(U_\alpha)$ is defined by

$$\langle \theta_I, \theta_J \rangle = \begin{cases} \varepsilon(I) \,, & \text{if } I = J, \\ 0 \,, & \text{otherwise,} \end{cases}$$

where $\varepsilon(I) = \varepsilon(i_1)\varepsilon(i_2)\cdots\varepsilon(i_k)$ for $I = (i_1, \ldots, i_k)$, and extending it bilinearly to all of $\Omega^k(U_\alpha)$. As before, the inner product so defined agrees on overlaps and hence extends to an inner product on $\Omega^k(M)$.

We can now define the **Hodge $\star$ operator**: $\star : \Omega^k(M) \to \Omega^{n-k}(M)$ by

$$\alpha \wedge \star\beta = \langle\alpha,\beta\rangle \,\mathrm{dvol}\,,$$

where $\alpha,\beta \in \Omega^k(M)$. We can be more explicit, by showing what the Hodge $\star$ operator does to the θ_I. By definition,

$$\theta_I \wedge \star\theta_I = \varepsilon(I)\,\mathrm{dvol}\,,$$

whence

$$\star\theta_I = \varepsilon(I)\varsigma(I)\theta_{\bar{I}}\,,$$

where $\bar{I}$ is the complementary multi-index to I; that is, the unique multi-index of length $|\bar{I}| = n-k$ such that $I \cup \bar{I} = \{1,2,\ldots,n\}$ (as sets), and $\varsigma(I)$ is the sign of the permutation of $(1,2,\ldots,n)$ given by concatenating $I \sqcup \bar{I}$.

Done? ❑ **Exercise 3.1.** Let $n = 4$ and let g have positive-definite signature $(4,0)$. Calculate the Hodge $\star$ acting on all θ_I. Show that $\star^2 = \mathrm{id}$ on 2-forms. Now do the same for lorentzian signature $(3,1)$ and show that $\star^2 = -\mathrm{id}$ on 2-forms. Can you guess what happens in split signature $(2,2)$?

Solution. There are $2^4 = 16$ possible θ_I in four dimensions. We will use the notation $\theta_{123} = \theta_1 \wedge \theta_2 \wedge \theta_3$, et cetera. In positive-definite signature, we find

$$\begin{array}{llll}
\star\theta_1 = \theta_{234} & \star 1 = \mathrm{dvol} & \star\theta_{23} = \theta_{14} & \star\theta_{123} = \theta_4 \\
\star\theta_2 = -\theta_{134} & \star\theta_{12} = \theta_{34} & \star\theta_{24} = -\theta_{13} & \star\theta_{124} = -\theta_3 \\
\star\theta_3 = \theta_{124} & \star\theta_{13} = -\theta_{24} & \star\theta_{34} = \theta_{12} & \star\theta_{134} = \theta_2 \\
\star\theta_4 = -\theta_{123} & \star\theta_{14} = \theta_{23} & \star\,\mathrm{dvol} = 1 & \star\theta_{234} = -\theta_1
\end{array}$$

from where one sees that $\star^2 = \mathrm{id}$ on 2-forms.

In lorentzian signature, letting θ_0 be time-like and $\mathrm{dvol} = \theta_{0123}$, we find

$$\begin{array}{llll}
\star\theta_0 = -\theta_{123} & \star 1 = \mathrm{dvol} & \star\theta_{12} = \theta_{03} & \star\theta_{012} = -\theta_3 \\
\star\theta_1 = -\theta_{023} & \star\theta_{01} = -\theta_{23} & \star\theta_{13} = -\theta_{02} & \star\theta_{013} = \theta_2 \\
\star\theta_2 = \theta_{013} & \star\theta_{02} = \theta_{13} & \star\theta_{23} = \theta_{01} & \star\theta_{023} = -\theta_1 \\
\star\theta_3 = -\theta_{012} & \star\theta_{03} = -\theta_{12} & \star\,\mathrm{dvol} = -1 & \star\theta_{123} = -\theta_0
\end{array}$$

from where one sees that $\star^2 = -\mathrm{id}$ on 2-forms.

In split signature $(2,2)$, $\star^2 = \mathrm{id}$ on 2-forms, as in positive-definite signature. ◆

Iterating the Hodge $\star$ operator yields a map $\star^2 : \Omega^k(M) \to \Omega^k(M)$. To recognise it, we act on θ_I:

$$\star^2\theta_I = \varepsilon(I)\varsigma(I) \star \theta_{\bar{I}} = \varepsilon(I)\varepsilon(\bar{I})\varsigma(I)\varsigma(\bar{I})\theta_I\,,$$

whence $\star^2$ is a scalar operator, acting as a sign. To work out the sign, notice that $\varepsilon(I)\varepsilon(\bar{I}) = (-1)^t$ and that $\varsigma(I)\varsigma(\bar{I}) = (-1)^{|I||\bar{I}|}$,

$$\star^2 = (-1)^t(-1)^{k(n-k)}\,\mathrm{id} \qquad \text{on } \Omega^k(M).$$

Done? ❑ **Exercise 3.2.** Let M be even-dimensional. Show how the Hodge $\star$ operator transforms under a conformal transformation and show that it is conformally invariant acting on middle-dimensional forms. In other words, rescale the metric on M to $\tilde{g} = e^{2f}g$, and work out the relation between the Hodge operators $\star_g$ and $\star_{\tilde{g}}$. In particular, show that they agree on middle-dimensional forms.

Solution. Let $g = \sum_{i=1}^n \varepsilon_i\theta_i^2$ and $\tilde{g} = \sum_{i=1}^n \varepsilon_i\tilde{\theta}_i^2$. Since $\tilde{g} = e^{2f}g$, we can take $\tilde{\theta}_i = e^f\theta_i$. In particular, $\tilde{\theta}_I = e^{|I|f}\theta_I$. The Hodge operator for $\tilde{g}$ is defined by

$$\star_{\tilde{g}}\tilde{\theta}_I = \varepsilon(I)\varsigma(I)\tilde{\theta}_{\bar{I}}\,,$$

whence

$$\star_{\tilde{g}} e^{|I|f}\theta_I = \varepsilon(I)\varsigma(I)e^{|\bar{I}|f}\theta_{\bar{I}}\,,$$

or equivalently

$$\star_{\tilde{g}}\theta_I = \varepsilon(I)\varsigma(I)e^{(|\bar{I}|-|I|)f}\theta_{\bar{I}} = e^{(|\bar{I}|-|I|)f}\star_g\theta_I\,.$$

The result follows from the fact that in middle dimension $|I| = |\bar{I}|$, whence $\star_{\tilde{g}} = \star_g$. ◆

3.1.3 Inner product on bundle-valued forms

We would also like to define inner products on forms with values in an associated vector bundle $P \times_G V$. Locally, on each U_α, we view such forms as forms with values in V. To define an inner product on such locally defined forms, all we need an inner product on V; but if we want this inner product to glue well on overlaps, we must require that it be G-invariant, so that for all $g \in G$, $\boldsymbol{v}, \boldsymbol{w} \in V$,

$$\langle \varrho(g)\boldsymbol{v}, \varrho(g)\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ .$$

Indeed, if $\zeta \in \Omega^k(M; P \times_G V)$ is represented locally by $\zeta_\alpha \in \Omega^k(U_\alpha; V)$, consider the function $\langle \zeta_\alpha, \zeta_\alpha \rangle \in C^\infty(U_\alpha)$, where $\langle -, - \rangle$ denotes both the inner product on V and the inner product on forms. On a nonempty overlap $U_{\alpha\beta}$,

$$\langle \zeta_\alpha, \zeta_\alpha \rangle = \langle \varrho(g_{\alpha\beta})\zeta_\beta, \varrho(g_{\alpha\beta})\zeta_\beta \rangle = \langle \zeta_\beta, \zeta_\beta \rangle \ ,$$

whence it defines a global function $\langle \zeta, \zeta \rangle \in C^\infty(M)$.

The existence of a G-invariant inner product on V is of course not guaranteed, but if G is compact, for example, then we may always construct one by departing from any positive-definite inner product and averaging over the group with respect to the Haar measure.

In the case of the adjoint bundle $\operatorname{ad} P$, we require an inner product on the Lie algebra $\mathfrak{g}$ which is invariant under the adjoint action of G. For example, if $\mathfrak{g}$ is semisimple then the Killing form κ, defined by

$$\kappa(X, Y) = \operatorname{Tr} \operatorname{ad}_X \operatorname{ad}_Y$$

where $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is defined by $\operatorname{ad}_X Y = [X, Y]$, is a possible such inner product. Of course, there are nonsemisimple (even nonreductive) Lie algebras admitting an ad-invariant inner product; although for a positive-definite inner product $\mathfrak{g}$ must be the Lie algebra of a compact group, hence reductive. In any case we will assume in what follows that $\mathfrak{g}$ has such an inner product.

3.2 The variational problem

3.2.1 The action functional

The gauge field-strengths F_α define a 2-form $F_A \in \Omega^2(M; \operatorname{ad} P)$ whose norm defines a function on M:

$$|F_A|^2 = \langle F_A, F_A \rangle \ .$$

Notation

We may at times use the notation

$$\operatorname{Tr}(F_A \wedge \star F_A) := |F_A|^2 \operatorname{dvol} \in \Omega^n(M) \ .$$

We will define the **Yang–Mills action** to be

$$S_{YM} = \int_M |F_A|^2 \operatorname{dvol} \ , \tag{17}$$

provided that the integral exists. This will be the case for M compact, for example.

The above action does not depend on the choice of local sections used to pull back the curvature two-form to M. Indeed, let $\tilde{s}_\alpha : U_\alpha \to P$ be a different choice of local sections. Let $m \in U_\alpha$ and consider $\tilde{s}_\alpha(m)$ and $s_\alpha(m)$. Since they belong to the same fibre, there exists $h_\alpha(m) \in G$ such that

$$\tilde{s}_\alpha(m) = s_\alpha(m) h_\alpha(m) \ .$$

As m varies, this defines a function $h_\alpha : U_\alpha \to G$. Let $\tilde{F}_\alpha = \tilde{s}_\alpha^* \Omega$. Then for all $m \in U_\alpha$,

$$\begin{aligned}
\tilde{F}_\alpha(m) &= \tilde{s}_\alpha^* \Omega(\tilde{s}_\alpha(m)) \\
&= (R_{h_\alpha(m)} \circ s_\alpha)^* \Omega(s_\alpha(m) h_\alpha(m)) \\
&= s_\alpha^* R_{h_\alpha(m)}^* \Omega(s_\alpha(m) h_\alpha(m)) \\
\text{(since } \Omega \text{ is invariant)} \qquad &= s_\alpha^* \left(\mathrm{ad}_{h_\alpha(m)^{-1}} \circ \Omega(s_\alpha(m))\right) \\
&= \mathrm{ad}_{h_\alpha(m)^{-1}} \circ s_\alpha^* \Omega(s_\alpha(m)) \\
&= \mathrm{ad}_{h_\alpha(m)^{-1}} \circ F_\alpha(m)
\end{aligned}$$

whence, by the ad-invariance of the inner product, $|\tilde{F}|^2 = |F|^2$.

Similarly, the action does not depend on the choice of trivialisation. Indeed, given two trivialisations, we simply pass to a common refinement and use the independence on the choice of local section to show that the norm of the gauge field-strength does not change.

Therefore, if M is compact, then the Yang–Mills action defines a function on the space of connections: $S_{YM} : \mathscr{A} \to \mathbb{R}$. If M is not compact, then we must restrict to connections for which the integral exists. Moreover, the Yang–Mills action is gauge-invariant. Indeed, under a gauge transformation $\Phi \in \mathscr{G} \cong C^\infty(M; \mathrm{Ad}\, P)$

$$F_\alpha \mapsto F_\alpha^\Phi = \mathrm{ad}_{\phi_\alpha} \circ F_\alpha ,$$

whence $|F^\Phi|^2 = |F|^2$ due to the invariance of the inner product on $\mathfrak{g}$. This means that (for M compact) the Yang–Mills action descends to a function $\mathscr{A}/\mathscr{G} \to \mathbb{R}$.

3.2.2 The field equations

A connection A is said to be a **Yang–Mills connection** if it is a critical point of the Yang–Mills action. This means that all directional derivatives of S_{YM} vanish at A. We will now see that this condition turns into a second-order partial differential equation for A.

We recall that $\mathscr{A}$ is an affine space modelled on $\Omega^1(M; \mathrm{ad}\, P)$. This means that the tangent space to $\mathscr{A}$ at any point is isomorphic to $\Omega^1(M; \mathrm{ad}\, P)$. Given a connection $A \in \mathscr{A}$ and a one-form $\tau \in \Omega^1(M; \mathrm{ad}\, P)$, we consider the curve $A + t\tau$ in $\mathscr{A}$ whose tangent vector (at A) is precisely τ. The directional derivative of S_{YM} at A in the direction τ is given by

$$\frac{d}{dt} S_{YM}(A + t\tau)\Big|_{t=0}$$

and the Yang–Mills condition states that this vanishes for all τ. To see what this means, we first compute the curvature along the above curve. Working locally, but omitting the index α associated to the trivialisation, we have from the structure equation:

$$\begin{aligned}
F_{A+t\tau} &= d(A + t\tau) + \tfrac{1}{2}[A + t\tau, A + t\tau] \\
&= F_A + t\left(d\tau + \tfrac{1}{2}[A,\tau] + \tfrac{1}{2}[\tau, A]\right) + \tfrac{1}{2} t^2 [\tau,\tau] \\
&= F_A + t\,(d\tau + [A,\tau]) + \tfrac{1}{2} t^2 [\tau,\tau] \\
&= F_A + t d_A \tau + \tfrac{1}{2} t^2 [\tau,\tau] .
\end{aligned}$$

Computing its norm,

$$\begin{aligned}
|F_{A+t\tau}|^2 &= |F_A + t d_A\tau + \tfrac{1}{2} t^2 [\tau,\tau]|^2 \\
&= \langle F_A + t d_A\tau + \tfrac{1}{2} t^2[\tau,\tau], F_A + t d_A \tau + \tfrac{1}{2} t^2 [\tau,\tau]\rangle \\
&= |F_A|^2 + 2t\,\langle d_A\tau, F_A\rangle + t^2\left(|d_A\tau|^2 + \langle F_A, [\tau,\tau]\rangle\right) + t^3\,\langle d_A\tau, [\tau,\tau]\rangle + \tfrac{1}{4} t^4 |[\tau,\tau]|^2 .
\end{aligned}$$

Therefore, the Yang–Mills condition is

$$0 = \frac{d}{dt} S_{YM}(A + t\tau)\Big|_{t=0} = 2 \int_M \langle d_A\tau, F_A\rangle\, \mathrm{dvol} \qquad \text{for all } \tau \in \Omega^1(M; \mathrm{ad}\, P).$$

Let d_A^* denote the formal adjoint of d_A, so that

$$\int_M \langle d_A\tau, F_A\rangle \,\mathrm{dvol} = \int_M \langle \tau, d_A^* F_A\rangle \,\mathrm{dvol}\,,$$

whence the Yang–Mills condition becomes the following differential equation:

$$d_A^* F_A = 0\,.$$

Done? ❑ **Exercise 3.3.** Show that $\star d_A^* F_A = d\star F_A$.

Solution. Calculating,

$$\begin{aligned}\int_M \langle d_A\tau, F_A\rangle \,\mathrm{dvol} &= \int_M \mathrm{Tr}\,(d_A\tau\wedge\star F_A)\\ &= \int_M \mathrm{Tr}\,((d\tau + [A,\tau])\wedge\star F_A)\\ &= \int_M \mathrm{Tr}\,(\tau\wedge(d\star F_A + [A,\star F_A]))\\ &= \int_M \mathrm{Tr}\,(\tau\wedge d_A\star F_A)\;.\end{aligned}$$

Comparing with

$$\int_M \langle \tau, d_A^* F_A\rangle \,\mathrm{dvol} = \int_M \mathrm{Tr}\left(\tau\wedge\star d_A^* F_A\right)\,,$$

gives the desired answer. ◆

We therefore conclude that the Yang–Mills condition is equivalent to the equation

$$d_A\star F_A = 0\,, \tag{18}$$

which together with the Bianchi identity $d_A F_A = 0$ constitutes of a nonlinear version of the conditions for a 2-form to be harmonic.

Notice that because the Yang–Mills action is gauge-invariant, if A solves the Yang–Mills equations, so will any gauge transformed A^Φ. In other words, the gauge group acts on the space $\mathscr{A}_{YM}$ of Yang–Mills connections. The quotient $\mathscr{A}_{YM}/\mathscr{G}$ is the space of classical solutions. In general it is infinite-dimensional, but we will see that it has interesting finite-dimensional subspaces.

3.3 Coupling to matter

Gauge fields are responsible for the "forces" in Nature. Matter fields, on the other hand, are modelled as sections of certain bundles over M. For bosonic matter fields, these are simply associated fibre bundles to P: typically associated vector bundles, but more generally associated fibre bundles in the case of nonlinear realisations (σ-models,...). Fermionic matter fields are sections of a tensor product of a spinor bundle on M (assumed spin) and an associated vector bundle to P.

For simplicity, let us consider a bosonic matter field φ which is a section of an associated vector bundle $P\times_G V$ over M with representation $\varrho : G \to GL(V)$, preserving an inner product $\langle -,-\rangle$ on V. Let $d_A : \Omega^0(M; P\times_G V) \to \Omega^1(M; P\times_G V)$ denote the covariant derivative and let $|d_A\varphi|^2 \in C^\infty(M)$ denote the (squared) norm of $d_A\varphi$ using both the inner product on forms and the one on V. The coupling of this matter to the gauge fields is described by the action functional

$$S_{\text{matter}} = \tfrac{1}{2}\int_M |d_A\varphi|^2\,\mathrm{dvol}\;.$$

Done? ❑ **Exercise 3.4.** Show that the field equation for φ obtained by extremising the above action is given by

$$d_A \star d_A \varphi = 0\,,$$

which is a nonlinear version of Laplace's equation.

Solution. As $\Omega^0(M; P \times_G V)$ is a vector space, it is canonically identified with its tangent space. Hence if $\varphi, \alpha \in \Omega^0(M; P \times_G V)$, $\varphi + t\alpha$ is a curve in $\Omega^0(M; P \times_G V)$ passing through φ with velocity α. We demand that φ be a critical point of the action S_{matter}, so that

$$\frac{d}{dt} S_{\text{matter}}(\varphi + t\alpha)\Big|_{t=0} = 0 \qquad \text{for all } \alpha.$$

Evaluating the action along this curve, one finds

$$\begin{aligned} S_{\text{matter}}(\varphi + t\alpha) &= \tfrac{1}{2}\int_M |d_A\varphi + t d_A\alpha|^2 \,\text{dvol} \\ &= \int_M \left(\tfrac{1}{2}|d_A\varphi|^2 + t\,\langle d_A\alpha, d_A\varphi\rangle + \tfrac{1}{2}t^2|d_A\alpha|^2\right)\text{dvol}\,, \end{aligned}$$

whence the equation of motion is

$$\int_M \langle d_A\alpha, d_A\varphi\rangle\,\text{dvol} = \int_M \langle \alpha, d_A^* d_A\varphi\rangle\,\text{dvol} = 0\,,$$

for all α — in other words, $d_A^* d_A\varphi = 0$. Finally, an argument identical to the one in Exercise 3.3 shows that $\star d_A^* d_A\varphi = -d_A \star d_A\varphi$, whence the equation of motion is as asked. ◆

Of course, the inclusion of matter fields also changes the Yang–Mills equations. It's easy enough to work out the new equations by demanding that A be a critical point of the action $S_{\text{YM}} + S_{\text{matter}} : \mathscr{A} \to \mathbb{R}$, for fixed φ.

Done? ❑ **Exercise 3.5.** Show that in the presence of the matter field φ the Yang–Mills equations are modified by a quadratic term in φ:

$$d_A^* F_A + T(A,\varphi) = 0\,,$$

where $T = T(A,\varphi) \in \Omega^1(M; \operatorname{ad} P)$ is defined by

$$\langle T, \tau\rangle = \langle d_A\varphi, \varrho(\tau)\varphi\rangle$$

for every $\tau \in \Omega^1(M; \operatorname{ad} P)$.

Solution. To find how the Yang–Mills equation is modified by the presence of this matter, we compute the action S_{matter} along a curve $A + t\tau \in \mathscr{A}$. The covariant derivative is $d_{A+t\tau}\varphi = d_A\varphi + t\varrho(\tau)\varphi$, whence

$$|d_{A+t\tau}\varphi|^2 = |d_A\varphi|^2 + 2t\,\langle d_A\varphi, \varrho(\tau)\varphi\rangle + t^2|\varrho(\tau)\varphi|^2\,.$$

Its contribution to the equation of motion is

$$\begin{aligned} \frac{d}{dt} S_{\text{matter}}(A + t\tau)\Big|_{t=0} &= \int_M \langle d_A\varphi, \varrho(\tau)\varphi\rangle\,\text{dvol} \\ &= \int_M \langle T, \tau\rangle\,\text{dvol}\,, \end{aligned}$$

which defines $T \in \Omega^1(M; \operatorname{ad} P)$. ◆

Lecture 4: Instantons

Forget it all for an instanton!
— (not quite) The National Lottery

In this lecture we will specialise to the case of a four-dimensional riemannian manifold M and introduce the notion of (anti-)self-dual connection, the so-called instantons. We will establish a lower bound for the Yang–Mills action and show that instantons saturate this bound, so they correspond to minima of the action.

4.1 (Anti-)self-duality

Let (M, g) be a four-dimensional oriented riemannian manifold. We saw in Exercise 3.1 that in this dimension and signature, the Hodge $\star$ operator obeys $\star^2 = \mathrm{id}$ acting on 2-forms. This allows us to decompose the vector space of 2-forms into eigenspaces of $\star$:

$$\Omega^2(M) = \Omega^2_+(M) \oplus \Omega^2_-(M) ,$$

where a 2-form $\omega \in \Omega^2_\pm(M)$ if and only if $\star\omega = \pm\omega$. We will say that ω is **self-dual** if $\omega \in \Omega^2+(M)$ and **anti-self-dual** if $\omega \in \Omega^2_-(M)$. Every 2-form ω can therefore be written uniquely as a linear combination of a self-dual and an anti-self-dual form $\omega = \omega_+ + \omega_-$, with $\omega_\pm \in \Omega^2_\pm(M)$. Furthermore this decomposition is orthogonal with respect to the inner product. Indeed, on the one hand

$$\langle\omega_+, \omega_-\rangle \,\mathrm{dvol} = \omega_+ \wedge \star\omega_- = -\omega_+ \wedge \omega_- ,$$

but also

$$\langle\omega_+, \omega_-\rangle \,\mathrm{dvol} = \langle\omega_-, \omega_+\rangle \,\mathrm{dvol} = \omega_- \wedge \star\omega_+ = \omega_- \wedge \omega_+ = \omega_+ \wedge \omega_- ,$$

whence $\langle\omega_+, \omega_-\rangle = 0$.

The same results also hold in the case of 2-forms with values in vector bundles with inner products. In particular, it applies to the gauge field strength $F_A \in \Omega^2(M; \mathrm{ad}\, P)$ of a connection on a principal fibre bundle P over M. Decomposing $F_A = F_A^+ + F_A^-$ into its self-dual and anti-self-dual parts, the Yang–Mills action (16) is a sum of two terms (provided that the integrals exist):

$$S_{YM} = \int_M |F_A|^2 \,\mathrm{dvol} = \int_M |F_A^+|^2 \,\mathrm{dvol} + \int_M |F_A^-|^2 \,\mathrm{dvol} ,$$

each one being positive-semidefinite.

Consider now the integral over M

$$c := \int_M \mathrm{Tr}\, F_A \wedge F_A$$

of the 4-form $\mathrm{Tr}\, F_A \wedge F_A$. Decomposing F_A into its self-dual and anti-self-dual components, we can rewrite this integral as the difference

$$c = \int_M |F_A^+|^2 \,\mathrm{dvol} - \int_M |F_A^-|^2 \,\mathrm{dvol} ,$$

where the mixed terms are absent because F_A^+ and F_A^- are perpendicular. This implies the following bound for the Yang–Mills action

$$S_{YM} \geq |c| , \tag{19}$$

with equality if and only if $F_A^\pm = 0$ in which case

$$S_{YM} = \mp c .$$

Finally notice that if $F_A^\pm = 0$ then F_A satisfies the Yang–Mills equation (17), by virtue of the Bianchi identity (15).

Notation

If $F_A = F_A^{\pm}$ we will say that the connection is **(anti-)self-dual** and we say that the gauge field describes an **(anti-)instanton**.

Notice that the (anti-)self-duality condition is a first order partial differential equation for the connection, whereas the Yang–Mills equation is of second order. Hence imposing (anti-)self-duality is a way of finding solutions of a second-order partial differential equation via first order equations. This is reminiscent of *supersymmetry* and in fact there is a deep relation between instantons and supersymmetry.

4.2 What is c?

We have shown that the Yang–Mills action is bounded below by a number: (the absolute value of) the integral of the 4-form $\Theta = \operatorname{Tr} F_A \wedge F_A$ over M. Since M is 4-dimensional, Θ is closed for dimensional reasons; however

Done? ❑ **Exercise 4.1.** Show that Θ is a closed 4-form even if $\dim M > 4$.

Solution. This follows trivially from the Bianchi identity:

$$
\begin{aligned}
d\Theta &= d \operatorname{Tr} F_A \wedge F_A \\
&= 2 \operatorname{Tr} dF_A \wedge F_A \\
\text{(by Bianchi)} \quad &= -2 \operatorname{Tr}[A, F_A] \wedge F_A \\
\text{(ad-invariance of Tr)} \quad &= 2 \operatorname{Tr} A \wedge [F_A, F_A] \\
&= 0 \,.
\end{aligned}
$$

◆

Therefore Θ defines a class $[\Theta] \in H^4_{dR}(M)$ in de Rham cohomology and c is the evaluation of this class on the fundamental class $[M] \in H_4(M)$. We will now show that c is independent of the connection, as the notation already suggests, so that it is a *characteristic number* of the bundle.

Recall that the space $\mathscr{A}$ of connections is an affine space locally modelled on $\Omega^1(M; \operatorname{ad} P)$. This means that if $A_0, A_1 \in \mathscr{A}$, then the straight line

$$A_t := A_0 + t(A_1 - A_0)$$

lies in $\mathscr{A}$. Let $\tau = A_1 - A_0 \in \Omega^1(M; \operatorname{ad} P)$. Let us introduce the notation $d_t := d_{A_t}$ and $F_t := F_{A_t}$. One has

$$F_t = F_0 + t d_0 \tau + \tfrac{1}{2} t^2 [\tau, \tau] \,.$$

Notice that

$$\frac{dF_t}{dt} = d_0 \tau + t[\tau, \tau] = d_t \tau \,.$$

Let $\Theta_t = \operatorname{Tr} F_t \wedge F_t$. Differentiating, we obtain

$$\frac{d\Theta_t}{dt} = 2 \operatorname{Tr}(d_t \tau \wedge F_t) \,.$$

On the other hand,

$$
\begin{aligned}
d \operatorname{Tr}(\tau \wedge F_t) &= \operatorname{Tr}(d\tau \wedge F_t - \tau \wedge dF_t) \\
\text{(by Bianchi)} \quad &= \operatorname{Tr}(d\tau \wedge F_t + \tau \wedge [A_t, F_t]) \\
\text{(ad-invariance of Tr)} \quad &= \operatorname{Tr}(d_t \tau \wedge F_t) \,.
\end{aligned}
$$

In other words,

$$\frac{d\Theta_t}{dt} = \tfrac{1}{2} d\,\mathrm{Tr}(\tau \wedge \mathrm{F}_t)\ ,$$

whence integrating with respect to t over $[0,1]$, we obtain

$$\Theta_1 - \Theta_0 = d\left(\tfrac{1}{2}\int_0^1 \tau \wedge \mathrm{F}_t\right)\ .$$

In particular, in cohomology, $[\Theta_1] = [\Theta_0]$ and hence c is a constant on $\mathscr{A}$. In fact, up to a factor, it is the first Pontrjagin number of the adjoint bundle $\mathrm{ad}\,\mathrm{P}$:

$$p_1(\mathrm{ad}\,\mathrm{P})[\mathrm{M}] = \frac{1}{4\pi^2}\int_{\mathrm{M}} \mathrm{Tr}\,\mathrm{F_A} \wedge \mathrm{F_A} \implies c = 4\pi^2 p_1(\mathrm{ad}\,\mathrm{P})[\mathrm{M}]\ .$$

The factor of $4\pi^2$ depends on the normalisation of the inner product Tr on the Lie algebra. We have made a choice here which is correct for $\mathfrak{g} = \mathfrak{su}(2)$ where the inner product is the natural one identifying $\mathfrak{su}(2) = \mathfrak{sp}(1) = \mathrm{Im}\mathbb{H}$.

One can show that $p_1(\mathrm{ad}\,\mathrm{P})[\mathrm{M}]$ is an integer, which in the present context is called the **instanton number** and usually denoted k. Hence, we can rewrite the bound (18) on the Yang–Mills action as

$$\mathrm{S_{YM}} \geq 4\pi^2 |k|\ , \tag{20}$$

for some integer k.

4.3 The Chern–Simons form

We can pull back Θ to P using the projection: $\pi^*\Theta$. Since d commutes with pull-backs, $\pi^*\Theta$ is also closed, but in fact we have

Done? ❑ **Exercise 4.2.** Show that $\pi^*\Theta \in \Omega^4(\mathrm{P})$ is exact:

$$\pi^*\Theta = d\,\mathrm{Tr}\left(\omega \wedge \left(d\omega + \tfrac{1}{3}[\omega,\omega]\right)\right)\ .$$

Solution. This can be done directly by writing $\pi^*\Theta = \mathrm{Tr}\,\Omega \wedge \Omega$ and using the structure equation. Here, however, is a trick due to Chern. Rescale the connection 1-form $\omega_t := t\omega$ and let $\Omega_t := d\omega_t + \tfrac{1}{2}[\omega_t,\omega_t] = t\Omega + \tfrac{1}{2}(t^2 - t)[\omega,\omega]$. It is important to remark that ω_t is not a connection 1-form (for $t \neq 1$) and that hence Ω_t is not a curvature 2-form, yet the Bianchi identity $d^{\omega_t}\Omega_t = 0$ still holds. Notice that by the Fundamental Theorem of Calculus,

$$\mathrm{Tr}\,\Omega \wedge \Omega = \int_0^1 \frac{d}{dt}\mathrm{Tr}(\Omega_t \wedge \Omega_t)\,dt\ .$$

Alternatively, we can differentiate with respect to t on the RHS to obtain

$$\begin{aligned}
\mathrm{Tr}\,\Omega \wedge \Omega &= \int_0^1 2\,\mathrm{Tr}\left(\frac{d\Omega_t}{dt} \wedge \Omega_t\right) dt \\
&= 2\int_0^1 \mathrm{Tr}\left((d\omega + t[\omega,\omega]) \wedge \Omega_t\right) dt \\
\text{(by parts)} \quad &= 2\int_0^1 d\,\mathrm{Tr}(\omega \wedge \Omega_t)\,dt + 2\int_0^1 \mathrm{Tr}\left(\omega \wedge (d\Omega_t + [t\omega,\Omega_t])\right) dt \\
\text{(by Bianchi)} \quad &= d\int_0^1 2\,\mathrm{Tr}(\omega \wedge \Omega_t)\,dt\ .
\end{aligned}$$

Evaluating the integral in the RHS, we obtain

$$\mathrm{Tr}\,\Omega \wedge \Omega = d\left(\omega \wedge d\omega + \tfrac{1}{3}\omega \wedge [\omega,\omega]\right)\ .$$

◆

We can now pull-back the 3-form

$$\operatorname{Tr}\left(\omega \wedge d\omega + \tfrac{1}{3}\omega \wedge [\omega,\omega]\right)$$

via the canonical sections $s_\alpha : U_\alpha \to P$. On each trivialising neighbourhood U_α we have the **Chern–Simons 3-form**

$$\Xi_\alpha := \operatorname{Tr}\left(A_\alpha \wedge dA_\alpha + \tfrac{1}{3}A_\alpha \wedge [A_\alpha, A_\alpha]\right) \in \Omega^3(U_\alpha)\ .$$

By construction, we have on each U_α,

$$d\Xi_\alpha = \operatorname{Tr} F_A \wedge F_A\ ,$$

whence on double overlaps $U_\alpha \cap U_\beta$, $d\Xi_\alpha = d\Xi_\beta$, so that $\Xi_\alpha - \Xi_\beta$ is a closed 3-form.

Done? ❑ **Exercise 4.3.** Show that on each double overlap $U_\alpha \cap U_\beta$,

$$\Xi_\alpha - \Xi_\beta = g_{\alpha\beta}^* \left(\tfrac{1}{6}\operatorname{Tr}(\theta \wedge [\theta,\theta])\right)\ ,$$

where θ is the Maurer–Cartan 1-form on G.

Solution. This is a straight-forward (if a little tedious) calculation, which can be made a little easier by working with a matrix group and using equation (5). ◆

4.4 The BPST instanton

We will now take $M = \mathbb{R}^4$. This is not compact and we have to be careful with the convergence of the integrals. We will be concerned with Yang–Mills connections with **finite action** : those for which the Yang–Mills action converges. In particular, this means that the field strength vanishes sufficiently fast at infinity. Euclidean space $\mathbb{R}^4$ is conformally equivalent to the 4-sphere S^4 with a point removed, as can be seen immediately using stereographic projection. Now, it follows from Exercise 3.2 that the (anti)self-duality conditions are conformally invariant. Hence if an instanton on $\mathbb{R}^4$ has finite action *and* it extends to the point at infinity, it defines an instanton on S^4. The simplest such example is the so-called BPST instanton, named after its discoverers: Belavin, Polyakov, Schwarz and Tyupkin. The BPST instanton is a connection on a nontrivial principal SU(2)-bundle over S^4 whose total space is in fact the 7-sphere. This is a generalisation of the classical Hopf fibration $S^3 \to S^2$ responsible for the Dirac monopole. Let us describe it in more detail.

Like many interesting results in Physics, the construction of the BPST instanton stems from a seemingly un-natural identification: in this case, from an embedding of the Lie algebra of SU(2) into the space M. To explain this it is convenient to work in terms of quaternions. We will identify $\mathbb{R}^4$ with the quaternions $\mathbb{H}$:

$$\mathbb{R}^4 \ni \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto x_1 + x_2 i + x_3 j + x_4 k \in \mathbb{H}\ .$$

We will denote by $\boldsymbol{x}$ also the corresponding quaternion. We denote by $\operatorname{Re}\boldsymbol{x} = x_1$ and $\operatorname{Im}\boldsymbol{x} = x_2 i + x_3 j + x_4 k$ the real and imaginary parts of the quaternion $\boldsymbol{x}$, respectively. As with the complex numbers, quaternionic conjugation merely changes the sign of the imaginary part:

$$\overline{\boldsymbol{x}} = x_1 - x_2 i - x_3 j - x_4 k\ .$$

The euclidean inner product on $\mathbb{R}^4$ agrees with the quaternionic inner product: $\boldsymbol{x}\cdot\boldsymbol{y} = \operatorname{Re}(\boldsymbol{x}\overline{\boldsymbol{y}})$. We will denote the corresponding norm by $|\boldsymbol{x}|^2 = \operatorname{Re}(\boldsymbol{x}\overline{\boldsymbol{x}})$.

The Lie group SU(2) also has a quaternionic interpretation. Indeed, it is isomorphic to the group Sp(1) of unit quaternions:

$$\operatorname{Sp}(1) = \left\{\boldsymbol{x} \in \mathbb{H} \middle| |\boldsymbol{x}|^2 = 1\right\}\ ,$$

and this isomorphism induces one of Lie algebras $\mathfrak{su}(2) \cong \mathfrak{sp}(1)$, which is itself isomorphic to the imaginary quaternions $\operatorname{Im}\mathbb{H}$.

We now introduce the following imaginary quaternion-valued 1-form on $\mathbb{H}$,

$$\mathrm{A}(\boldsymbol{x}) = \frac{1}{|\boldsymbol{x}|^2+1}\operatorname{Im}(\boldsymbol{x}\,d\overline{\boldsymbol{x}}) ,$$

which we interpret as an $\mathfrak{su}(2)$-valued 1-form on $\mathbb{R}^4$ and hence as a gauge field. The corresponding field-strength is given by

$$\mathrm{F}(\boldsymbol{x}) = \frac{1}{(|\boldsymbol{x}|^2+1)^2}\, d\boldsymbol{x}\wedge d\overline{\boldsymbol{x}} ,$$

where $\wedge$ means both the wedge product of 1-forms and quaternionic multiplication. Let us unpack this:

$$\begin{aligned} d\boldsymbol{x}\wedge d\overline{\boldsymbol{x}} &= (dx_1 + dx_2 i + dx_3 j + dx_4 k)\wedge(dx_1 - dx_2 i - dx_3 j - dx_4 k)\\ &= -2(dx_{12}+dx_{34})i - 2(dx_{13}-dx_{24})j - 2(dx_{14}+dx_{23})k , \end{aligned}$$

where we have used the notation $dx_{12} = dx_1 \wedge dx_2$, etc. It is evident from the above that $d\boldsymbol{x}\wedge d\overline{\boldsymbol{x}}$ is an $\operatorname{Im}\mathbb{H}$-valued self-dual 2-form, and hence so is the field-strength F. Therefore the gauge field A defines an SU(2) instanton on $\mathbb{R}^4$. To determine its instanton number, we need only integrate

$$\begin{aligned} k &= \frac{1}{4\pi^2}\int_{\mathbb{R}^4} |\mathrm{F}|^2 d^4x\\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^4} \frac{4}{(|\boldsymbol{x}|^2+1)^4}\left|(dx_{12}+dx_{34})i + (dx_{13}-dx_{24})j + (dx_{14}+dx_{23})k\right|^2 d^4x\\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^4} \frac{4}{(|\boldsymbol{x}|^2+1)^4}\left(|dx_{12}+dx_{34}|^2 + |dx_{13}-dx_{24}|^2 + |dx_{14}+dx_{23}|^2\right) d^4x\\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^4} \frac{24}{(|\boldsymbol{x}|^2+1)^4} d^4x , \end{aligned}$$

where we have used that $|dx_{12}+dx_{34}|^2 = 2$ and similarly for the other two self-dual 2-forms. This is an elementary integral, whose evaluation is simplified by going to spherical polar coordinates:

$$k = \frac{6}{\pi^2}\operatorname{Vol}(S^3)\int_0^\infty \frac{r^3\,dr}{(r^2+1)^4} = \frac{1}{2\pi^2}\operatorname{Vol}(S^3) = 1 ,$$

where we have used that the volume of the unit sphere in $\mathbb{R}^4$ is $2\pi^2$. (Show this!)

Done? ❑ **Exercise 4.4.** Let $\lambda > 0$ be a positive real number and $\boldsymbol{x}_0 \in \mathbb{H}$ a fixed quaternion. Calculate the field-strength of the gauge field

$$\mathrm{A}_{\lambda,\boldsymbol{x}_0}(\boldsymbol{x}) = \frac{1}{|\boldsymbol{x}-\boldsymbol{x}_0|^2+\lambda^2}\operatorname{Im}\left((\boldsymbol{x}-\boldsymbol{x}_0)\,d\overline{\boldsymbol{x}}\right)$$

and show that this defines a $k = 1$ instanton. Convince yourself that as $\lambda \to 0$ the instanton becomes concentrated at $\boldsymbol{x}_0$. (You may wish to visualise what is going on by plotting $|\mathrm{F}|^2$ as a function of $|\boldsymbol{x}-\boldsymbol{x}_0|$ for several values of λ.)

Solution. The clever way to solve this problem is to notice that the self-duality condition is invariant under the action of the conformal group and that $\mathrm{A}_{\lambda,\boldsymbol{x}_0}$ is the pull-back of $\mathrm{A} = \mathrm{A}_{1,\boldsymbol{0}}$ by a conformal transformation:

$$\boldsymbol{x} \mapsto \lambda^{-1}(\boldsymbol{x}-\boldsymbol{x}_0) .$$

It follows from this that in the limit as $\lambda \to 0$, the instanton concentrates at $\boldsymbol{x}_0$. You can also interpret $\mathrm{P} := 4\pi^2|\mathrm{F}|^2$ as a probability density and calculate its variance:

$$\sigma^2 = \langle|\boldsymbol{x}-\boldsymbol{x}_0|^2\rangle - \langle|\boldsymbol{x}-\boldsymbol{x}_0|\rangle^2 ,$$

where $\langle - \rangle$ denotes the expectation value with respect to P:

$$\langle f(\boldsymbol{x})\rangle = \int_{\mathbb{R}^4} f(\boldsymbol{x})\mathrm{P}(\boldsymbol{x})\,d^4x .$$

A simple calculation shows that

$$\langle |\boldsymbol{x} - \boldsymbol{x}_0| \rangle = \tfrac{3}{8}\pi\lambda \qquad \text{and} \qquad \langle |\boldsymbol{x} - \boldsymbol{x}_0|^2 \rangle = 2\lambda^2 ,$$

whence

$$\sigma^2 \approx 0.6\lambda^2 ,$$

which goes to 0 as $\lambda \to 0$. ◆

Lecture 5: Instanton moduli space

The steady progress of physics requires for its theoretical formulation a mathematics that gets continually more advanced.

— PAM Dirac, 1931

In the previous lecture we constructed a $k = 1$ SU(2) instanton on S^4 and in fact saw that it belongs to a five-parameter family of such instantons. This is not an accident and in this lecture we will see that there is a moduli space of instantons, which is a disjoint union of a countable number of finite-dimensional connected subspaces labelled by the instanton number. To a first approximation, the moduli space is the quotient of the space of (anti)self-dual connection modulo gauge transformations. However this space turns out to be singular in general and in order to guarantee a smooth quotient we will have either to restrict ourselves to irreducible connections, or else quotient by a (cofinite) subgroup of gauge transformations.

5.1 Irreducible connections

Throughout this section we will let $P \to M$ be a fixed principal G-bundle with connection $H \subset TP$. Let ω denote the connection 1-form. A smooth curve $\tilde{\gamma} : [0,1] \to P$ is said to be **horizontal** if the velocity vector is everywhere horizontal: $\dot{\tilde{\gamma}}(t) \in H_{\tilde{\gamma}(t)}$ for all t. This is equivalent to $\omega(\dot{\tilde{\gamma}}(t)) = 0$. Let $\gamma(t) = \pi(\tilde{\gamma}(t))$ denote the projection of the curve onto M. Assume that the curve is small enough so that the image of γ lies inside some trivialising neighbourhood U_α. Then $\psi_\alpha(\tilde{\gamma}(t)) = (\gamma(t), g(t))$, where $g(t)$ is a smooth curve on G.

Done? ❑ **Exercise 5.1.** Show that the condition $\omega(\dot{\tilde{\gamma}}(t)) = 0$ translates into the following ordinary differential equation for the curve $g(t)$:

$$\mathrm{ad}_{g(t)^{-1}} A_\alpha(\dot{\gamma}(t)) + (g^*\theta)(\dot{\gamma}(t)) = 0 \,, \tag{21}$$

where A_α is the gauge field on U_α corresponding to the connection, and θ is the left-invariant Maurer–Cartan 1-form on G. Show further that for a matrix group, this equation becomes

$$\dot{g}(t) + A_\alpha(\dot{\gamma}(t))g(t) = 0 \,. \tag{22}$$

Solution. Equation (20) follows immediately from Proposition 1.3 once we notice that $\pi_* \dot{\tilde{\gamma}} = \dot{\gamma}$. As for equation (21), it follows from equation (20) using that $(g^*\theta)(\dot{\gamma}(t)) = g(t)^{-1}\dot{g}(t)$. ◆

Being a first-order ordinary differential equations with smooth coefficients, equation (20) (equivalently (21)) has a unique solution for specified initial conditions, so that if we specify $g(0)$ then $g(1)$ is determined uniquely. This then defines a map $\Pi_\gamma : P_{\gamma(0)} \to P_{\gamma(1)}$ from the fibre over $\gamma(0)$ to the fibre over $\gamma(1)$, associated to the curve $\gamma : [0,1] \to M$. Rephrasing, given the curve γ, there is a unique horizontal lift $\tilde{\gamma}$ once we specify $\tilde{\gamma}(0) \in P_{\gamma(0)}$ and $\Pi_\gamma \tilde{\gamma}(0) = \tilde{\gamma}(1)$ is simply the endpoint of this horizontal curve. The map Π_γ is called **parallel transport along** γ **with respect to the connection** H.

Now let γ be a loop, so that $\gamma(0) = \gamma(1)$. Parallel transport along γ defines a group element $g_\gamma \in G$ defined by $g_\gamma = g(1)g(0)^{-1}$. To show that this element is well-defined, we need to show that it does not depend on the initial point $g(0)$. Indeed, suppose we choose a different starting point $\overline{g}(0)$. Then there is some group element $h \in G$ such that $\overline{g}(0) = g(0)h$. The parallel-transport equations (20) and (21) are clearly invariant under the right G action, whence $\overline{g}(t) := g(t)h$ solves the equation with initial condition $\overline{g}(0)$. Therefore the final point of the curve is $\overline{g}(1) = g(1)h$, whence $\overline{g}(1)\overline{g}(0)^{-1} = g(1)g(0)^{-1}$ and g_γ is well-defined. This defines a map from piecewise-smooth loops based at $m = \gamma(0)$ to G, whose image is a subgroup of G called the **holonomy group of the connection at** m denoted

$$\mathrm{Hol}_m(\omega) = \left\{ g_\gamma \middle| \gamma : [0,1] \to M \,,\ \gamma(1) = \gamma(0) = m \right\} \,. \tag{23}$$

Done? ❑ **Exercise 5.2.** Show that the holonomy group is indeed a subgroup of G; that is, show that it is closed under group multiplication. More precisely, if g_{γ_1} and g_{γ_2} are elements in $\mathrm{Hol}_m(\omega)$, then show that so is their product by exhibiting a loop γ based at m such that $g_\gamma = g_{\gamma_1} g_{\gamma_2}$. Further show that if $m, m' \in M$ belong to the same connected component, the holonomy groups $\mathrm{Hol}_m(\omega)$ and $\mathrm{Hol}_{m'}(\omega)$ are conjugate in G and hence isomorphic.

Solution. If (γ_1, g_1) is a horizontal lift of γ_1 (in a trivialisation) and (γ_2, g_2) a horizontal lift of γ_2, then $g_{\gamma_1} = g_1(1)g_1(0)^{-1}$ and $g_{\gamma_2} = g_2(1)g_2(0)^{-1}$, whence $g_{\gamma_1} g_{\gamma_2} = g_1(1)g_1(0)^{-1}g_2(1)g_2(0)^{-1}$. Let us consider the piecewise smooth curve

$$\gamma(t) = \begin{cases} \gamma_2(2t)\,, & t \in [0, \frac{1}{2}] \\ \gamma_1(2t-1)\,, & t \in [\frac{1}{2}, 1]\,. \end{cases}$$

A continuous horizontal lift of this curve is given (in a trivialisation) by $(\gamma(t), g(t))$ where

$$g(t) = \begin{cases} g_2(2t)\,, & t \in [0, \frac{1}{2}] \\ g_1(2t-1)\,, & t \in [0, \frac{1}{2}] \end{cases}$$

where, for continuity, we choose the horizontal lift of γ_1 in such a way that $g_1(0) = g_2(1)$. Then

$$g_{\gamma_1} g_{\gamma_2} = g_1(1)g_1(0)^{-1}g_2(1)g_2(0)^{-1} = g_1(1)g_2(0)^{-1} = g(1)g(0)^{-1} = g_\gamma\,.$$

Now suppose that m, m' lie in the same connected component of M. For a manifold this means that they lie in the same path component, whence there is a curve $\delta : [0,1] \to M$ such that $\delta(0) = m$ and $\delta(1) = m'$. Let $\delta^{-1} : [0,1] \to M$ be the curve $\delta^{-1}(t) = \delta(1-t)$: the same image as δ but traced backward. Then there is a one-to-one correspondence between loops based at m and based at m'. Indeed, if γ' is a loop based at m' then the composition $\gamma = \delta^{-1} \circ \gamma' \circ \delta$ is a loop based at m; and vice versa. Arguments similar to the ones above show that the element g_γ of the holonomy group at m is given by $h g_{\gamma'} h^{-1}$ where h is the group element corresponding to $\delta(0)$ in the trivialisation. This shows that $\mathrm{Hol}_m(\omega)$ and $\mathrm{Hol}_{m'}(\omega)$ are conjugate subgroups of G. ◆

It follows from the previous exercise, that if M is connected, then the holonomy group of the connection is well-defined as a conjugacy class of subgroups of G. A connection is said to be **irreducible** if the holonomy group is precisely G and not a proper subgroup. The importance of the concept of irreducibility is that the group $\mathscr{G}$ of gauge transformations acts (almost) freely on the space of irreducible connections. The key observation is the covariance of parallel transport under gauge transformations.

Done? ❑ **Exercise 5.3.** Let $\Phi \in \mathscr{G}$ be a gauge transformation and let $\gamma : [0,1] \to M$ be a curve on M. Let Π_γ and Π^Φ_γ denote the operations of parallel transport along γ with respect to the connections H and H^Φ, respectively. Show that

$$\Phi_{\gamma(1)} \circ \Pi_\gamma = \Pi^\Phi_\gamma \circ \Phi_{\gamma(0)}\,. \tag{24}$$

Solution. Let $\Phi \in \mathscr{G} = C^\infty(M; \mathrm{Ad}\,P)$ be a gauge transformation. Locally it is equivalent to a family $\{\phi_\alpha \in C^\infty(U_\alpha, G)\}$. It is a simple calculation to show that if $(\gamma(t), g(t))$ is a horizontal lift (in a trivialisation) of $\gamma(t)$ relative to H, then $(\gamma(t), \phi_\alpha(\gamma(t))g(t))$ is a horizontal lift of γ relative to the gauge transformed connection H^Φ. This follows by showing that $\phi_\alpha(\gamma(t))g(t)$ solves the parallel transport equation (20) with A replaced by A^Φ, which is given by equation (8). Therefore

$$\Pi^\Phi_\gamma \phi_\alpha(\gamma(0))g(0) = \phi_\alpha(\gamma(1))g(1) = \phi_\alpha(\gamma(1))\Pi_\gamma g(0)\,.$$

Abstracting the $g(0)$, we obtain the desired equation. ◆

Suppose now that H is a connection which is fixed by a gauge transformation $\Phi \in \mathscr{G}$. Then for all curves γ, $\Pi_\gamma = \Pi^\Phi_\gamma$, and in particular for all *loops*,

$$\Phi_{\gamma(0)} \circ \Pi_\gamma = \Pi_\gamma \circ \Phi_{\gamma(0)}\,.$$

If the connection is irreducible, then every group element in G is realisable as Π_γ for some loop γ, and the above equation says that $\Phi_{\gamma(0)}$ commutes with all group elements. In other words, it is central and hence trivial in the adjoint group. For example, if $G = SU(2)$ this means that $\Phi_{\gamma(0)}$ is ± 1.

Let $o \in M$ be any point and consider those gauge transformations which are the identity at o. These gauge transformations form a normal subgroup $\mathscr{G}_o \subset \mathscr{G}$, whose quotient $\mathscr{G}/\mathscr{G}_o$ is isomorphic to G. It is not hard to see, again using the gauge covariance of parallel transport, that $\mathscr{G}_o$ acts freely on the space $\mathscr{A}$ of connections. Indeed, suppose that $\Phi \in \mathscr{G}_o$ leaves invariant a connection H. Then again $\Pi_\gamma = \Pi_\gamma^\Phi$ for all curves γ starting at $\gamma(0) = o$, whence using that $\Phi_{\gamma(0)} = \mathrm{id}$,

$$\Phi_{\gamma(1)} \circ \Pi_\gamma = \Pi_\gamma \implies \Phi_{\gamma(1)} = \mathrm{id} \ .$$

Since γ is an arbitrary curve, $\Phi = \mathrm{id}$ everywhere.

In summary, the group of gauge transformations $\mathscr{G}$ acts (almost) freely on the space of irreducible connections and the group of restricted gauge transformations $\mathscr{G}_o$ acts freely on the space of connections. This prompts the following definitions. We will work with definiteness with self-dual connections, but similar definitions apply for anti-self-dual connections.

Let $\mathscr{A}^+ \subset \mathscr{A}$ denote the space of self-dual connections and let $\mathscr{A}_o \subset \mathscr{A}$ denote the space of irreducible connections. Their intersection $\mathscr{A}_o^+ = \mathscr{A}^+ \cap \mathscr{A}_o$ is then the space of irreducible self-dual connections. Both irreducibility and self-duality are gauge invariant conditions, whence $\mathscr{G}$ preserves $\mathscr{A}_o^+$. The quotient $\mathscr{M} = \mathscr{A}_o^+/\mathscr{G}$ is called the **moduli space of instantons**. Alternatively we can consider the quotient $\widetilde{\mathscr{M}} = \mathscr{A}^+/\mathscr{G}_o$, which is called the **moduli space of *framed* instantons**. $\widetilde{\mathscr{M}}$ is fibred over $\mathscr{M}$ with fibres G. Under suitable conditions, both $\mathscr{M}$ and $\widetilde{\mathscr{M}}$ are finite-dimensional manifolds; although it is $\widetilde{\mathscr{M}}$ which has the more interesting geometry, as we will see.

5.2 The deformation complex

Let ω be a self-dual connection. The tangent space $T_\omega\mathscr{A}^+$ is the subspace of $T_\omega\mathscr{A}$ defined by the linearised self-duality equations. In turn, $T_\omega\mathscr{A}^+$ has a subspace consisting of tangent directions to the orbit $\mathscr{G}\cdot\omega$ of ω under gauge transformations. If ω is also irreducible, then the orthogonal complement of $T_\omega(\mathscr{G}\cdot\omega)$ (with respect to a suitable inner product) inside $T_\omega\mathscr{A}^+$ is isomorphic to the tangent space $T_\omega\mathscr{M}$ to the moduli space of instantons at ω. In this section we will set up the calculation of the dimension of $T_\omega\mathscr{M}$. The details can be found in the paper [AHS78].

Since $\mathscr{A}$ is an affine space modelled on $\Omega^1(M;\operatorname{ad}P)$, the tangent space $T_\omega\mathscr{A}$ is naturally isomorphic to $\Omega^1(M;\operatorname{ad}P)$. Consider a curve $\omega_t := \omega + t\tau$ in $\mathscr{A}$ passing through ω, where $\tau \in \Omega^1(M;\operatorname{ad}P)$. Such a straight line will not generally correspond to a self-dual connection for any $t > 0$, but we can demand that it be so up to first order in t; that is, we can demand that its velocity be tangent to $\mathscr{A}^+$. The curvature Ω_t of ω_t is given by

$$\Omega_t = \Omega + t d^\omega\tau + \tfrac{1}{2}t^2[\tau,\tau] \ ,$$

where Ω is the curvature of ω. This is self-dual up to first order if and only if $d^\omega\tau$ is self-dual.

In order to recognise those directions tangent to the gauge orbit, we need to discuss infinitesimal gauge transformations. We will consider a curve Φ_t in $\mathscr{G}$ passing by $\Phi_0 = \mathrm{id}$. The derivative with respect to t at the identity gives rise to an element of the tangent space to $\mathscr{G}$ at the identity, which we may identify with Lie algebra $\mathfrak{G} = C^\infty(M;\operatorname{ad}P)$ of the group of gauge transformations. We can define a map $\exp : \mathfrak{G} \to \mathscr{G}$ by fibrewise application of the exponential map.[1] We may describe this locally relative to a trivialisation. If $\Theta \in C^\infty(M;\operatorname{ad}P)$ is described by a family of local functions $\{\theta_\alpha : U_\alpha \to \mathfrak{g}\}$, then $\Phi_t := \exp(t\Theta) \in \mathscr{G}$ is described by the family of local functions $\{\exp(t\theta_\alpha) : U_\alpha \to G\}$. The connection ω is similarly described by a family of local gauge fields $\{A_\alpha \in \Omega^1(U_\alpha;\mathfrak{g})\}$, on which the gauge transformation Φ_t has the following effect

$$A_\alpha^{\Phi_t} = \exp(t\theta_\alpha)A_\alpha\exp(-t\theta_\alpha) - d\exp(t\theta_\alpha)\exp(-t\theta_\alpha) \ ,$$

where we have assumed a matrix group for simplicity. Differentiating with respect to t and setting $t = 0$ we recover the form of an infinitesimal gauge transformation:

$$\left.\frac{d}{dt}A_\alpha^{\Phi_t}\right|_{t=0} = \theta_\alpha A_\alpha - A_\alpha\theta_\alpha - d\theta_\alpha = -d_A\theta_\alpha \ ,$$

[1] Although contrary to what happens in finite-dimensional Lie groups, there may be gauge transformations which are arbitrarily close to the identity which are not in the image of the exponential map.

which are (up to a sign) the local representatives of $d_A\Theta \in \Omega^1(M;\operatorname{ad}P)$.

The preceding discussion can be summarised in terms of the following sequence of linear maps:

$$0 \longrightarrow \Omega^0(M;\operatorname{ad}P) \xrightarrow{d_A} \Omega^1(M;\operatorname{ad}P) \xrightarrow{d_A^-} \Omega^2_-(M;\operatorname{ad}P) \longrightarrow 0\,, \tag{25}$$

where $d_A^-\tau := (d_A\tau)^-$ stands for the anti-self-dual part of $d_A\tau$. Notice that because ω is a self-dual connection, the composition $d_A^- \circ d_A = F_A^- = 0$, so the above is a complex, called the **deformation complex**. This means that the image of the first map is contained in the kernel of the second, but it need not necessarily be all of it.

In fact, a tangent vector $\tau \in T_\omega\mathscr{A}$ is tangent to $\mathscr{A}^+$ if and only if it is in the kernel of the second map, whereas it is tangent to the gauge orbit if and only if it is in the image of the first. In other words, if ω is irreducible,

$$T_\omega\mathscr{M} \cong \frac{\ker d_A^- : \Omega^1(M;\operatorname{ad}P) \to \Omega^2_-(M;\operatorname{ad}P)}{\operatorname{im} d_A : \Omega^0(M;\operatorname{ad}P) \to \Omega^1(M;\operatorname{ad}P)}\,,$$

which is the first cohomology group H^1 of the deformation complex. For M compact, the deformation complex is elliptic and hence has finite index

$$\operatorname{index} = \dim H^0 - \dim H^1 + \dim H^2\,.$$

In other words,

$$\dim\mathscr{M} = \dim H^1 = -\operatorname{index} + \dim H^0 + \dim H^2\,.$$

The index can be computed in principle by the Atiyah–Singer index theorem, but the index will not be enough to compute the dimension of the moduli space unless we have some control over H^0 and H^2.

For an irreducible connection, $\dim H^0 = 0$. Indeed, $H^0 = \ker d_A : \Omega^0(M;\operatorname{ad}P) \to \Omega^1(M;\operatorname{ad}P)$, hence $\dim H^0 \neq 0$ if and only if there is some $\Theta \in \Omega^0(M;\operatorname{ad}P)$ such that $d_A\Theta = 0$. But such a Θ is invariant under parallel transport and hence commutes with the holonomy group of the connection. In particular, it belongs to the centraliser of its Lie algebra. If the connection is irreducible, this Lie algebra is all of $\mathfrak{g}$, which is assumed to be semisimple and hence without centre.

In fact, for the 4-sphere and in the case of $G = SU(2)$, there can be no self-dual reducible connections with nonzero instanton number. The reason is that if the holonomy is a proper subgroup of SU(2), it must have the homotopy type of a circle subgroup and for the 4-sphere, there can be no nontrivial circle bundles, since G-bundles over the 4-sphere are classified by the third homotopy group $\pi_3(G)$, which vanishes for $G = S^1$.

For the 4-sphere (more generally any self-dual manifold with positive scalar curvature), a Weitzenböck argument shows that $H^2 = 0$. This argument runs as follows. H^2 is the cokernel of d_A^-, which is the kernel of the (formal) adjoint $(d_A^-)^*$. One calculates the corresponding laplacian operator $d_A(d_A^-)^*$ and shows that this is a positive operator and hence that it has no kernel.

Therefore on the 4-sphere, $\dim\mathscr{M}$ coincides with the index of the deformation complex, which can be computed using the Atiyah–Singer index theorem. For gauge group $G = SU(2)$ and instanton number k (positive), one obtains $\dim M = 8k - 3$. In particular, for $k = 1$ we obtain a five-dimensional moduli space. These are precisely the five parameters in the BPST solution: the scale and the centre of the instanton.

(I realise that this section is missing many details. I hope to remedy this eventually by a couple of lectures on a supersymmetric proof of the index theorem.)

References

[AHS78] MF Atiyah, NJ Hitchin, and IM Singer, *Self-duality in four-dimensional riemannian geometry*, Proc. R. Soc. London A. **362** (1978), 425–461.

[KN63] S Kobayashi and K Nomizu, *Foundations of differential geometry*, Wiley, 1963.