Gauge theory at Bocconi
Lecture 2: Introduction to Yang-Mills theory

16. April 2026

1 Recap of the last lecture

Last week we have seen that electromagnetism can be reformulated as

$$dF = 0$$
$$d \star F = 0$$

for a two-form $F \in \Omega^2(M, i\mathbb{R})$ on a Riemannian manifold.

Via two methods, we saw that potentials $A \in \Omega^1(M, i\mathbb{R})$ are actually the underlying rather than $F = dA$. In this discussion, we also saw that there is a gauge group $\mathcal{G} = \mathscr{C}^\infty(M, S^1)$ acting via

$$g.A = A - dgg^{-1}.$$

Today we will upgrade this to the non-abelian setting.

2 Formal setup

2.1 Lie groups and their metrics

As we need some Riemannian geometry of compact Lie groups let us recall some basic facts. A **Lie group** G is a group and a manifold at the same time such that the multiplication and the inversion are smooth.

Popular examples are $U_1 = \{z \in \mathbb{C} \mid |z| = 1\}$ or

$$\mathrm{SU}_2 = \left\{ \begin{bmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{bmatrix} \,\middle|\, |\alpha|^2 + |\beta|^2 = 1 \right\} \cong S^3$$

The **Lie algebra** $\mathfrak{g} = \mathrm{Lie}(G)$ associated to a Lie group G can be defined as T_eG where e denotes the identity element of G.
For SU_2 its Lie algebra consists of traceless anti-Hermitian matrices

$$\mathfrak{su}_2 = \{X \in \mathrm{Mat}_2(\mathbb{C}) \mid X + {}^t\bar{X} = 0, \mathrm{tr}(X) = 0\}.$$

Proposition 2.1. *On a compact Lie group G there is a biinvariant metric h, that is, a metric h that is invariant under left and right multiplication on G.*

On $\mathfrak{su}_2$ the biinvariant metric is induced by $h(X, Y) = -\mathrm{tr}(XY)$ for $X, Y \in \mathfrak{su}_2$.

2.2 Principal G-bundles

The basic backbone of gauge theory is the theory of principal bundles. The idea of principal bundles is not complicated. In the definition of a vector bundle, we require the fibres to be Lie groups G with a right G-action (more precisely G-torsors) instead of vector spaces. The next definition formalises this intuition.

Definition 2.2. A **principal G-bundle** $\pi : P \to M$ is a smooth map $\pi : P \to M$ between the smooth manifolds P, M satisfying the following conditions:

(a) There is a right action of G on P such that it is **free**[1] and **fibre transitive**[2]

(b) around each $x \in M$ there is an open set U with the following property: There is a diffeomorphism $\psi_U : P_U = \pi^{-1}(U) \to U \times G$ such that

  (i) $\pi|_{P_U} = \mathrm{pr}_1 \circ \psi_U$

  (ii) $\psi(p.g) = \psi_U(p).g$ where the right action on $U \times G$ is given by $(x, h).g = (x, hg)$ for $g \in G$.

Example 2.3. (a) A trivial G-principal bundle $P = M \times G \to M$ with G-action $(x, h).g = (x, hg)$ is the most basic example.

(b) A slightly more involved example is the Hopf fibration $\pi : S^3 \to S^2$. Consider the the mapping

$$\mathbb{C}^2 \supseteq S^3 \to S^2 \cong \mathbb{CP}^1 \cong \mathbb{C} \cup \{\infty\}$$

where the smooth mapping is given by

$$\pi(z, w) = (z : w)$$

The right multiplication by the Lie group $S^1 \cong U_1$ is given by

$$(z, w).a = (za, wa)$$

The usual two sections used for the Hopf fibration are

$$s_1 : \mathbb{CP}^1 \setminus \{\infty\} \to S^3 \ ; \ z \mapsto \tfrac{1}{\sqrt{1+|z|^2}} \begin{pmatrix} z \\ 1 \end{pmatrix}$$

$$s_2 : \mathbb{CP}^1 \setminus \{0\} \to S^3 \ ; \ w \mapsto \tfrac{1}{\sqrt{1+|w|^2}} \begin{pmatrix} 1 \\ w \end{pmatrix}$$

for $w = \frac{1}{z}$. Since sections and local trivializations correspond to each other, we can also relate the section by the multiplication by the cocycle $g_{12} : \mathbb{C}^* \to S^1$ given by $g_{12}(z) = \frac{z}{|z|}$

$$s_1(z) = s_2\left(\frac{1}{z}\right).g_{12}(z)$$

for $z \in \mathbb{C}^*$.

[1] $p.g = p$ for every p then $g = e \in G$

[2] for every $x \in M$ and every $p, q \in \pi^{-1}(x) := P_x$ there is a $g \in G$ such that $p.g = q$

Associated vector bundles Let $R : G \to \mathrm{GL}(V)$ be a representation. Then there is a way of associating a vector bundle $R(P) \to M$ to a principal G-bundle $P \to M$.

The idea is pretty simple. Use the representation R to replace a fibre $P_x \cong G$ by the vector space V. More precisely this is done via the definition

$$R(P) := (P \times V)/\sim$$

where $(p, v) \sim (p.g, R(g^{-1})v)$ for every $g \in G$.

Definition 2.4. The **adjoint bundle** $\mathrm{Ad}(P)$ is the vector bundle associated with the adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$.

2.3 Connections, curvature and covariant derivative

Let $\pi : P \to M$ be a principal G-bundle. For every $\xi \in \mathfrak{g} = T_eG$ we can define a vector field $\xi^\sharp(p) = \frac{d}{dt}\Big|_{t=0} (p.\exp(t\xi))$ for $p \in P$. By definition

$$D\pi(\xi^\sharp) = \frac{d}{dt}\Big|_{t=0} \pi(p.\exp(t\xi)) = \frac{d}{dt}\Big|_{t=0} \pi(p) = 0$$

and the vertical bundle $VP := \ker D\pi \subseteq TP$ can be shown to satisfy $VP \cong P \times \mathfrak{g}$.

Definition 2.5. A **connection form** A on a principal G-bundle P is an element of $\Omega^1(P, \mathfrak{g})$ such that

(a) $R_g^* A(X) = \mathrm{Ad}_{g^{-1}}(A(X))$ for every vector field X on $P$and

(b) $A(\xi^\sharp) = \xi$ for every $\xi \in \mathfrak{g}$.

Remark. (a) The two conditions of a connection form are constructed such that $H^AP = \ker A$ becomes a complementary bundle to VP, i.e. $TP = VP \oplus H^AP$.

(b) Morover, the set of all connections $\mathcal{A}(P)$ is an affine space.

Example 2.6. Consider $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. A connection A in this case has to satisfy $A(a\partial_y) = a$. Thus $A = dy + f(x)dx$

Definition 2.7. The **curvature** F_A of a connection A is defined by $F_A := dA + \frac{1}{2}[A, A]$ where the bracket is defined by $[\omega \otimes X, \sigma \otimes Y] = \omega \wedge \sigma \otimes [X, Y]$ and continued linearly.

Definition 2.8. The **covariant derivative** $d_A : \Gamma(\mathrm{Ad}(P)) \to \Omega^1(\mathrm{Ad}(P))$ is defined as $d_A\sigma = d\sigma + [A, \sigma]$[3].

Lastly, let us define the **gauge group** as

$$\mathcal{G} = \{\phi \in \mathrm{Diff}(P) \mid \phi(p.g) = \phi(p).g \text{ for all } g \in G,\ \pi \circ \phi = \pi\}$$

with composition as a group multiplication. The gauge action of $\mathcal{G}$ on a connection A is given by $\phi.A = (\phi^{-1})^*A$ locally given by

$$\phi.A = \phi A \phi^{-1} - d\phi\phi^{-1}$$

interpreted correctly if G is not a matrix Lie group. Usually we will denote an element of $\mathcal{G}$ by g.

Proposition 2.9. *The curvature transforms under gauge transformation as* $F_{g.A} = gF_Ag^{-1}$.

[3]Strictly speaking we are not giving the correct definition here as we are suppressing identifications.

2.4 A Mathematics ⇔ Physics dictionary

As the naming convention in Mathematics and Physics are slightly different, we collect them in the next table.

Object	**Math**	**Physics**
$P \to M$	principal G-bundle	gauge bundle
$A \in \Omega^1(P, \mathfrak{g})$	connection form	gauge potential/field
F_A	curvature	field strength
G	structure group	gauge group
$\mathfrak{g}$	Lie algebra of G	infinitesimal gauge algebra
$\mathcal{G}$	gauge group	gauge transformations

3 Yang Mills theory

In 1954 C.N. Yang and R.L. Mills published their ground-breaking paper "Conservation of isotopic Spin and isotopic gauge invariance" the foundation for modern gauge theory were considered. Motivated by quantum mechanics they proposed a covariant derivative that takes values in $\mathfrak{su}_2 = \mathrm{sp}_1$ instead of $\mathfrak{u}_1 \cong i\mathbb{R}$. As their motivation was to understand processes involving isospin this was a natural choice.

> In the present paper we wish to explore the possibility of requiring all interactions to be invariant under independent rotations of the isotopic spin at all spacetime points (YM54, p.192).

Their approach is clearly motivated by the abelian gauge theory of Weyl and at the end of the day they their upgrade is mathematically just replacing U_1 by SU_2.
Remark. The theory of Yang and Mills ultimately lead to the development of particle physics and even to the standard model. This is a gauge theory based on the group $\mathrm{SU}_3 \times \mathrm{SU}_2 \times \mathrm{U}_1$.

3.1 The Yang-Mills functional

The **Yang-Mills functional** $\mathrm{YM} : \mathcal{A}(P) \to \mathbb{R}$ over $P \to (M, g)$ is

$$\mathrm{YM}(A) = \frac{1}{2} \int_M \|F_A\|^2 \, d\mathrm{Vol}_g$$

and for $G = \mathrm{U}_1$ this is just the energy associated to (Euclidean) electromagnetism in a vacuum. Here, the norm $\|F_A\|^2$ uses g and a choice of biinvariant inner product on G.

Proposition 3.1. *Let $P \to (M, g)$ a principal G-bundle.*

(a) YM *is conformally invariant iff* $\dim M = 4$.

(b) $YM(g.A) = YM(A)$ for every $g \in \mathcal{G}$

(c) The Euler-Lagrange equation of YM *is $d_A^* F_A = 0$*

Proof. We will just perform the crucial computation for (c) and leave the remaining part for the reader.

The computation for the Euler-Lagrange equation is

$$\begin{aligned} F_{A+ta} &= d(A + ta) + \frac{1}{2}[A + ta, A + ta] \\ &= dA + \frac{1}{2}[A, A] + t(da + [A, a]) + \frac{t^2}{2}[a, a] \\ &= F_A + t d_A a + \frac{t^2}{2}[a, a]. \end{aligned}$$

□

Remark. (a) Partially because of (a) Yang-Mills theory can be seen as an analogue of holomorphic/harmonic maps from a surface into a Riemannian manifold.

(b) Part (c) really shows that for $G = \mathrm{U}_1$ we obtain $d \star F_A = 0$ and $dF_A = 0$ (Bianchi identity) and hence the equations of vacuum electromagnetism are truly a part of Yang-Mills theory.

(c) If A is a solution to $d_A \star F_A = 0$ then so is $g.A$ for every $g \in \mathcal{G}$. This shows that the Yang-Mills equations are fundamentally non-elliptic. Only after a gauge fixing condition like $d_A^* A = 0$ would we re-obtain an elliptic nature.

3.2 Dimension four phenomena

> There is something very special about this identity $4 = 2 + 2$ [...] So you can say that this is kind of the proof of Donaldson's theorem.
> - Misha Gromov (Youtube)

Let us explain what Gromov means with this remark. Let us take for simplicities sake $M = \mathbb{R}^4$ (note that we could have taken any Riemannian four manifold) be the Euclidean space equipped with the coordinates $(x^1, \dots, x^4)$. Then the Hodge star on two forms $\star : \Omega^2(M) \to \Omega^2(M)$ acts via

$$\begin{aligned} \star(dx^1 \wedge dx^2) &= dx^3 \wedge dx^4 \\ \star(dx^1 \wedge dx^3) &= -dx^2 \wedge dx^4 \\ \star(dx^1 \wedge dx^4) &= dx^2 \wedge dx^3 \end{aligned}$$

from which we see that $\star^2 = 1$ and hence there is an orthogonal decomposition

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

with $\Omega^2_\pm = \{\omega \in \Omega^2(M) \mid \star\omega = \pm\omega\}$. $+$ refers to **self-dual (SD)** forms and $-$ to **anti-self-dual (ASD)** forms.

Generally on a Riemannian manifold we have

$$\begin{aligned} \frac{1}{2}\|F_A \pm \star F_A\|^2 \, d\mathrm{Vol}_g &= (\|F_A\|^2 \pm \langle F_A, \star F_A\rangle) d\mathrm{Vol}_g \\ &= \|F_A\|^2 d\mathrm{Vol}_g \pm \langle F_A \wedge F_A\rangle \end{aligned}$$

and hence

$$\mathrm{YM}(A) = \frac{1}{4}\int_M \|F_A \pm \star F_A\|^2 \, d\mathrm{Vol}_g \mp \frac{1}{2}\int_M \langle F_A \wedge F_A\rangle$$

and by Chern-Weil theory

$$\int_M \langle F_A \wedge F_A\rangle$$

is a topological term representing a characteristic class which is fixed by the choice of bundle.

Proposition 3.2. *Let $P \to M$ be fixed. A connection A on P satisfying $\star F_A = \pm F_A$ is an absolute minimiser of* YM *in its topological class.*

The only remaining part is noticing that for $\star F_A = \pm F_A$ the Yang Mills equations are $d_A \star F_A = d_A F_A = 0$ by the Bianchi identity.

3.3 BPST instantons

The BPST instanton family on $\mathbb{R}^4$ is an explicit solution family obtained by Belavin-Polyakov-Schwarz-Tyupkin in 1975. The construction idea is relatively simple. Identify $\mathbb{R}^4$ with the quaternions $\mathbb{H}$, i.e. the skew field constructed via $\mathbb{R}\langle 1, i, j, k\rangle$ such that the relations $i^2 = j^2 = k^2 = -1$ and $ij = k$ hold.
The identification is via

$$(x^1, x^2, x^3, x^4) \mapsto q = x^1 + i\, x^2 + j\, x^3 + k\, x^4$$

and note that the conjugation is defined as

$$\bar{q} = x^1 - i\, x^2 - j\, x^3 - k\, x^4$$

Via a symmetric Ansatz the SD equation $\star F_A = F_A$ on $\mathbb{H}$ reduce to an ODE and the resulting connection is

$$A^{BPST}_{\lambda,x} = \frac{\mathrm{Im}(d(q-x)(\overline{q-x}))}{\lambda^2 + |q-x|^2}$$

with the parameters $\lambda > 0, x \in \mathbb{H}$

Proposition 3.3. $F_A = \frac{\lambda^2}{(\lambda^2+r^2)^2}\mathrm{Im}(dq \wedge d\bar{q})$ *and* $\mathrm{YM}(A) = \frac{1}{2}\int_{R^4} |F_A|^2 = 4\pi^2$

An interesting observation at this stage is that for $\lambda \to 0$ a problem seems to happen. The curvature becomes flat on $\mathbb{R}^4 \setminus \{x\}$ and the energy of $4\pi^2$ seems to be missing. This is actually a prime example of bubbling in Yang-Mills theory as

$$\frac{1}{2}|F_A|^2\, d\mathrm{Vol} \rightharpoonup 0\, d\mathrm{Vol} + 4\pi^2\, \delta_x$$

In this manner Yang-Mills theory is a codimension four analogue of the codimension two bubbling known for harmonic maps.
For a full bubbling picture, we would need more ingredients like ε-regularity, gauge fixing and monotonicity formulas.

3.4 Dimensional reduction

The ASD equation $F_A = -\star F_A$ on $\mathbb{R}^4$ for a connection

$$A = A_1(x^1, \dots, x^4)\, dx^1 + \dots + A_4(x^1, \dots, x^4)\, dx^4$$

is

$$\begin{aligned} F_{12} &= -F_{34} \\ F_{13} &= F_{24} \\ F_{23} &= -F_{14} \end{aligned}$$

where $F_{ij} = [\nabla_i, \nabla_j] = \partial_i A_j - \partial_j A_i + [A_i, A_j]$ for $\nabla_i = \partial_i + A_i$

Reduction to $3D$ Now, we assume that A is independent of x^4 variable, i.e. invariant under the action of $\mathbb{R}$ by

$$\lambda.(x^1, x^2, x^3, x^4) = (x^1, x^2, x^3, x^4 + \lambda)$$

Then $A = A_1(x^1, x^2, x^3)\, dx^1 + \dots + A_3(x^1, x^2, x^3)\, dx^3 + \phi(x^1, x^2, x^3)\, dx^4$ and the equations imply

$$\begin{aligned} F_{12} &= -\nabla_3 \phi \\ F_{13} &= \nabla_2 \phi \\ F_{23} &= -\nabla_1 \phi \end{aligned}$$

On a three manifold (M, g) this becomes $F_A = -\star d_A \phi$ and is called the **Bogomolny equation** up to a harmless sign.

Moduli spaces of monopoles are especially rich and often also hyperkähler.

Reduction to $2D$ Under an $\mathbb{R}^2$-invariance, the connections becomes

$$A = A_1(x^1, x^2)\, dx^1 + A_2(x^1, x^2)\, dx^2 + \phi_1(x^1, x^2)\, dx^3 + \phi_2(x^1, x^2)\, dx^4$$

and the equations reduce to

$$\begin{aligned} F_{12} &= -[\phi_1, \phi_2] \\ \nabla_1 \phi_1 &= \nabla_2 \phi_2 \\ \nabla_2 \phi_1 &= -\nabla_1 \phi_2 \end{aligned}$$

which are referred to as the **Hitchin's equation**. By introducing a complex structure $z = x^1 + ix^2$ and $\Phi = (\phi_1 - i\phi_2)dz$ we obtain a system of equations

$$\begin{aligned} F_A + [\Phi, \Phi^*] &= 0 \\ \bar{\partial}_A \Phi &= 0 \end{aligned}$$

These equations are inherently related to interesting moduli spaces of holomorphic vector bundles and is a very active field of research. Here some keywords are Higgs bundles, character varieties and also higher Teichmüller theory.

Reduction to $1D$ Under an $\mathbb{R}^3$-invariance, the ASD equations reduce to an ODE system

$$\partial_t T_i + [T_0, T_i] + [T_j, T_k] = 0$$

for $T_0, T_1, T_2, T_3 : \mathbb{R} \to \mathfrak{g}$ and $(123) = (ijk)$ cyclically permuted. These are Nahm's equations and form again important moduli spaces. For instance, the standard construction of the hyperkähler metric on the cotangent bundle T^*G^c of a complexified compact Lie group G^c uses the Nahm's equation in an essential way.