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 instantos 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

$$S_{YM} \geq |c| , \tag{18}$$

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 $\dim M > 4$.

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,

$$\frac{d\Theta_t}{dt} = \tfrac{1}{2} d \operatorname{Tr}(\tau \wedge F_t) ,$$

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

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

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 $\operatorname{ad} P$:

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

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 $\mathfrak{su}(2) = \mathfrak{sp}(1) = \operatorname{Im}\mathbb{H}$.

One can show that $p_1(\operatorname{ad}P)[M]$ 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

$$S_{YM} \geq 4\pi^2 |k| \,, \tag{19}$$

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 $\pi^*\Theta \in \Omega^4(P)$ is exact:

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

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.

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 $\mathrm{Re}\,\boldsymbol{x} = x_1$ and $\mathrm{Im}\,\boldsymbol{x} = x_2 i + x_3 j + x_4 k$ 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: $\boldsymbol{x} \cdot \boldsymbol{y} = \mathrm{Re}(\boldsymbol{x}\overline{\boldsymbol{y}})$. We will denote the corresponding norm by $|\boldsymbol{x}|^2 = \mathrm{Re}(\boldsymbol{x}\overline{\boldsymbol{x}})$.

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

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

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

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

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

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 $\mathrm{Im}\mathbb{H}$-valued self-dual 2-form, and hence so is the field-strength F. Therefore the gauge field A defines an $\mathrm{SU}(2)$ instanton on $\mathbb{R}^4$. To determine its instanton number, we need only integrate

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

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:

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

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

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

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.)