Instantons and the ADHM Construction
Lecture 1

Alex Atanasov

November 15, 2016

Abstract

We explore connections on $\mathbb{R}^4$ and the Yang-Mills equations arising from minimizing a quantity known as action. We study solutions to these equations possessing nonzero action, known as instantons, and demonstrate a method to construct all instantons on $\mathbb{R}^4$ with dimension n and topological charge k. This is the ADHM construction of Atiyah et al.

1 Motivation

In this course we have seen examples of geometrization: the association of geometric structure to an underlying algebraic structure. We'e seen that categorification of $\mathfrak{sl}_q(2,\mathbb{C})$ gives rise to cohomology rings of Grassmanians. In a similar vein, more general affine Lie algebras $\hat{\mathfrak{g}}$ give rise to geometric spaces that can be understood as moduli spaces of instantons on asymptotically-locally-euclidean (ALE) spaces $\mathbb{C}^2/\Gamma$, in one-to-one correspondence with the extended Affine Dynkin diagrams.

We give an introduction to instanton construction first in the simple case of $\mathbb{C}^2 \cong \mathbb{R}^4$. Even in this simple case, we will see how this theory is deeply connected to affine Lie algebras, Hilbert schemes, and quiver varieties.

2 Yang Mills Instantons on $\mathbb{R}^4$

2.1 Connection and Curvature Forms

Definition 2.1. A **Hermitian vector bundle** $\pi : E \to M$ over a base space M is a complex vector bundle over M equipped with a Hermitian inner product on each fiber.

Yang Mills theory on M concerns itself with the metric-compatible **connections** A on E.

Definition 2.2 (Connection on a Vector Bundle)**.** A connection A on a vector bundle $\pi : E \to M$ of rank n is a $\mathfrak{gl}(n)$-valued 1-form

For a Hermitian bundle, we restrict to $\mathfrak{u}(n)$, to work with only metric-compatible connections. Each such connection $A \in \mathcal{A}$ is a $\mathfrak{u}(n)$-valued 1-form acting on E by ρ.

Definition 2.3 (**Covariant Exterior Derivative**)**.** For a given connection $A \in \Omega^1(M, \mathfrak{u}(n))$, we obtain a corresponding differential operator on M:

$$\mathrm{d}_A := \mathrm{d} + \rho(A) \tag{1}$$

Observation 2.4. *In coordinate language, we can write:*

$$(\mathrm{d}_A)_\mu = \partial_\mu + \rho(A_\mu) \tag{2}$$

We can then define the **curvature** 2-form by having this derivative act on its own connection 1-form

Definition 2.5 (Curvature/Field-Strength 2-form)**.**

$$\begin{aligned} F := d_A A &= dA + A \wedge A \\ &= dA + \frac{1}{2}[A, A] \end{aligned} \tag{3}$$

Observation 2.6. *In coordinate language, we can write:*

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{4}$$

$$s.t.\ F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu \tag{5}$$

We conclude with an identity that can be checked by direct computation

Proposition 2.7 (Bianchi Identity)**.**

$$\mathrm{d}_A F = 0 \tag{6}$$

2.2 The Action

For our purposes, $M = \mathbb{R}^4$ will be the manifold in question. In particular $\mathbb{R}^4$ has Riemannian structure, so we are given the Hodge-star operator

$$\star : \Omega^k \to \Omega^{n-k}.$$

We define the **action**, from which we will obtain all information about the dynamics, by:

$$S_E[\mathcal{A}] = -\int_M \mathrm{Tr}(F \wedge \star F) \tag{7}$$

Proposition 2.8. $\mathrm{Tr}(F \wedge \star F)$ *is globally-defined and gauge invariant*

Proof. This follows directly from the cyclic properties of the trace, and the transformation laws on F making it transform under the adjoint representation. □

We want to find A so that $S_E[\mathcal{A}]$ is a minimum. To do this, we use standard calculus of variations. Consider a local perturbation $A + t\alpha$

$$\begin{aligned} F[A + t\alpha] &= \mathrm{d}(A + t\alpha) + A \wedge A + t[A, \alpha] + O(t^2) \\ &= F[A] + t(d\alpha + [A, \alpha]) + O(t^2) \\ &= F[A] + \mathrm{d}_A\alpha + O(t^2) \end{aligned} \tag{8}$$

so that to order t:

$$\begin{aligned} ||F[A + t\alpha]||^2 &= ||F[A + t\alpha]||^2 + 2t(F[A], \mathrm{d}_A\alpha) \\ &\Rightarrow (F[A], \mathrm{d}_A\alpha) = 0 \ \forall\alpha \end{aligned} \tag{9}$$

By taking adjoints, this gives:

$$\begin{aligned} &\Rightarrow \star\mathrm{d}_A \star F[A] = 0 \\ &\Rightarrow \mathrm{d}_A \star F = 0 \end{aligned} \tag{10}$$

This, together with the tautological Bianchi identity: $\mathrm{d}_A F = 0$ form the Yang-Mills equations. These equations are very difficult to solve in all but abelian gauges, where they become linear.

2.3 Instantons and Topological Charge

Proposition 2.9. *Let* $\dim M = 4$. *Then* $\int_M \mathrm{Tr}(F \wedge F)$ *is independent of changes in* A.

Proof. Following the same variational procedure will give us $\mathrm{d}_{\mathcal{A}}F$, which is zero always, independent of any condition on A. □

We define the **topological charge** k of the theory by

$$k := -\frac{1}{8\pi^2}\int_M \mathrm{Tr}(F \wedge F) \tag{11}$$

Proposition 2.10. *When* $M = S^4$, *we have that* k *is an integer.*

Proof. The proof lies in simple ideas from Chern classes and classifying bundles over S^4. It establishes a one-to-one correspondence between the global topology type of the bundle E over S^4 and the topological charge.

□

Now note that on $\mathbb{R}^4$, we have $\star\star = 1$. This means that $\star$ has eigenvalues ± 1 and so $\Omega^2(U, \mathfrak{g})$ splits as a direct sum of two orthogonal spaces:

$$\Omega^2(\mathbb{R}^2, \mathfrak{u}(n)) = \Omega^2_+ \oplus \Omega^2_- \tag{12}$$

called **self-dual** and **anti-self-dual** spaces respectively.

We can "symmetrize" any form to become a sum of a self-dual and an anti-self dual one. In particular, if we write:

$$F = F_+ + F_- \tag{13}$$

then we have

$$\begin{aligned} -8\pi^2 k &= \int_M \mathrm{Tr}[(F_+ + F_-) \wedge (F_+ + F_-)]\mathrm{dVol} \\ &= \int_M \mathrm{Tr}[(F_+) \wedge (F_+)]\mathrm{dVol} + \int_M \mathrm{Tr}[(F_-) \wedge (F_-)]\mathrm{dVol} \\ &= \int_M ||F_+||^2 \ \mathrm{dVol} - \int_M ||F_-||^2 \ \mathrm{dVol} \end{aligned} \tag{14}$$

Note that the absolute value of this gives:

$$8\pi^2 k \leq \int_M ||F||^2 = |S_A[F]| \tag{15}$$

Proposition 2.11. *The action is bounded below by this topological charge and is in fact* equal *to it exactly when one of $F_+ = 0$ or $F_- = 0$.*

We call a solution an **instanton** of the theory. Its action is equal to the topological charge, and in fact we call this the **instanton number** when appropriate. We are interested in the space of instantons modulo gauge equivalence

Definition 2.12. The **gauge group** $\mathcal{G}$ of all metric-compatible transformation on E, restricts to SU(n) at each point. Two connections A_1, A_2 are Gauge equivalent if they differ by an element in $\mathcal{G}$. We are interested in the space of connections modulo gauge.

Instantons on $\mathbb{R}^4$ must have that F is either self-dual or anti-self-dual. In the latter case:

$$\star F = - \star F \tag{16}$$

This equation is much simpler to solve than the equation of motion $\mathrm{d}_A \star F = 0$. The anti-self-duality (ASD) equations can be written out explicitly:

$$\begin{aligned} F_{12} + F_{34} &= 0 \\ F_{14} + F_{23} &= 0 \\ F_{13} + F_{42} &= 0 \end{aligned} \tag{17}$$

This can also be written in terms of commutators of the covariant derivatives. If we denote $(d_A)_\mu$ simply by D_μ then $F_{\mu\nu} = (d_A)_\mu (d_A)_\nu = [D_\mu, D_\nu]$.

$$\begin{aligned} [D_1, D_2] + [D_3, D_4] &= 0 \\ [D_1, D_4] + [D_2, D_3] &= 0 \\ [D_1, D_3] + [D_4, D_2] &= 0 \end{aligned} \tag{18}$$

Proposition 2.13. *There are no instantons on Minkowski space $\mathbb{R}^{3,1}$.*

Proof. $\star\star = -1$ on Minkowski space, so $\star$ has eigenvalues $\pm i$, meaning the duality equations would require $\star F = \pm i F$, but $F \in \Omega^2(\mathbb{R}^4, \mathfrak{u}(n))$ is a real object. □

Proposition 2.14. *For all connections on a given vector bundle E, the instanton number is an invariant.*

Proof. This follows since for instantons $S_{\mathcal{A}} = 8\pi \mathrm{k}$ is independent of the connection. □

Corollary 2.15. *There are no instantons when G is abelian.*

Proof. $F = dA \Rightarrow ||F|| = (\star dA, dA) = (\delta \star A, dA) = (\star A, d^2 A) = 0$ $\square$

We then have two invariants to note: n and k. We will be especially interested in the moduli space of all instantons for specific n and k (modulo gauge). From now on, we will focus specifically on anti-self-dual (ASD) instantons.

$$\mathcal{M}_{ASD}(n, k)$$

Self-dual instantons can be constructed in a straightforward one-to-one manner from the ASD instantons.

There is one subtlety: For k to be finite, we need F to vanish sufficiently quickly. This gives a bound for $|F| = |d_A A(x)| = O(|x|^{-4})$ for large x. This further gives a constraint on the gauge group $\mathcal{G}$ as $x \to \infty$ to have locally trivial structure. Instantons with this condition on their behaviour and gauge group are called **framed** instantons.

We say that in a neighborhood of infinity of S^4, the gauge group element must give a section of the bundle E that has a local trivialization $\Phi : E_\infty \to \mathbb{C}^n$. We denote the moduli space of framed instantons by

$$\mathcal{M}^{fr}_{ASD}(n, k)$$

3 The ADHM Construction

3.1 The Data

Let x_1, x_2, x_3, x_4 parameterize a $\mathbb{R}^4$, and write this as $\mathbb{C}^2$ using $z_1 = x_2 + ix_1, z_2 = x_4 + ix_3$. We can then write all the $(\mathrm{d}_{\mathcal{A}})_\mu$ (from now on just D_μ) . Moreover in terms of the complex coordinates, we get

$$\begin{aligned} \mathcal{D}_1 &= \frac{1}{2}(D_2 - iD_1) \\ \mathcal{D}_2 &= \frac{1}{2}(D_4 - iD_3) \end{aligned} \tag{19}$$

We can express anti-self duality of $F_{\mu\nu}$ in terms of these $\mathcal{D}_\mu$ through two equations:

$$\begin{aligned} [\mathcal{D}_1, \mathcal{D}_2] &= 0 \\ [\mathcal{D}_1, \mathcal{D}_1^\dagger] + [\mathcal{D}_2, \mathcal{D}_2^\dagger] &= 0 \end{aligned} \tag{20}$$

The idea behind ADHM is to convert these D_i to matrices B_i in a method akin to taking "Fourier transforms", and adding source terms depending on k.

Definition 3.1 (ADHM Data)**.** Let U be a 4-dimensional space with complex structure. An **ADHM System** on U is a set of linear data:

1. Vector spaces V, W over $\mathbb{C}$ of dimensions k, n respectively.

2. Complex $k \times k$ matrices B_1, B_2, a $k \times n$ matrix I, and an $n \times k$ matrix J.

We can see this diagrammatically by the following doubled, framed quiver:

$$\begin{tikzcd} W \arrow[r, shift right, "i"'] & V \arrow[l, shift right, "j"'] \arrow[loop above, "B_1"] \arrow[loop below, "B_2"] \end{tikzcd}$$

Definition 3.2 (ADHM System)**.** A set of ADHM Data is an ADHM system if it satisfies the following contraints:

1. The ADHM equations:

$$\begin{aligned} [B_1, B_2] + IJ = 0 \\ [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J = 0 \end{aligned} \tag{21}$$

These quantities are called real and complex moment maps, respectively.

2. For any two $x, y \in \mathbb{C}^2$ with $x = (z_1, z_2), y = (w_1, w_2)$, the map:

$$\alpha_{x,y} = \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \tag{22}$$

is injective from V to $W \oplus (V \otimes U)$ while

$$\beta_{x,y} = \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \end{pmatrix} \tag{23}$$

is surjective from $W \oplus (V \otimes \mathbb{C}^2)$ to V.

It's worth noting that $W \oplus (V \otimes \mathbb{C}^2) \cong W \oplus V \oplus V$.

Lemma 3.3. *If (B_1, B_2, I, J) satisfy the above conditions, then for $g \in \mathrm{U(k)}, h \in \mathrm{SU}(n)$, we get*

$$(gB_1g^{-1}, gB_2g^{-1}, gIh^{-1}, hJg^{-1}) \tag{24}$$

also satisfy the above conditions.

Thus we care about solutions to these equations modulo $\mathrm{GL}(V)$ and $\mathrm{GL}(W)$ as above.

Proof. It's a quick check through direct algebra that the equations are preserved. □

Proposition 3.4. *The ADHM equations are satisfied iff*

$$V \xrightarrow{\alpha_{x,y}} W \oplus (V \otimes \mathbb{C}^2) \xrightarrow{\beta_{x,y}} V \tag{25}$$

is a complex

Proof. We need both $\beta\alpha = 0$ as well as surjectivity of β and injectivity of α. The actual equation $\beta\alpha = 0$ reduces exactly to a quadratic polynomial in the w_1, w_2 with the two ASD equations emerging as coefficients. □

Now because we have Hermitian structure on each of W, V, and U, we have hermitian structure on the space we are interested. We can thus define adjoints $\alpha^\dagger, \beta^\dagger$. In particular the Hermitian structure gives us canonical projection operators P_β onto $\ker\beta$ and P_α $(\mathrm{im}\ \alpha)^\perp = \ker\alpha$ so that $P_x = P_{\beta,x}P_{\alpha,x}$ is then a projection onto $\mathrm{im}\ \alpha^\perp \cap \ker\beta \cong \ker\beta/\mathrm{im}\ \alpha$.

The above proposition also implies

$$\Delta^\dagger_{x,y} := \begin{pmatrix} \beta_{x,y} \\ \alpha^\dagger_{x,y} \end{pmatrix} : W \oplus (V \otimes \mathbb{C}^2) \to V \times V \tag{26}$$

is a surjection. Explicitly:

$$\Delta^\dagger = \begin{pmatrix} w_2 I + w_1 J^\dagger & w_2 B_2 - w_1 B_1^\dagger + z_2 & w_2 B_1 + w_1 B_2^\dagger + z_1 \\ -\bar{w}_1 I + \bar{w}_2 J^\dagger & -\bar{w}_1 B_2 - \bar{w}_2 B_1^\dagger - \bar{z}_1 & -\bar{w}_1 B_1 + \bar{w}_2 B_2 + \bar{z}_2 \end{pmatrix} \tag{27}$$

Moreover, there is an adjoint operator to $\Delta^\dagger$ (dropping subscripts):

$$\Delta := \begin{pmatrix} \beta^\dagger & \alpha \end{pmatrix} = \begin{pmatrix} \bar{w}_2 I^\dagger + \bar{w}_1 J & w_2 J - w_1 I^\dagger \\ \bar{w}_2 B_2^\dagger - \bar{w}_1 B_1 + \bar{z}_2 & -w_2 B_1 - w_1 B_2^\dagger - z_1 \\ \bar{w}_2 B_1^\dagger + \bar{w}_1 B_2 + \bar{z}_1 & w_2 B_2 - w_1 B_1^\dagger + z_2 \end{pmatrix} \tag{28}$$

More compactly, if we write

$$a = \begin{pmatrix} I^\dagger & J \\ B_2^\dagger & -B_1 \\ B_1^\dagger & B_2 \end{pmatrix}, \; b = \begin{pmatrix} 0 & 0 \\ I_k & 0 \\ 0 & I_k \end{pmatrix} \tag{29}$$

then

$$\Delta = aw + bz \tag{30}$$

when we write w and z as quaternions in this space by associating to a complex pair $(q_1, q_2) = q \in \mathbb{C}^2$ the quaternionic operator:

$$q \leftrightarrow \begin{pmatrix} \bar{q}_2 & -q_1 \\ \bar{q}_1 & q_2 \end{pmatrix} \tag{31}$$

for any $q_1, q_2 \in \mathbb{C}$. This structure is compatible with the operator R:

Proposition 3.5. $\Delta^\dagger_{xq,yq} = \bar{q}\Delta^\dagger_{x,y}$

Proof. We have that

$$\begin{aligned} \Delta^\dagger_{x,y} &= (awq + bzq)^\dagger \\ &= q^\dagger(aw + bz) \\ &= q^\dagger\Delta^\dagger \end{aligned} \tag{32}$$

□

Define the bundle vector E at (x, y) as the vector space corresponding to the kernel of the $\Delta^\dagger$ map at (x, y).

Corollary 3.6. $E_{x,y} = E_{xq,yq}$*, meaning* x, y *are projective coordinates over the quaternions.*

The above makes E a bundle on the projective space $\mathbb{P}^1(\mathbb{H}) \cong S^4$. On this compact space, we can calculate topological charge.

Because of this symmetry, we can specialize to the case $(w_1, w_2) = (0, 1)$ in the ADHM equations to get a solution. This simplifies the operator $\Delta^\dagger$ to

$$\Delta^\dagger = \begin{pmatrix} I & B_2 + z_2 & B_1 + z_1 \\ J^\dagger & -\bar{B}_1^\dagger - \bar{z}_1 & \bar{B}_2^\dagger + \bar{z}_2 \end{pmatrix} \tag{33}$$

Solutions to ADHM correspond to Ψ such that

$$\Delta^\dagger\Psi = 0. \tag{34}$$

It is easy to see that

$$\Delta^\dagger \Delta = \begin{pmatrix} f^{-1} & 0 \\ 0 & f^{-1} \end{pmatrix} \tag{35}$$

for some Hermitian f. We can also construct an *orthonormal* matrix M whose columns span $\ker \Delta^\dagger$. Clearly then:

$$\Delta^\dagger M = 0$$

The set of solutions Ψ to $\Delta^\dagger \Psi = 0$ gives rise to M and gives a connection: MdM.

We can then define the projection operator:

$$Q := \Delta f \Delta^\dagger \tag{36}$$

as well as

$$P := MM^\dagger \tag{37}$$

Lemma 3.7. *$P + Q = 1$. That is, P projects into the null space of $\Delta^\dagger$.*

Proposition 3.8. *This gives rise to a connection $A = M^\dagger dM$*

Proof. Take s a section so that Ms gives a section on $E = \ker \Delta^\dagger$, then

$$\begin{aligned} M\mathrm{d}s + MAs &= \mathrm{d}_A(Ms) \\ &= P\mathrm{d}(Ms) \\ &= MM^\dagger \mathrm{d}(Ms) \\ &= M(\mathrm{d}s + (M^\dagger \mathrm{d}M)s) \end{aligned} \tag{38}$$

giving our result. □

Proposition 3.9. *$A \in \mathfrak{su}(n)$.*

Proof. $A^\dagger = (dM)^\dagger M = -M^\dagger dM$ because of normalization: $M^\dagger M = 1$. □

Proposition 3.10. *A is anti-self-dual.*

Proof.

$$
\begin{aligned}
F_{\mu\nu} &= \partial_{[\mu}A_{\nu]} + A_{[\mu}A_{\nu]} \\
&= \partial_{[\mu}(M^\dagger\partial_{\nu]}M) + (M^\dagger\partial_{[\mu}M)(M^\dagger\partial_{\nu]}M) \\
&= (\partial_{[\mu}M^\dagger)(\partial_{\nu]}M) + (M^\dagger\partial_{[\mu}M)(M^\dagger\partial_{\nu]}M) \\
&= (\partial_{[\mu}M^\dagger)(\partial_{\nu]}M) + (\partial_{[\mu}M^\dagger)M(M^\dagger\partial_{\nu]}M) \\
&= (\partial_{[\mu}M^\dagger)(1-P)(\partial_{\nu]}M) \\
&= (\partial_{[\mu}M^\dagger)Q(\partial_{\nu]}M) \\
&= (\partial_{[\mu}M^\dagger)\Delta f\Delta^\dagger(\partial_{\nu]}M) \\
&= M^\dagger(\partial_{[\mu}\Delta)f(\partial_{\nu]}\Delta^\dagger)M
\end{aligned}
\tag{39}
$$

The term involving the derivatives of these Δ operators

$$(\partial_{[\mu}\Delta)f(\partial_{\nu]}\Delta^\dagger) \tag{40}$$

can be reduced to the action of sigma matrices $-i\sigma_\mu$ on f:

$$
\begin{aligned}
\partial_\mu\Delta &= -i\sigma_\mu \\
\Rightarrow (\partial_{[\mu}\Delta)f(\partial_{\nu]}\Delta^\dagger) &= (-i\sigma_{[\mu}\otimes I_k)(I_2\otimes f)(-i\sigma^\dagger_{\nu]}\otimes I_k) \\
&= -2i\sigma_{\mu\nu}\otimes f
\end{aligned}
\tag{41}
$$

And we know $\star\sigma_{\mu\nu} = -\sigma_{\mu\nu}$ This illusatrates how the underlying quaternionic structure gives rise to rise to ASD solutions. □

Proposition 3.11. *The topological charge of E when considered as a bundle over S^4 is $-k$*

Proof. (Sketch) Note that $W\oplus(V\otimes U)\cong\mathbb{C}^{n+2k} = E\oplus E^\perp$. Since E has dimension n this leaves a complement of complex dimension $2k$. This can be identified as k one-dimensional copies of the quaternions, so that $W\oplus(V\otimes U)$ decomposes as a direct sum

$$E\oplus\mathbb{H}^{\oplus k} \tag{42}$$

so corresponds to k quaternion line bundles over S^4. In fact this turns out to be the **tautological line bundle** Σ.

Now from simple Chern theory, we know:

$$0 = c_2(\mathbb{C}^{n+2k}) = c_2(E) + kc_2(\Sigma). \tag{43}$$

But the second chern number of a quaternionic tautological bundle is 1 (analogous to how the first chern number of a complex tautological bundle is 1). This gives $c_2(E) = -k$. □

Corollary 3.12. *A is a framed connection, and the topological charge is $-k$.*

Proof. By the conformal invariance proved before, A over $\mathbb{R}^4$ extends to a connection over $S^4 = {}^1(\mathbb{H})$. □