THE ADHM EQUATIONS

E568G

Contents

1 Introduction **3**
- 1.1 What are Instantons? . . . 3
- 1.2 What are Twistors? . . . 3
- 1.3 What is the ADHM construction? . . . 4

2 Geometry **4**
- 2.1 A Sliver of Differential Geometry . . . 4
  - 2.1.1 Chern and Chern-Simons Forms . . . 5
  - 2.1.2 Chern Numbers . . . 6
- 2.2 Holomorphic and Hermitian Structure . . . 7
  - 2.2.1 Dolbeault Cohomology . . . 7
  - 2.2.2 Holomorphic Vector Bundles . . . 7
  - 2.2.3 Hermitian Structure . . . 9
- 2.3 Line Bundles and Serre Twisting Sheaves . . . 10
- 2.4 Twisted Dolbeault Cohomology . . . 11

3 Yang-Mills Instantons **13**
- 3.1 Yang-Mills Theories . . . 13
  - 3.1.1 A Light-Speed Review of Maxwell's Electrodynamics . . . 13
  - 3.1.2 Non-Abelian Yang-Mills . . . 13
- 3.2 Instantons and Anti-Self Duality . . . 14
- 3.3 The BPST Instanton- an Explicit Example . . . 15

4 Twistor Theory **17**
- 4.1 Spinor-Helicity Variables . . . 17
- 4.2 Twistor Space . . . 19
  - 4.2.1 The Double Fibration . . . 19
  - 4.2.2 Twistor Space as Serre Twisting Sheaves: . . . 21
- 4.3 The Ward Correspondence . . . 22
  - 4.3.1 Anti-Self Duality in Spinor-Helicity Variables . . . 22
  - 4.3.2 The Zero Curvature Representation . . . 22
  - 4.3.3 Ward's Theorem . . . 23
  - 4.3.4 The Ward Construction . . . 24
  - 4.3.5 Holonomy . . . 24

5 The ADHM Construction **28**
- 5.1 Building Bundles via Cohomology . . . 28
- 5.2 The General ADHM Construction . . . 29
  - 5.2.1 The ADHM Monad . . . 29
  - 5.2.2 The Complex ADHM Equation . . . 30
  - 5.2.3 The ADHM Theorem . . . 30
- 5.3 The Rank Two Charge One Solution . . . 30
  - 5.3.1 The Rank Two Charge One Monad . . . 30
  - 5.3.2 The Zero-Size Limit . . . 31
- 5.4 Euclidean Instantons . . . 31
  - 5.4.1 The Euclidean ADHM Equations . . . 32
- 5.5 Conformal Compactification . . . 34

5.6 The BPST Instanton Revisited . . . 35

6 Concluding Remarks **36**

A Appendix: Geometry **37**
A.1 Grassmann's Exterior Algebra . . . 37
A.1.1 Differential Forms . . . 38
A.1.2 Integration . . . 39
A.1.3 Forms Valued in Vector Bundles . . . 40
A.2 Fibre-Bundles . . . 41
A.2.1 Smooth Fibre Bundles . . . 41
A.2.2 Connections and Curvature . . . 42
A.3 Complex Geometry . . . 45
A.3.1 Complex Manifolds . . . 45
A.3.2 Holomorphic Vector Bundles . . . 46
A.3.3 Complex Projective Space . . . 47

B Appendix: Sheaves and Cohomology **48**
B.1 Sheaves . . . 48
B.2 Cochain Complexes . . . 49
B.3 de-Rham Cohomology . . . 50
B.4 Čech Cohomology . . . 50

1 Introduction

Quantum field theory is a vast subject. There are many interesting facets one could talk about at length and almost none for which a single essay such as this one could possibly capture entirely. The aim of this article is to explore the ADHM construction through the lens of twistor theory. Atiyah, Drinfeld, Hitchin and Manin, after whom such methods- and consequently this essay too- are named, managed to find a classification of Yang-Mills instantons [1] using linear algebra and complex geometry. This is favourable over the traditional characterisation of instantons through a partial differential equation.

1.1 What are Instantons?

If we are to explore the ADHM construction it is prudent to first appreciate why we might care about instantons. Indeed it will serve us well to know what *are* instantons? The short but perhaps not so illuminating answer is; instantons are self or anti-self dual solutions to the Yang-Mills equation of motion. More concretely, and importantly more instructively; instantons are finite minima of the Euclidean Yang-Mills action within a given topological sector of the action. This sector is labelled by an integer k, often called the topological charge. They are a field configuration with a well defined size and a stability owed to a type of topological rigidity. This rigidity stems from the fact that you cannot continuously deform an instanton solution back to the vacuum via a path with finite action. It is important to realise that these are *not* particles in the ordinary sense. They are saddle solutions of the analytically continued path integral and contribute non-perturbative effects to amplitudes in Lorentzian spacetime.

1.2 What are Twistors?

The language of the ADHM construction is that of complex projective geometry. The proper way to understand this setting is to appreciate twistor theory. Twistor theory was first developed by Sir Roger Penrose and his collaborators in the 1960's [2]. The theory collects itself as a correspondence between conformal structure on spacetime and complex structure apparent in the space of null rays. It was Penrose's contention that null rays themselves are more fundamental than points in spacetime. The name is intriguing and owes its designation to a 'twisted' stereographic projection of the Hopf fibration onto $\mathbb{R}^3$. The funny spelling is supposed to remind you of the spinors inherent in the formalism, as soon we shall see.

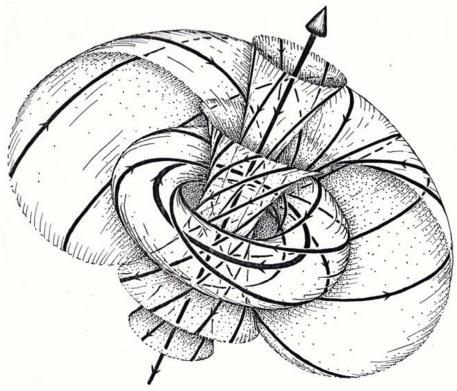


Figure 1: Robinson congruence of light rays

The Robinson congruence is pretty but arises specifically for Lorentzian spacetime. We are interested in Euclidean instantons and as such won't see this funny picture again.

1.3 What is the ADHM construction?

The ADHM construction is firstly, a powerful tool for constructing instantons. It uses only simple linear algebra as input data, yielding instantons of a chosen charge, gauged in $SU(N)$ for a chosen N. Secondly, the construction provides a simple way to characterise the moduli space $\mathcal{M}_{k,N}$ of charge k, $SU(N)$ instantons on $\mathbb{R}^4$ with the Euclidean metric.

We will strive in this text to build the ADHM construction from the ground up, utilising the tools of twistor theory. The first use of the ADHM data as mentioned above will only become apparent in section 5. The second use, however currently unjustified, can be appreciated immediately.

The ADHM data [1] for a charge k, $SU(N)$ instanton consists of:

1. A $k \times N$ matrix I
2. An $N \times k$ matrix J
3. A pair of $k \times k$ matrices B_1 and B_2

Subject to the following equations:

$$\mu_{\mathbb{C}} \coloneqq IJ + [B_1, B_2] = 0 \qquad \mu_{\mathbb{R}} \coloneqq II^\dagger - J^\dagger J + [B_1, B_1^\dagger] + [B_2, B_2^\dagger] = 0$$

The moduli space is:

$$\mathcal{M}_{k,N} = \{(I, J, B_1, B_2) \mid \mu_{\mathbb{C}} = \mu_{\mathbb{R}} = 0\} / U(k)$$

Where the $U(k)$ symmetry is given by:

$$\begin{aligned} B_{1,2} &\mapsto U B_{1,2} U^{-1} \\ I &\mapsto UI \\ J &\mapsto JU^{-1} \end{aligned}$$

2 Geometry

The bread and butter of the twistor theorist, and so for us too, is geometry. Much of the geometry we use is assumed to be familiar to the reader. We will not endeavour to begin with foundations of the subject, but appendix A is provided for the physicist who might like to brush up on his or her bundles, forms and connections; or for the mathematician unfamiliar with the notation and conventions preferred in the physics literature. To ease the notational burden, we will broadly use script font for real objects, and standard font in the complex case.

2.1 A Sliver of Differential Geometry

The topological charge of an instanton bundle is given by a quantity called the second Chern number. Essentially it boils down to the homotopy of $S^3 \hookrightarrow S^3$. We begin by defining Chern forms. The conventions used here are typical in physics literature. A good reference is Nakahara's book [3]. For an alternative approach closer to the mathematics literature see [4].

2.1.1 Chern and Chern-Simons Forms

Chern and Chern-Simons forms provide useful topological invariants for principal bundles. They will in particular allow us to characterise the topological charge of our instantons.

Throughout this section we assume we have a principal bundle $\mathscr{P}$ with a matrix structure group G, connection one form ω, and curvature Ω. The connection and curvature have local representations $\mathcal{A}$ and $\mathcal{F}$ on the base manifold.

Definition 2.1. Given a local connection $\mathcal{A}$, and its curvature $\mathcal{F}$, the Chern-Weil form $C(\mathcal{F})$ is given by:

$$C(\mathcal{F}) \coloneqq det\left(\mathbf{1} + \frac{i}{2\pi}\mathcal{F}\right)$$

One easily sees $C(\mathcal{F})$ is gauge invariant:

$$C(g\mathcal{F}g^{-1}) = det(g \cdot g^{-1} + \frac{i}{2\pi}g\mathcal{F}g^{-1}) = det(g)C(\mathcal{F})det(g)^{-1} = C(\mathcal{F})$$

Expanding this determinant, we find:

$$C(\mathcal{F}) = \mathbf{1} + C_1(\mathcal{F}) + C_2(\mathcal{F}) + \cdots + C_{\lfloor \frac{n}{2} \rfloor}(\mathcal{F})$$

Where n is the dimension of the base manifold $\mathscr{M}$ and $C_k(\mathcal{F})$ is a $2k$-form called the k^{th} Chern form.

Turning our heads toward Yang-Mills instantons, we specialise now to $\mathscr{M} = \mathbb{R}^4$, $G = SU(N)$. In this case:

$$C_2(\mathcal{F}) = \frac{1}{8\pi^2}tr(\mathcal{F} \wedge \mathcal{F})$$

The second Chern form is closed as it is a top degree form[1].

Since de-Rham cohomology tells us $H^4_{dR}(\mathbb{R}^4) = 0$ we conclude the second Chern form is exact:

$$C_2(\mathcal{F}) = dY_3$$

This 3-form Y_3 [5] is known as the Chern-Simons 3-form.

Proposition 2.1 (Chern-Simons)**.** The Chern-Simons 3-form is given by:

$$Y_3 = \frac{1}{8\pi^2}tr(\mathcal{A} \wedge d\mathcal{A} + \frac{2}{3}\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A})$$

up-to an irrelevant exact form.

Proof of Proposition 2.1.

$$\begin{aligned}
\frac{1}{8\pi^2}d\left(tr(\mathcal{A} \wedge d\mathcal{A} + \frac{2}{3}\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A})\right) &= \frac{1}{8\pi^2}tr(d\mathcal{A} \wedge d\mathcal{A} + \frac{2}{3}d\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} - \frac{2}{3}\mathcal{A} \wedge d\mathcal{A} \wedge \mathcal{A} + \frac{2}{3}\mathcal{A} \wedge \mathcal{A} \wedge d\mathcal{A}) \\
&= \frac{1}{8\pi^2}tr(d\mathcal{A} \wedge d\mathcal{A} + 2\mathcal{A} \wedge \mathcal{A} \wedge d\mathcal{A}) \\
&= \frac{1}{8\pi^2}tr(d\mathcal{A} \wedge d\mathcal{A} + d\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} + \mathcal{A} \wedge \mathcal{A} \wedge d\mathcal{A} + \underbrace{\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A}}_{0}) \\
&= \frac{1}{8\pi^2}tr((d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}) \wedge (d\mathcal{A} + \mathcal{A} \wedge \mathcal{A})) \\
&= \frac{1}{8\pi^2}tr(\mathcal{F} \wedge \mathcal{F})
\end{aligned}$$

[1] In fact the second Chern form is closed in general by the Bianchi identity.

□

2.1.2 Chern Numbers

Using the Chern forms we can define Chern numbers for even dimensional manifolds:

Definition 2.2 (Chern Number)**.** Let $\mathscr{M}$ be a 2n dimensional manifold at the base of $\mathscr{P}$. Then the Chern numbers for the bundle are:

$$c_{i_1 \dots i_r} = \int_{\mathscr{M}} C_{i_1}(\mathcal{F}) \wedge \cdots \wedge C_{i_r}(\mathcal{F})$$

where $i_1 + \cdots + i_r = n$

We note that for a four dimensional manifold there is only one Chern number given by:

$$c_2 = \int C_2(\mathcal{F})$$

Note that Chern forms are indicated with capital C's and Chern numbers with lower case c's. We will see later that the second Chern number is equal to the instanton charge. In fact c_2 is a topological invariant, as summarised in the following proposition:

Proposition 2.2. The Chern numbers are topological invariants of the bundle. That is to say they are independent of the choice of connection $\mathcal{A}$.

Proof of Proposition 2.2. We will prove the case $n = 4$ here. Of course this is far from a general proof but we will only need this result. This proof closely follows the proof in chapter 11 of [3]:

The affine space of connections $\mathscr{A}$ is locally modelled on $\Omega^1(\mathscr{M}, ad\mathscr{P})$. If we have $\mathcal{A}_0, \mathcal{A}_1 \in \mathscr{A}$ then it follows that $\mathcal{A}_t = \mathcal{A}_0 + t(\mathcal{A}_1 - \mathcal{A}_0)$ also lies in $\mathscr{A}$ for some $t \in \mathbb{R}$. Let $d_t \coloneqq d_{\mathcal{A}_t}$, $\mathcal{F}_t \coloneqq d\mathcal{A}_t + \mathcal{A}_t \wedge \mathcal{A}_t$ and $\tau \coloneqq \mathcal{A}_1 - \mathcal{A}_0$. Then:

$$\left.\begin{aligned} \mathcal{F}_t &= d\mathcal{A}_0 + td\tau + (\mathcal{A}_0 + t\tau) \wedge (\mathcal{A}_0 + t\tau) \\ &= \mathcal{F}_0 + t(d\tau + [\mathcal{A}_0, \tau]) + \frac{1}{2}t^2[\tau, \tau] \\ &= \mathcal{F}_0 + td_0\tau + \frac{1}{2}t^2[\tau, \tau] \end{aligned}\right\} \implies \begin{aligned} \frac{d\mathcal{F}_t}{dt} &= d_0\tau + t[\tau, \tau] \\ &= d\tau + [\mathcal{A}_0, \tau] + t[\tau, \tau] \\ &= d\tau + [\mathcal{A}_t, \tau] = d_t\tau \end{aligned}$$

$$C_2(\mathcal{F}_t) = \frac{1}{8\pi^2} tr(\mathcal{F}_t \wedge \mathcal{F}_t) \implies \frac{dC_2(\mathcal{F}_t)}{dt} = \frac{1}{4\pi^2} tr(\frac{d\mathcal{F}_t}{dt} \wedge \mathcal{F}_t)$$

$$\begin{aligned} d\,tr(\tau \wedge \mathcal{F}_t) &= tr(d\tau \wedge \mathcal{F}_t) - tr(\tau \wedge d\mathcal{F}_t) \\ &= tr(d\tau \wedge \mathcal{F}_t) + tr([\mathcal{A}_t, \tau] \wedge d\mathcal{A}_t) + \underbrace{tr([\mathcal{A}_t, \tau] \wedge \mathcal{A}_t \wedge \mathcal{A}_t)}_{=0} \\ &= tr(d\tau \wedge \mathcal{F}_t) + tr([\mathcal{A}_t, \tau] \wedge \mathcal{F}_t) = tr(d_t\tau \wedge \mathcal{F}_t) \end{aligned}$$

$$\frac{dC_2(\mathcal{F}_t)}{dt} = \frac{1}{4\pi^2} d\,tr(\tau \wedge \mathcal{F}_t)$$

Integrating the result on $t \in [0, 1]$ tells us that $C(\mathcal{F}_1) - C(\mathcal{F}_0)$ is exact. Hence for two connections $\mathcal{A}, \mathcal{A}' \in \mathscr{A}$ we get the same Chern number by Stoke's theorem. □

2.2 Holomorphic and Hermitian Structure

We now turn our attention to complex geometry. It is assumed that the reader is familiar with complex structure, almost complex structure, complex projective spaces, and holomorphic vector bundles. For those who wish to review such material, once again appendix A is your friend. We will begin however with an alternative characterisation of holomorphic vector bundles, using twisted cohomology. For the reader that wishes to review cohomology, appendix B is at your disposal.

2.2.1 Dolbeault Cohomology

In this section we develop the necessary tools to talk about the cohomology sensitive to holomorphic structure; Dolbeault cohomology. A great reference for this section is the book by Griffiths [6].

Let X be a complex manifold. Then in particular, X has an almost complex structure $J : TX \to TX,\ \ J^2 = -id$. Then we have $T_{\mathbb{C}}X = T^{1,0}X \oplus T^{0,1}X$ the $\pm i$ eigenspaces of J. Similarly we can define:

$$T^*_{\mathbb{C}}X = T^{*1,0}X \oplus T^{*0,1}X$$

From this bifurcation one sees:

$$\Lambda^r T^*_{\mathbb{C}}X = \bigoplus_{p+q=r} \left(\Lambda^p T^{*1,0} \otimes \Lambda^q T^{*0,1}\right)$$
$$\Omega^{p,q}(X) = \Gamma\left(\Lambda^p T^{*1,0} \otimes \Lambda^q T^{*0,1}\right)$$

Definition 2.3 (Dolbeault operator)**.** The de-Rham operator naturally splits into holomorphic and anti-holomorphic exterior derivatives:

$$d = \partial + \bar{\partial}$$
$$\partial : \Omega^{p,q}(X) \to \Omega^{p+1,q}(X)$$
$$\bar{\partial} : \Omega^{p,q}(X) \to \Omega^{p,q+1}(X)$$

Noting the nilpotence of the Dolbeault operator, we build the Dolbeault cohomology for a complex manifold X from the complex:

$$\left(\bigoplus_k \Omega^{0,k}(X), \bar{\partial}\right)$$

Hence we get the following cochain:

$$0 \hookrightarrow \Omega^{0,0}(X) \xrightarrow{\bar{\partial}} \Omega^{0,1}(X) \xrightarrow{\bar{\partial}} \Omega^{0,2}(X) \xrightarrow{\bar{\partial}} \dots$$

Dolbeault cohomology is then given by the quotient:

$$H^k(X) = \frac{\ker(\bar{\partial} : \Omega^{0,k}(X) \to \Omega^{0,k+1}(X))}{\mathrm{im}(\bar{\partial} : \Omega^{0,k-1}(X) \to \Omega^{0,k}(X))}$$

2.2.2 Holomorphic Vector Bundles

It will be convenient for us to use a definition of 'holomorphic vector bundle' that differs from the standard but equivalent description in terms of holomorphic transition functions. We shall make use of a 'partial connection' to prescribe holomorphicity to sections. See [6] for a good reference.

Definition 2.4 (Holomorphic Vector Bundle)**.** A smooth rank n complex vector bundle $E \to X$ is in fact a holomorphic vector bundle if it comes equipped with a zero curvature covariant Dolbeault operator:

$$\bar{\partial}_A : \Omega^{0,0}(X, E) \to \Omega^{0,1}(X, E)$$

In a local patch $U \subset X$ this can be expressed with a 'partial connection' $A \in \Omega^{0,1}(U, \mathfrak{gl}(n, \mathbb{C}))$:

$$\bar{\partial}_A = \bar{\partial} + A$$

And the zero curvature condition is:

$$F^{0,2} = \bar{\partial}A + A \wedge A = 0$$

A section $S \in \Gamma(E)$ is holomorphic if it is covariantly constant (i.e. $\bar{\partial}_A S = 0$).

Remarks 2.1.

1. The partial connection is entirely analogous to a standard connection on a smooth vector bundle and gives rise to a Dolbeault operator which is just like a normal covariant derivative. The only novelty here is that we only care to covariantly differentiate in the anti-holomorphic directions.
2. The covariant derivative can of course be extended to all p, q forms via a Leibniz rule.
3. Just as for smooth bundles there is a notion of a holomorphic principal bundle P equipped with a connection $\omega \in \Omega^{0,1}(P, \mathfrak{g})$. If the holomorphic vector bundle is associated to this principal bundle, then the partial connection arises from a representation of the group entirely analogously to the smooth case.

Lemma 2.1. The flat condition $F^{0,2}$ is equivalent to the nilpotence of $\bar{\partial}_A$.

Proof of Lemma 2.1. In a local patch with partial connection A, for any section on this patch we have:

$$\begin{aligned}
\bar{\partial}_A \circ \bar{\partial}_A S &= \bar{\partial}_A(\bar{\partial}S + AS) \\
&= \bar{\partial}(AS) + A\bar{\partial}S + A \wedge AS \\
&= \bar{\partial}AS - A\bar{\partial}S + A\bar{\partial}S + A \wedge AS \\
&= F^{0,2}S
\end{aligned}$$

□

As a result of lemma 2.1 the following complex:

$$0 \hookrightarrow \Omega^{0,0}(X, E) \xrightarrow{\bar{\partial}_A} \Omega^{0,1}(X, E) \xrightarrow{\bar{\partial}_A} \Omega^{0,1}(X, E) \xrightarrow{\bar{\partial}_A} \dots \xrightarrow{\bar{\partial}_A} \Omega^{0,n}(X, E) \xrightarrow{\bar{\partial}_A} 0$$

gives us a 'twisted Dolbeault cohomology' analogously to the twisted de Rham theory algebraic geometers interest themselves with. We will revisit this in section 2.4.

Proposition 2.3. The standard definition A.17 of a holomorphic vector bundle (in appendix A) and the definition 2.4 are equivalent.

Lemma 2.2. The flat condition $F^{0,2} = \bar{\partial}A + A \wedge A = 0$ on any patch U implies we can choose a gauge such that $A = 0$ on U.

Proof of Lemma 2.2. We will show that choosing $g : U \to GL(n, \mathbb{C})$ such that $gAg^{-1} - (\bar{\partial}g)g^{-1} = 0$ is consistent with $F^{0,2} = 0$.

$$\begin{aligned} \bar{\partial}g &= gA \\ 0 = \bar{\partial}(gA) &= \bar{\partial}g \wedge A + g\bar{\partial}A = gA \wedge A + g\bar{\partial}A = gF^{0,2} \end{aligned}$$

Where in the second line we took the Dolbeault derivative. Hence we see that the existence of such a gauge is contingent on $F^{0,2} = 0$. □

Proof of Proposition 2.3. We begin with a holomorphic bundle as defined in definition 2.4. We will show that a transition function is necessarily holomorphic in the sense of A.17.

Consider transition function $Y : U \cap V \to GL(n, \mathbb{C})$ and a section S that locally is given by $S_U \in \Omega^{0,0}(U, E)$ on U and $S_V \in \Omega^{0,0}(V, E)$ on V. Then we have $S_U = YS_V$ and on $U \cap V$:

$$\begin{aligned} (\bar{\partial} + A_U)S_U &= (\bar{\partial} + YA_VY^{-1} - (\bar{\partial}Y)Y^{-1})YS_V \\ &= \bar{\partial}(YS_V) + YA_VS_V - (\bar{\partial}Y)S_V \\ &= Y(\bar{\partial} + A_V)S_V \end{aligned}$$

But by Lemma 2.2 we can choose a gauge such that $A_U = A_V = 0$. Hence:

$$\begin{aligned} \bar{\partial}S_U = Y\bar{\partial}Y^{-1}S_U \; \forall \; S \\ \implies \bar{\partial}Y = 0 \end{aligned}$$

□

2.2.3 Hermitian Structure

A Hermitian inner product is the complex analogue of the standard real inner product. As familiar from quantum theory, it is a sesquilinear positive definite inner product on a complex vector space. One can equip a complex vector bundle with a Hermitian structure in much the same way as one equips the tangent bundle of a smooth manifold with a metric tensor.

Suppose $E \to X$ is a rank n holomorphic vector bundle. Then we make the following claim:

Proposition 2.4. A Hermitian structure on E is equivalent to linear isomorphisms on each fibre E_x:

$$\overline{E_x} \cong E_x^*$$

where the bar indicates complex conjugation and the star indicates the vector space dual.

Proof of Proposition 2.4. A Hermitian structure on the fibre E_x is a map $h_x : E_x \times E_x \to \mathbb{C}$ anti-linear in the first slot and linear in the second. Define the following anti-linear isomorphism:

$$\begin{aligned} \phi_x : E_x &\to E_x^* \\ v &\mapsto h_x(v, \cdot) \end{aligned}$$

Then pre-composing ϕ_x with the conjugation involution $i : \overline{E_x} \to E_x \; ; \quad v \mapsto \bar{v}$ gives the desired linear isomorphism:

$$\begin{aligned} \phi_x \circ i : \overline{E_x} &\xrightarrow{\sim} E_x^* \\ v &\mapsto \bar{h}_x(\bar{v}, \cdot) \end{aligned}$$

□

2.3 Line Bundles and Serre Twisting Sheaves

Algebraic geometers have much interest in the spaces denoted $\mathcal{O}(n),\ n \in \mathbb{Z}$ over $\mathbb{CP}^k$ [7] and regard them as locally free sheaves of rank 1. For now it will be in our interest to view these objects instead as complex line bundles over $\mathbb{CP}^k$, an equivalent viewpoint. These line bundles will play a crucial role in the ADHM construction on twistor space.

For now let us discuss $\mathcal{O}(n)$ on $\mathbb{CP}^1$. Let $[w_1 : w_2]$ be homogenous co-ordinates on $\mathbb{CP}^1$. Take U_1, U_2 to be the usual charts $U_1 = \{[w_1 : w_2] \mid w_1 \neq 0\}$, $U_2 = \{[w_1 : w_2] \mid w_2 \neq 0\}$. Furthermore we define $z = \frac{w_2}{w_1}$ to be our co-ordinate on the first patch (around $z = 0$) and $\tilde{z} = \frac{w_1}{w_2}$ our co-ordinate on the second patch (around the pole $z = \infty$), noting that $\tilde{z} = \frac{1}{z}$ on the chart overlap. Now consider holomorphic maps $\sigma_1 : U_1 \to \mathbb{C}$ and $\sigma_2 : U_2 \to \mathbb{C}$. We will define the holomorphic line bundle $\mathcal{O}(n)$ by gluing these 'sections' in the sense of transition functions (see definition A.17).

$$f_{11}(z) = 1 \tag{2.1}$$
$$f_{22}(\tilde{z}) = 1 \tag{2.2}$$
$$f_{12}(z) = z^n \tag{2.3}$$
$$f_{21}(\tilde{z}) = \tilde{z}^n \tag{2.4}$$

Now because σ_1 and σ_2 are holomorphic functions of z on $U_1 \cap U_2$, we can expand them in a power series on this intersection, for now taking $n > 0$:

$$\sigma_1(z) = a_0 + a_1 z + a_2 z^2 + \dots \tag{2.5}$$
$$\sigma_2(z) = f_{21}\sigma_1 = a_0 z^{-n} + a_1 z^{-n+1} + a_2 z^{-n+2} + \dots + a_{n-1} z^{-1} + a_n + a_{n+1} z + \dots \tag{2.6}$$
$$\tag{2.7}$$

This expansion for σ_1 is finite everywhere on U_1. However extending σ_2 holomorphically to all of U_2 requires that σ_2 has no poles at $z = \infty$ which forces $a_m = 0\ \forall\ m > n$. Hence we have global sections of the form:

$$\sigma = \begin{cases} a_0 + a_1 z + \dots + a_n z^n \text{ on } U_1 \\ a_n + a_{n-1}\tilde{z} + \dots + a_0 \tilde{z}^n \text{ on } U_2 \end{cases}$$

.

Repeating the same argument for $n < 0$ and taking $k = -n > 0$:

$$\sigma_1(z) = a_0 + a_1 z + a_2 z^2 + \dots \tag{2.8}$$
$$\sigma_2(z) = f_{21}\sigma_1 = a_0 z^k + a_1 z^{k+1} + a_2 z^{k+2} + \dots \tag{2.9}$$

Again σ_1 extends holomorphically to include $z = 0$ but σ_2 has a pole at $z = \infty$ unless $a_m = 0\ \forall\ m$. Hence for negative n, $\mathcal{O}(n)$ admits no global sections.

We now consider the holomorphic tangent and cotangent bundles on $\mathbb{CP}^1$, which one notes are both examples of line bundles.

$$\frac{\partial}{\partial z} = \frac{\partial \tilde{z}}{\partial z}\frac{\partial}{\partial \tilde{z}} = -z^{-2}\frac{\partial}{\partial \tilde{z}} \tag{2.10}$$

$$\sigma_1(z)\frac{\partial}{\partial z} = \sigma_1(z)\frac{\partial \tilde{z}}{\partial z}\frac{\partial}{\partial \tilde{z}} = -\sigma_1(z)z^{-2}\frac{\partial}{\partial \tilde{z}} = \sigma_2(z)\frac{\partial}{\partial \tilde{z}} \tag{2.11}$$

$$\sigma_2 = -z^{-2}\sigma_1(z) \tag{2.12}$$

$$dz = \frac{\partial z}{\partial \tilde{z}}d\tilde{z} = -z^2 d\tilde{z} \tag{2.13}$$

$$\eta_1(z)dz = \eta_1(z)\frac{\partial z}{\partial \tilde{z}}d\tilde{z} = -\eta_1(z)z^2 d\tilde{z} = \eta_2(z)d\tilde{z} \tag{2.14}$$

$$\eta_2(z) = -z^2\eta_1(z) \tag{2.15}$$

As such we see:

$$T^{1,0}\mathbb{CP}^1 \cong \mathcal{O}(2)$$
$$T^{*1,0}\mathbb{CP}^1 \cong \mathcal{O}(-2)$$

Remarks 2.2.

1. The factors of -1 in the transition functions are irrelevant. One can remove them by simply negating the basis on one patch.

2. If we were to entertain ourselves with the anti-holomorphic directions we'd have little trouble finding $T^{0,1}\mathbb{CP}^1 \cong \overline{\mathcal{O}(2)}$ and $T^{*0,1}\mathbb{CP}^1 \cong \overline{\mathcal{O}(-2)}$ where the bar indicates conjugation.

3. We can and will make use of these bundle isomorphisms to write sections of the bundle $\mathcal{O}(-n)$ as $f(z,\bar{z})dz^{\frac{n}{2}}$, where $n > 0$ is a positive integer and $dz^{\frac{n}{2}}$ encodes the transformation property required.

Strictly speaking to make the third remark precise we ought introduce the spin bundle $k \cong \mathcal{O}(-1)$ and first define $dz^{\frac{1}{2}}$ as a section of this bundle. Being more careful here will not change much of anything so $dz^{\frac{1}{2}}$ may be taken to be convenient notation encoding the right transformations and not regarded as a spinor at this stage, although of course that is what it is.

2.4 Twisted Dolbeault Cohomology

We can further generalise this cohomology for Dolbeault forms valued in some holomorphic vector bundle. We shall use the partial connection on the bundle to consider the twisted cohomology.

As we saw in section 2.2.2, we can use the nilpotent partial connection on a holomorphic bundle E to obtain the following cochain:

$$0 \hookrightarrow \Omega^{0,0}(X,E) \xrightarrow{\bar{\partial}_A} \Omega^{0,1}(X,E) \xrightarrow{\bar{\partial}_A} \Omega^{0,1}(X,E) \xrightarrow{\bar{\partial}_A} \ldots \xrightarrow{\bar{\partial}_A} \Omega^{0,n}(X,E) \xrightarrow{\bar{\partial}_A} 0$$

The cohomology we get is then:

$$H^k(X,E) = \frac{\ker(\bar{\partial}_A : \Omega^{0,k}(X,E) \to \Omega^{0,k+1}(X,E))}{\mathrm{im}(\bar{\partial}_A : \Omega^{0,k-1}(X,E) \to \Omega^{0,k}(X,E))}$$

Consider the case when $X = \mathbb{CP}^1$ and $E = \mathcal{O}(-n)$. This example will be relevant when we come to discuss the Ward correspondence. Then elements of $\Omega^{0,1}(\mathbb{CP}^1, \mathcal{O}(-n))$ can be written in the form:

$$f(z,\bar{z})\, dz^{\frac{n}{2}} \otimes d\bar{z}$$

Where $f : \mathbb{CP}^1 \to \mathbb{C}$ is a complex function, $dz^{\frac{n}{2}}$ is a local section of $\mathcal{O}(-n)$ and $d\bar{z}$ a local section of $\Omega^{0,1}(\mathbb{CP}^1)$. To ease the notational burden, from here we drop the tensor product. From context it will always be clear what object we have.

The zeroth class is easy to compute, since:

$$\begin{aligned} H^0(\mathbb{CP}^1, \mathcal{O}(-n)) &= ker(\bar{\partial}_A : \Omega^{0,0}(\mathbb{CP}^1, \mathcal{O}(-n)) \to \Omega^{0,1}(\mathbb{CP}^1, \mathcal{O}(-n))) \\ &= ker(\bar{\partial}_A : \Gamma(\mathcal{O}(-n)) \to \Omega^{0,1}(\mathbb{CP}^1, \mathcal{O}(-n))) \\ &= \{\textit{holomorphic sections of } \mathcal{O}(-n)\} \end{aligned}$$

Theorem 2.1.

$$H^1(\mathbb{CP}^1, \mathcal{O}(n)) = \begin{cases} 0 \text{ for } n \geq -1 \\ \mathbb{C}^{-n-1} \text{ for } n \leq -2 \end{cases}$$

We will not prove the full theorem here, but rather the case $n = -2$ to illustrate the nature of such a proof.

Partial Proof of Theorem 2.1. We will exploit the fact that Dolbeault and Čech cohomology compute the same topological obstruction here. For more on Čech cohomology see appendix B. We take the standard chart cover $\{U_1, U_2\}$ of $\mathbb{CP}^1$ as given in equation A.3. We write our anti-holomorphic $\mathcal{O}(-2)$- valued one form as:

$$\alpha = f(z, \bar{z}) dz d\bar{z} \in \Omega^{0,1}(\mathbb{CP}^1, \mathcal{O}(-2))$$

Which we observe is closed:

$$\bar{\partial}\alpha = \frac{\partial f(z, \bar{z})}{\partial \bar{z}} dz \; d\bar{z} \wedge d\bar{z} = 0$$

It should be noted that we have used lemma 2.2 to set A=0. In the Čech language α is a co-cycle. There is a 'Dolbeault lemma' in perfect analogy with the Poincaré lemma for real manifolds, that insists α is locally Dolbeault exact:

$$\exists\, \sigma_1 \in \Omega^{0,0}(U_1, \mathcal{O}(-2)),\ \sigma_2 \in \Omega^{0,0}(U_2, \mathcal{O}(-2)) \text{ such that } \alpha \mid_{U_1} = \bar{\partial}\sigma_1 \text{ and } \alpha \mid_{U_2} = \bar{\partial}\sigma_2$$

The question then is when is α exact globally? In the Čech language, when can we glue these local sections together to form a global section? On the intersection $U_{1,2} = U_1 \cap U_2$:

$$0 = \alpha \mid_{U_1} - \alpha \mid_{U_2} = \bar{\partial}(\sigma_1 - \sigma_2)$$

Hence we deduce $\sigma_1 - \sigma_2$ is holomorphic, and write $\sigma_1 - \sigma_2 = g(z)dz = -g(\tilde{z})d\tilde{z}$ where z and $\tilde{z}$ are the standard co-ordinates on the patches U_1 and U_2 respectively, with $z = \tilde{z}^{-1}$ on the overlap. The minus sign accounts for the transformation property of $dz \in \Gamma(\mathcal{O}(-2))$ as noted in Remarks 2.2.

$$g(z)dz = -g\left(\frac{1}{z}\right) d\left(\frac{1}{z}\right) = \frac{1}{z^2} g\left(\frac{1}{z}\right) dz$$

Hence we see that on $U_{1,2} \cong \mathbb{C}^*$ we have $g(z) = \frac{1}{z^2} g\left(\frac{1}{z}\right)$. Expanding $g(z)$ in a Laurent series on $\mathbb{C}^*$ immediately forces $g(z) = \frac{k}{z}$, $k \in \mathbb{C}$ and hence we see the failure of these sections to glue is 1-dimensional:

$$H^1(\mathbb{CP}^1, \mathcal{O}(-2)) = \mathbb{C}$$

Which concludes the proof. □

From here on, a better notation for the Dolbeault cohomology discussed in section 2.2.1 is:

$$H^k(X, \mathcal{O}) = H^k(X)$$

Dolbeault cohomology calculates sheaf data and this notation keeps an honest accounting. One should read $\mathcal{O}$ as $\mathcal{O}(0)$.

3 Yang-Mills Instantons

We now turn to our theory of interest, Yang-Mills theory. We shall begin with a brief review of the Abelian theory, then swiftly review the general construction. We will then discuss self duality and the classification of instanton solutions, before turning to a concrete, simple example. For a detailed review of this subject [3] is a great book.

3.1 Yang-Mills Theories

3.1.1 A Light-Speed Review of Maxwell's Electrodynamics

We begin with an overview of Maxwell's electrodynamics in the vacuum; written concisely and naturally in the language of Cartan's differential forms; Maxwell theory and its generalisations thereof are geometric theories.

The electric and magnetic fields are packaged into a gauge one form $\mathcal{A}$ and its curvature $\mathcal{F} = d\mathcal{A}$. We then begin with a classical action á la Hamilton:

$$S[\mathcal{A}] = -\frac{1}{2}\int_{\mathbb{R}^4} \mathcal{F} \wedge \star\mathcal{F}$$

From which one derives the equation of motion (3.1) and the Bianchi identity (3.2):

$$d \star \mathcal{F} = 0 \tag{3.1}$$

$$d\mathcal{F} = 0 \tag{3.2}$$

We get the Bianchi identity for free by virtue of the nilpotence of the de-Rham operator. The Gauge invariance of the action leads us to interpret Maxwell's theory as one of connections on a U(1) principal bundle.

3.1.2 Non-Abelian Yang-Mills

It is natural to ask, given a compact Lie group G for which one wishes to construct a gauge-invariant theory; what should such a construction look like. These generalisations of Maxwell go by the name of Yang-Mills theory. Here we will (very) briefly review the action as we did with Maxwell.

For a Yang-Mills theory with compact structure group G and Lie algebra $\mathfrak{g} = Lie(G)$ one writes the action:

$$S[\mathcal{A}] = -\frac{1}{2}\int_{\mathbb{R}^4} \langle \mathcal{F} \wedge \star\mathcal{F} \rangle$$

Where the curvature $\mathcal{F}$ is given by:

$$\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$$

Both the connection $\mathcal{A}$ and its curvature are valued in $\mathfrak{g}$. The bracket $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is the invariant Killing form on $\mathfrak{g}$. As before we vary the action to arrive at the equation of motion (3.3). The Bianchi identity (3.4) as before is structural.

$$d_{\mathcal{A}} \star \mathcal{F} = 0 \tag{3.3}$$

$$d_{\mathcal{A}}\mathcal{F} = 0 \tag{3.4}$$

Where $d_{\rho(\mathcal{A})} = d + \rho(\mathcal{A})$ is the covariant derivative. In particular if $\rho : G \to GL(V)$ is the adjoint representation $Ad : G \to Aut(\mathfrak{g})$ then we drop the subscript ρ for convenience and so for a Lie algebra-valued k-form $\omega \in \Omega^k(\mathbb{R}^{1,3}, ad\mathscr{P})$ we have $d_{\mathcal{A}}\omega = d\omega + [\mathcal{A}, \omega]$ where one recalls $ad : \mathfrak{g} \to End(\mathfrak{g})$ sends $\mathcal{A} \mapsto [\mathcal{A}, \cdot]$. $\mathscr{P}$ here is the underlying principal bundle.

3.2 Instantons and Anti-Self Duality

As briefly introduced in 1.1, instantons are finite action solutions of the equation of motion which minimise the Euclidean Yang-Mills action within a topological sector. After we identify the killing form on $SU(N)$ with the trace over matrices we see:

$$\begin{aligned} S[\mathcal{A}] &= -\int_{\mathbb{R}^4} tr(\mathcal{F} \wedge \star\mathcal{F}) \\ &= -\frac{1}{2}\int_{\mathbb{R}^4} tr\left((\mathcal{F} + \star\mathcal{F}) \wedge \star(\mathcal{F} + \star\mathcal{F})\right) + \int_{\mathbb{R}^4} tr(\mathcal{F} \wedge \mathcal{F}) \\ &= \underbrace{-\frac{1}{2}\int_{\mathbb{R}^4} tr\left((\mathcal{F} + \star\mathcal{F}) \wedge \star(\mathcal{F} + \star\mathcal{F})\right)}_{\geq 0} + 8\pi^2 c_2 \end{aligned}$$

Hence the action is bounded below by $8\pi^2$ times the so called 'topological charge[2]' c_2; the second Chern number as introduced in section 2.1.2 (which we have assumed to be positive). The bound is saturated for anti-self dual solutions[3] $\mathcal{F} = -\star\mathcal{F}$.

We note that for anti-self dual gauge fields the Bianchi identity and Yang-Mills equation of motion coincide; as such self dual connections are automatically on shell solutions; instantons.

In co-ordinates x^i on $\mathbb{R}^4$, we shall denote the local covariant derivative along the i^{th} direction (associated to $d_{\mathcal{A}}$) by $\mathcal{D}_i$.

Proposition 3.1. The anti-self dual Yang-Mills equation on $\mathbb{R}^4$ is equivalent in co-ordinates to:

$$\mathcal{F}_{12} + \mathcal{F}_{34} = \mathcal{F}_{42} + \mathcal{F}_{13} = \mathcal{F}_{14} + \mathcal{F}_{23} = 0 \tag{3.5}$$

Proof of Proposition 3.1. For a 2-form ω on $\mathbb{R}^4$ equipped with the standard Euclidean metric:

$$(\star\omega)_{ab} = \frac{1}{2}\omega^{cd}\epsilon_{cdab} = \frac{1}{2}\omega_{cd}\epsilon_{cdab}$$

Hence:

$$\begin{aligned} (\mathcal{F} + *\mathcal{F})_{ab} &= \mathcal{F}_{ab} + \frac{1}{2}\mathcal{F}_{cd}\epsilon_{cdab} \\ (\mathcal{F} + *\mathcal{F})_{12} &= \mathcal{F}_{12} + \frac{1}{2}\mathcal{F}_{cd}\epsilon_{cd12} = \mathcal{F}_{12} + \mathcal{F}_{34} \\ (\mathcal{F} + *\mathcal{F})_{13} &= \mathcal{F}_{13} + \frac{1}{2}\mathcal{F}_{cd}\epsilon_{cd13} = \mathcal{F}_{13} + \mathcal{F}_{42} \\ (\mathcal{F} + *\mathcal{F})_{14} &= \mathcal{F}_{14} + \frac{1}{2}\mathcal{F}_{cd}\epsilon_{cd14} = \mathcal{F}_{14} + \mathcal{F}_{23} \\ \therefore\ \mathcal{F} + *\mathcal{F} = 0 &\implies \mathcal{F}_{12} + \mathcal{F}_{34} = \mathcal{F}_{13} + \mathcal{F}_{42} = F_{14} + \mathcal{F}_{23} = 0 \end{aligned}$$

[2]Sometimes in the physics literature this quantity is also called the instanton number, but we shall not use that language here.

[3]Had we instead taken the topological charge to be negative, we'd have found the bound saturated for self dual solutions.

□

Definition 3.1. We define the following complex combination of covariant derivatives:

$$D_1 \coloneqq \frac{1}{2}(\mathcal{D}_2 - i\mathcal{D}_1) \tag{3.6}$$

$$D_2 \coloneqq \frac{1}{2}(\mathcal{D}_4 - i\mathcal{D}_3) \tag{3.7}$$

Proposition 3.2. The anti-self dual equations 3.5 are equivalent to the following pair of equations:

$$[D_1, D_2] = 0 \tag{3.8}$$

$$[D_1, D_1^\dagger] + [D_2, D_2^\dagger] = 0 \tag{3.9}$$

Proof of Proposition 3.2.

$$\begin{aligned}
[D_1, D_2] &= \frac{1}{4}[\mathcal{D}_2 - i\mathcal{D}_1, \mathcal{D}_4 - i\mathcal{D}_3] \\
&= \frac{1}{4}([\mathcal{D}_2, \mathcal{D}_4] - [\mathcal{D}_1, \mathcal{D}_3]) - \frac{i}{4}([\mathcal{D}_1, \mathcal{D}_4] + [\mathcal{D}_2, \mathcal{D}_3]) \\
\therefore\ [D_1, D_2] = 0 &\iff F_{42} + \mathcal{F}_{13} = 0 \qquad F_{14} + \mathcal{F}_{23} = 0
\end{aligned}$$

$$\begin{aligned}
[D_1, D_1^\dagger] + [D_2, D_2^\dagger] &= \frac{1}{4}[\mathcal{D}_2 - i\mathcal{D}_1, \mathcal{D}_2 + i\mathcal{D}_1] + \frac{1}{4}[\mathcal{D}_4 - i\mathcal{D}_3, \mathcal{D}_4 + i\mathcal{D}_3] \\
&= \frac{i}{2}\left([\mathcal{D}_2, \mathcal{D}_1] + [\mathcal{D}_4, \mathcal{D}_3]\right) \\
\therefore\ [D_1, D_1^\dagger] + [D_2, D_2^\dagger] = 0 &\iff F_{12} + \mathcal{F}_{34} = 0
\end{aligned}$$

□

As we shall see in section 5, these equations 3.8 and 3.9 are essentially the ADHM equations 5.4 and 5.5 with $I = J = 0$. I and J carry information about the non-triviality of the instanton bundle on S^4, the conformal compactification of $\mathbb{R}^4$. Since here we have made no such effort to construct a bundle on S^4 and have simply asked for anti-self duality on $\mathbb{R}^4$, we probe here only this narrow subset of the instanton moduli space.

3.3 The BPST Instanton- an Explicit Example

We shall now explicitly look for an instanton. We will concern ourselves with $SU(2)$ Yang-Mills theory on $\mathbb{R}^4$. This solution was first discovered by Belavin, Polyakov, Schwartz and Tyupkin [8]. The way we formulate the solution here is similar to the construction used in [9].

It is very useful in fact to realise $SU(2)$ as an embedding in $\mathbb{H}$, the quaternions, which we will briefly review.

Quaternions are the largest finite normed division algebra:

$$\mathbb{H} = \{x_1 + ix_2 + jx_3 + kx_4 \mid x_1, x_2, x_3, x_4 \in \mathbb{R},\ i^2 = j^2 = k^2 = ijk = -1\}$$

We can naturally identify $\mathbb{R}^4 \cong \mathbb{H}$ via $(x_1, x_2, x_3, x_4) \mapsto x_1 + ix_2 + jx_3 + kx_4$, and we will denote both as $\mathbf{x}$. We will write $\mathscr{R}(\mathbf{x}) = x_1$ for the 'real' part of $\mathbf{x}$ and $\mathscr{I}(\mathbf{x}) = ix_2 + jx_3 + kx_4$ for the 'imaginary'

part, noting this definition of imaginary includes i,j,k in contrast to the convention for $\mathbb{C}$. Conjugation is given by $\overline{\mathbf{x}} = x_1 - ix_2 - jx_3 - kx_4$. There is a natural inner product on $\mathbb{H}$ given by $\mathbf{x} \cdot \mathbf{y} \coloneqq \mathscr{R}(\mathbf{x}\overline{\mathbf{y}})$ which agrees with Euclidean inner product on $\mathbb{R}^4$.

The utility of this exercise is not only that our spacetime is isomorphic to the quaternions, but in fact our Lie group is isomorphic to the unit quaternions:

$$SU(2) \cong \{\mathbf{x} \in \mathbb{H} \mid \mathbf{x} \cdot \mathbf{x} = 1\}$$

$$\begin{pmatrix} \alpha + i\beta & \gamma + i\delta \\ -\gamma + i\delta & \alpha - i\beta \end{pmatrix} \leftrightarrow \alpha + i\beta + j\gamma + k\delta$$

For which the group multiplication is easy to verify. Differentiating at the identity $x_1 = 1$ immediately gives:

$$\mathfrak{su}(2) \cong \mathscr{I}(\mathbb{H})$$

The proposed instanton connection and its curvature are then the $\mathscr{I}(\mathbb{H})$ valued forms on $\mathbb{H}$:

$$\mathcal{A}(\mathbf{x}) = \frac{1}{1 + |\mathbf{x}|^2} \mathscr{I}(\overline{\mathbf{x}}\, d\,\mathbf{x}) \tag{3.10}$$

$$\mathcal{F}(\mathbf{x}) = \frac{1}{(1 + |\mathbf{x}|^2)^2} d\,\overline{\mathbf{x}} \wedge d\,\mathbf{x} \tag{3.11}$$

To verify this, we must first show that $\mathcal{A}$ is anti-self dual, and further find its topological charge. Expanding $d\,\overline{\mathbf{x}} \wedge d\mathbf{x}$ it is easily seen:

$$d\,\overline{\mathbf{x}} \wedge d\mathbf{x} = 2i(dx_1 \wedge dx_2 - dx_3 \wedge dx_4) + 2j(dx_1 \wedge dx_3 - dx_4 \wedge dx_2) + 2k(dx_1 \wedge dx_4 - dx_2 \wedge dx_3)$$

We easily see that $\mathcal{F}$ is anti-self dual by proposition 3.1. To determine the topological charge we simply compute the second Chern number:

$$\begin{aligned}
c_2 &= \frac{1}{8\pi^2} \int_{\mathbb{R}^4} tr(\mathcal{F} \wedge \mathcal{F}) \\
&= -\frac{1}{4\pi^2} \int_{\mathbb{R}^4} |\mathcal{F}|^2 \\
&= -\frac{1}{4\pi^2} \int_{\mathbb{R}^4} \frac{4}{(1 + |\mathbf{x}|^2)^4} \left\{ (dx_1 \wedge dx_2 - dx_3 \wedge dx_4)^2 + (dx_1 \wedge dx_3 - dx_4 \wedge dx_2)^2 + (dx_1 \wedge dx_4 - dx_2 \wedge dx_3)^2 \right\} \\
&= -\frac{1}{4\pi^2} \int_{\mathbb{R}^4} \frac{4}{(1 + |\mathbf{x}|^2)^4} \left\{ -6\, dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \right\} \\
&= \frac{6}{\pi^2} \int_{-\infty}^{\infty} \frac{d^4x}{(1 + |\mathbf{x}|^2)^4} \\
&= \frac{6}{\pi^2} \underbrace{Area(S^3)}_{2\pi^2} \underbrace{\int_0^{\infty} dr \frac{r^3}{(1 + r^2)^4}}_{\frac{1}{12}}
\end{aligned}$$

$$c_2 = 1$$

We shall revisit this instanton later, deriving it from the ADHM construction.

4 Twistor Theory

The twistor correspondence is a non-local mapping between spacetime and complex projective space. It captures conformal spacetime structure in complex structure on 'twistor space'. It is a formalism that has proven capable of calculating amplitudes more efficiently than certain Feynman integrals in QFT [10]. Twistor theory happens to be the correct landscape in which to frame scattering of massless fields. On this point we shall elaborate immediately. A good reference for this section is the book by Penrose and Rindler, Spinors and Space-time: Volume 1 [11].

4.1 Spinor-Helicity Variables

In a tradition that might be called the 'Weinberg approach', there is an argument that quantum field theory and its excessive baggage of gauge redundant degrees of freedom is the *wrong* way to formulate fundamental physics. This arguably quite radical view is popularised for example by Nima Arkani-Hamed. Physicists sympathetic to this reformulation argue that particles (those pesky things we actually *measure*) are truly what exists, and that fields are emergent. This view predicates on an argument from symmetry. Wigner's classification tells us that it is *particles* whose existence is guaranteed by irreducible representations of the isometry group of spacetime. Fields appearing off-shell in scattering processes simply don't demand existence in the same way. Even if one doesn't find this convincing, it is hard to argue with the fact that amplitudes formulated in terms of observable kinematics (spin, helicity, momentum etc) are drastically simpler to compute than traditional Feynman integrals plagued by gauge redundant polarisation vectors. It is in the so called spinor-helicity variables that massless scattering simplifies. The twistor correspondence too is vastly simpler in these terms.

In this section we work in an analytically continued spacetime, $\mathbb{C}^4$. One can recover their favourite signature (Lorentzian, Euclidean, split or what have you) by selecting the appropriate real slice $\mathbb{R}^4 \subset \mathbb{C}^4$. We will discuss how to do this for the Euclidean signature (as is relevant for our Euclidean instantons) shortly. The conventions for these variables and their contractions, in this section, as well as the discussion of amplitudes closely follow [12].

We shall make use of the following well-known isomorphism[4] [11]:

Theorem 4.1 (Lorentz Algebra)**.** The complex Lorentz algebra $\mathfrak{so}(4)_{\mathbb{C}}$ is isomorphic to two copies of $\mathfrak{sl}(2,\mathbb{C})$:

$$\mathfrak{so}(4)_{\mathbb{C}} \cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$$

Using theorem 4.1 we can write objects in $\mathbb{C}^4$ carrying a Lorentz index μ instead with a pair of indices $\alpha\dot{\alpha}$ such that the dotted index corresponds to the first (left handed, or negative helicity) copy of the spin group $SL(2,\mathbb{C})$; the undotted index corresponds to the second (right handed or positive helicity) copy, and both dotted and undotted indices run over $\{1,2\}$.

We shall see this explicitly for the co-ordinates x^μ on $\mathbb{C}^4$:

$$x^\mu \mapsto x^{\alpha\dot{\alpha}} = (\sigma_\mu)^{\alpha\dot{\alpha}} x^\mu = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}$$

[4] In some sense the seemingly miraculous capability of the ADHM construction to translate self duality into linear information on projective space can be traced to the fact that in a complexified spacetime, the Lorentz group $(SL(2,\mathbb{C}) \times SL(2,\mathbb{C}))/\mathbb{Z}_2$ and the conformal group $SL(4,\mathbb{C})$ simply act linearly on planes.

Where $(\sigma_\mu)^{\alpha\dot{\alpha}}$ is the $\alpha\dot{\alpha}$ component of the μ^{th} Pauli matrix:

$$\sigma_\mu = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

Remarks 4.1.

1. The metric in this formalism is encoded in the determinant: $det(dx^{\alpha\dot{\alpha}}) = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$
2. x^μ null $\iff$ $det(x^{\alpha\dot{\alpha}}) = 0 \iff \exists\, y^\alpha, \tilde{y}^{\dot{\alpha}} : x^{\alpha\dot{\alpha}} = y^\alpha \tilde{y}^{\dot{\alpha}}$

We can similarly write the metric as:

$$\begin{aligned} \frac{1}{2}\epsilon_{\alpha\beta}\epsilon_{\dot{\alpha}\dot{\beta}} dx^{\alpha\dot{\alpha}} \otimes dx^{\beta\dot{\beta}} &= \frac{1}{2}\left(dx^{1\dot{1}} \otimes dx^{2\dot{2}} + dx^{2\dot{2}} \otimes dx^{1\dot{1}} - dx^{1\dot{2}} \otimes dx^{2\dot{1}} - dx^{2\dot{1}} \otimes dx^{1\dot{2}} \right) \\ &= \frac{1}{2}\big((dx^0 + dx^3) \otimes (dx^0 - dx^3) + (dx^0 - dx^3) \otimes (dx^0 + dx^3) \\ &\quad -(dx^1 - idx^2) \otimes (dx^1 + idx^2) - (dx^1 + idx^2) \otimes (dx^1 - idx^2) \big) \\ &= (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \end{aligned}$$

We can equally well do the same thing for momenta P^μ. Hence for null momenta (i.e. for massless particles) we have:

$$P^\mu \to \lambda^\alpha \tilde{\lambda}^{\dot{\alpha}}$$

One notes that P is invariant under the rescaling $\lambda^\alpha \mapsto t\lambda^\alpha$, $\tilde{\lambda}^{\dot{\alpha}} \mapsto t^{-1}\tilde{\lambda}^{\dot{\alpha}} \quad \forall\, t \in \mathbb{C}^*$. This is precisely the freedom necessary for amplitudes written in spinor-helicity variables to transform appropriately under helicity. I.e. for an n-point (stripped) amplitude $\mathcal{A}_{h_1,\ldots,h_n}(\lambda_a^\alpha, \tilde{\lambda}_a^{\dot{\alpha}})$ with massless external particles carrying helicity labels h_a:

$$\mathcal{A}_{h_1,\ldots,h_n}(\lambda_a^\alpha \mapsto t_a\lambda_a^\alpha, \tilde{\lambda}_a^{\dot{\alpha}} \mapsto t_a^{-1}\tilde{\lambda}_a^{\dot{\alpha}}) \to t_1^{-2h_1} \ldots t_n^{-2h_n} \mathcal{A}_{h_1,\ldots,h_n}(\lambda_a^\alpha, \tilde{\lambda}_a^{\dot{\alpha}}) \qquad [12]$$

Noting that the antisymmetric epsilon symbol is the only $SL(2,\mathbb{C})$ invariant, we shall follow [12] and use angular brackets to denote the contraction of undotted spinors:

$$\langle \lambda_1^\alpha \lambda_2^\beta \rangle := \lambda_1^\alpha \lambda_2^\beta \epsilon_{\alpha\beta} =: \langle 12 \rangle$$

and rectangular brackets for dotted spinors:

$$[\tilde{\lambda}_1^{\dot{\alpha}} \tilde{\lambda}_2^{\dot{\beta}}] := \tilde{\lambda}_1^{\dot{\alpha}} \tilde{\lambda}_2^{\dot{\beta}} \epsilon_{\dot{\alpha}\dot{\beta}} =: [12]$$

We can readily see that the scalar product becomes:

$$P_1 \cdot P_2 \to \langle 12 \rangle [12]$$

which isn't weighted under helicity transformations.

By considering nothing but the following:

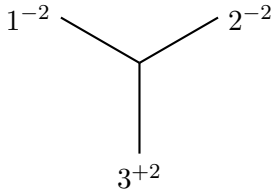
1. The S-matrix should not be weighted under helicity $\lambda^\alpha \mapsto t\lambda^\alpha$, $\tilde{\lambda}^{\dot{\alpha}} \mapsto t^{-1}\tilde{\lambda}^{\dot{\alpha}}$.

2. The S-matrix should respect locality and so the amplitude ought only have simple poles.

3. The S-matrix should respect unitarity and so the amplitude should factorise on poles.

One can readily prove charge conservation for massless spin 1 (photon, gluon) scattering, the equivalence principle for massless spin 2 (graviton) scattering and inconsistency of higher spin scattering.

To see the power of this formalism, before even we arrive at twistors, we present the graviton 3 point MHV (maximally helicity violating) amplitude first arduously calculated by de-Whitt [13] but realised here by simple consideration of helicity weighting:



$$= \frac{1}{M_{pl}} \left(\frac{\langle 12 \rangle^3}{\langle 13 \rangle \langle 23 \rangle} \right)^2 \qquad [12]$$

4.2 Twistor Space

We have several times eluded to this twistor correspondence, so what is it? There are actually multiple ways to get at the geometry here. Perhaps a little mystifying to begin with, we start with a double fibration. This has the drawback of initially appearing terribly co-ordinate dependent, but conceptually the learning curve is the shallowest.

4.2.1 The Double Fibration

We shall first build twistor space $\mathbb{PT}$ as the right hand base of a two sided fibration, the left hand base of which we take to be our complexified spacetime $\mathbb{C}^4$.

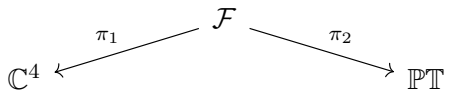


Figure 2: The double fibration of the correspondence space over complex spacetime and twistor space.

We shall equip our spacetime with spinor-helicity co-ordinates $x^{\alpha\dot{\alpha}}$. The correspondence space, $\mathcal{F}$, is defined by an equation called the incidence relation:

$$\mathcal{F} = \left\{ (x^{\alpha\dot{\alpha}}, [\mu^{\dot{\alpha}} : \lambda_\alpha]) \subset \mathbb{C}^4 \times \mathbb{CP}^3 \mid \mu^{\dot{\alpha}} = x^{\alpha\dot{\alpha}} \lambda_\alpha \ \ \& \ \ (\lambda_1, \lambda_2) \neq (0,0) \right\} \tag{4.1}$$

We see $\mathcal{F}$ is to be understood as the graph of the 'incidence relation':

$$\mu^{\dot{\alpha}} = x^{\alpha\dot{\alpha}} \lambda_\alpha \tag{4.2}$$

In fact $\mathcal{F}$ is the smooth sub-variety of $\mathbb{C}^4 \times \mathbb{CP}^3$ carved out by the incidence relations.

We have parameterised the space $\mathbb{CP}^3 \backslash \mathbb{CP}^1$ by homogenous co-ordinates:

$$Z = [\mu^{\dot{\alpha}} : \lambda_\alpha] = [\mu^{\dot{1}} : \mu^{\dot{2}} : \lambda_1 : \lambda_2] \qquad (\lambda_1, \lambda_2) \neq (0,0)$$

The projection maps are then just projection onto the first and second factors:

$$\pi_1 : \mathcal{F} \to \mathbb{C}^4 \qquad \pi_2 : \mathcal{F} \to \mathbb{PT}$$
$$(x^{\alpha\dot{\alpha}}, [\mu^{\dot{\alpha}} : \lambda_\alpha]) \mapsto x^{\alpha\dot{\alpha}} \qquad (x^{\alpha\dot{\alpha}}, [\mu^{\dot{\alpha}} : \lambda_\alpha]) \mapsto [\mu^{\dot{\alpha}} : \lambda_\alpha]$$

The incidence relation has solutions for all $Z \in \mathbb{CP}^3 \backslash \mathbb{CP}^1$. In other words twistor space is simply $\mathbb{PT} = \mathbb{CP}^3 \backslash \mathbb{CP}^1$.

To explore the nature of this non-local correspondence between complexified spacetime and complex projective space, we ought to fix a point $x \in \mathbb{C}^4$ and see what we get in twistor space.

If we fix x and let $[\lambda_1 : \lambda_2]$ vary, $\mu^{\dot{\alpha}}$ is fixed by the incidence relation. Hence we see that for each point in spacetime, we get a Riemann sphere $\mathbb{CP}^1$ in twistor space. We shall refer to such complex projective lines as 'twistor lines', and denote the twistor line corresponding to x by L_x. The linear, holomorphic incidence relation guarantees that each twistor line $L_x \subset \mathbb{PT}$ is linearly and holomorphically embedded.

On the other hand, if we fix a twistor $Z = [\mu^{\dot{\alpha}} : \lambda_\alpha]$, we get a totally null complex 2-plane in $\mathbb{C}^4$, known as an alpha plane. To see this consider two points $x, y \in \mathbb{C}^4$ whose corresponding projective lines in $\mathbb{PT}$ intersect. Since the twistor lines are linearly, holomorphically embedded they must intersect at a unique twistor, or not at all[5]. We shall consider the case $L_x \cap L_y \neq \emptyset$. Suppose then that $L_x \cap L_y = [\mu^{\dot{\alpha}} : \lambda_\alpha]$. From the incidence relations we see:

$$\mu^{\dot{\alpha}} = x^{\alpha\dot{\alpha}}\lambda_\alpha \qquad \mu^{\dot{\alpha}} = y^{\alpha\dot{\alpha}}\lambda_\alpha$$
$$\implies 0 = (x-y)^{\alpha\dot{\alpha}}\lambda_\alpha = \epsilon^{\alpha\beta}(x-y)_\beta{}^{\dot{\alpha}}\lambda_\alpha$$

From this we can infer $(x-y)_\beta{}^{\dot{\alpha}} \propto \lambda_\beta$ and hence $\exists \tilde{\lambda}^{\dot{\alpha}} : (x-y)^{\alpha\dot{\alpha}} = \lambda^\alpha \tilde{\lambda}^{\dot{\alpha}}$.

$x - y$ has rank 1, and so x and y are null separated. We get the entire null complex 2-plane paramaterised by $\tilde{\lambda}^{\dot{\alpha}}$. Any pair x, y on this plane have corresponding projective lines in $\mathbb{PT}$ that uniquely intersect at $[\mu^{\dot{\alpha}} : \lambda_\alpha]$.

We can interpret $\mathbb{PT}$ as providing a Riemann sphere's worth of complex structure for $T_x\mathbb{C}^4$. This is expressed in the following proposition:

Proposition 4.1. At each point $x \in \mathbb{C}^4$ the corresponding projective line $L_x \subset \mathbb{PT}$ gives a Riemann sphere's worth of complex structures for the tangent space $T_x\mathbb{C}^4$

Proof of Proposition 4.1. Fix a twistor $Z = [\mu^{\dot{\alpha}} : \lambda_\alpha] \in \mathbb{PT}$ such that it's alpha plane $\aleph \subset \mathbb{C}^4$ contains the curve $x(t)$. Suppose further we have $V^{\alpha\dot{\alpha}} \in T_x\mathbb{C}^4$ where $x = x(0)$.

$$\mu^{\dot{\alpha}} = x^{\alpha\dot{\alpha}}(t)\lambda_\alpha$$
$$0 = V^{\alpha\dot{\alpha}}\lambda_\alpha$$
$$\implies T_x\aleph = \{V^{\alpha\dot{\alpha}} \mid V^{\alpha\dot{\alpha}}\lambda_\alpha = 0\}$$

[5]One nice way to appreciate this fact, is to consider two Riemann spheres $\mathbb{CP}^1_a$ and $\mathbb{CP}^1_b$ living in some ambient space. Suppose their intersection $\mathbb{CP}^1_a \cap \mathbb{CP}^1_b = S^1$ is a real circle. Then this intersection is not a complex manifold, and the Riemann spheres were not holomorphically embedded in the ambient space after all.

Where we differentiated with respect to the curve parameter t and evaluated at $t = 0$. We now simply choose to define the complex structure at x by:

$$T_x^{1,0}\mathbb{C}^4 \coloneqq T_x\aleph$$

□

4.2.2 Twistor Space as Serre Twisting Sheaves:

Another way to get a feel for the geometry is to view twistor space as the following bundle:

$$\mathbb{PT} \cong \mathcal{O}(1) \oplus \mathcal{O}(1) \to \mathbb{CP}^1$$

Because of the fact $\lambda_\alpha \neq 0$ in $Z = [\mu^{\dot{\alpha}} : \lambda_\alpha]$, we have the projection:

$$\begin{aligned} \mathbb{PT} &\to \mathbb{CP}^1 \\ [\mu^{\dot{\alpha}} : \lambda_\alpha] &\mapsto [\lambda_\alpha] \end{aligned}$$

As such we conclude we can build twistor space as a bundle over a Riemann sphere, but what are the fibres? Homogenous co-ordinates satisfy $[\mu^{\dot{\alpha}} : \lambda_\alpha] = [k\mu^{\dot{\alpha}} : k\lambda_\alpha]$. On the standard co-ordinate patches on $\mathbb{CP}^1$ we have:

$$\underbrace{Z = \left(\frac{\mu^{\dot{1}}}{\lambda_1}, \frac{\mu^{\dot{2}}}{\lambda_1}, z\right)}_{z=\frac{\lambda_2}{\lambda_1}} \qquad \underbrace{Z = \left(\frac{\mu^{\dot{1}}}{\lambda_2}, \frac{\mu^{\dot{2}}}{\lambda_2}, w\right)}_{w=\frac{\lambda_1}{\lambda_2}}$$

With $w = \frac{1}{z}$ on the overlap of the charts. As such when passing between co-ordinates we have

$$\begin{aligned} z &\mapsto w = \frac{1}{z} \\ \left(\frac{\mu^{\dot{1}}}{\lambda_1}, \frac{\mu^{\dot{2}}}{\lambda_1}, z\right) &\mapsto \left(\frac{\mu^{\dot{1}}}{\lambda_2}, \frac{\mu^{\dot{2}}}{\lambda_2}, w\right) = \left(\frac{1}{z}\frac{\mu^{\dot{1}}}{\lambda_1}, \frac{1}{z}\frac{\mu^{\dot{2}}}{\lambda_1}, \frac{1}{z}\right) \end{aligned}$$

Which is to say $\mu^{\dot{1}}$ and $\mu^{\dot{2}}$ give sections of $\mathcal{O}(1)$.

It will be useful to move between homogenous co-ordinates $Z = [\mu^{\dot{1}} : \mu^{\dot{2}} : \lambda_1 : \lambda_2]$ and co-ordinates on the standard patch $U_1 = \{[\lambda_1 : \lambda_2] \mid \lambda_1 \neq 0\}$ on the base $\mathbb{CP}^1$. On this patch we have twistor co-ordinates:

$$(V^{\dot{1}}, V^{\dot{2}}, z) \coloneqq \left(\frac{\mu^{\dot{1}}}{\lambda_1}, \frac{\mu^{\dot{2}}}{\lambda_1}, \frac{\lambda_2}{\lambda_1}\right)$$

In these co-ordinates we have the incidence relation:

$$V^{\dot{\alpha}} = x^{1\dot{\alpha}} + zx^{2\dot{\alpha}} \tag{4.3}$$

We shall freely move between homogenous co-ordinates $(x^{\alpha\dot{\alpha}}, [\mu^{\dot{\alpha}} : \lambda_\alpha])$ on $\mathcal{F}$ and non-projective co-ordinates $(x^{\alpha\dot{\alpha}}, V^{\dot{\alpha}}, z)$ whenever convenient.

4.3 The Ward Correspondence

The Ward correspondence is essential in the ADHM construction. Ward [14] showed that we can construct anti-self dual bundles[6] on spacetime by first choosing a holomorphic bundle on twistor space, which is trivial on each embedded Riemann sphere. In this section we endeavour to convey this elegant result. We have closely followed the lectures by Costello [15].

4.3.1 Anti-Self Duality in Spinor-Helicity Variables

If we hope to draw a connection between anti-self dual bundles on spacetime and holomorphic data on twistor space, then first we need to be able to express curvature, and more generally any 2-form on $\mathbb{C}^4$, in spinor-helicity variables.

A generic 2 form F on $\mathbb{C}^4$ can be written in spinor-helicity variables as follows:

$$F = \frac{1}{2} F_{\alpha\dot{\alpha}\beta\dot{\beta}}\, dx^{\alpha\dot{\alpha}} \wedge dx^{\beta\dot{\beta}}$$

For $F_{\alpha\dot{\alpha}\beta\dot{\beta}}$ to be anti-symmetric under exchange of $\alpha\dot{\alpha} \leftrightarrow \beta\dot{\beta}$, either $\alpha \leftrightarrow \beta$ is symmetric and $\dot{\alpha} \leftrightarrow \dot{\beta}$ anti-symmetric, or vice versa. Using the fact that the only antisymmetric object in two dimensions is ϵ up to a factor, it's clear we have:

$$F_{\alpha\dot{\alpha}\beta\dot{\beta}} = \epsilon_{\alpha\beta} F_{\dot{\alpha}\dot{\beta}} + \epsilon_{\dot{\alpha}\dot{\beta}} F_{\alpha\beta} \tag{4.4}$$

Where $F_{\dot{\alpha}\dot{\beta}}$ and $F_{\alpha\beta}$ are symmetric.

In fact, the spinor-helicity variables not only make the incidence relations manifestly simpler, but have the added advantage of making explicit the self, anti-self dual decomposition. This is expressed in the following proposition:

Proposition 4.2. $\epsilon_{\dot{\alpha}\dot{\beta}} F_{\alpha\beta}$ and $\epsilon_{\alpha\beta} F_{\dot{\alpha}\dot{\beta}}$ are the self and anti-self dual components of F respectively.

Proof of Proposition 4.2.

$$\begin{aligned}
(\star F)^{\alpha\dot{\alpha}\beta\dot{\beta}} &= \frac{1}{2}(\epsilon^{\alpha\delta}\epsilon^{\beta\gamma}\epsilon^{\dot{\alpha}\dot{\gamma}}\epsilon^{\dot{\beta}\dot{\delta}} - \epsilon^{\alpha\gamma}\epsilon^{\beta\delta}\epsilon^{\dot{\alpha}\dot{\delta}}\epsilon^{\dot{\beta}\dot{\gamma}}) F_{\gamma\dot{\gamma}\delta\dot{\delta}} \\
&= \frac{1}{2}(\epsilon^{\alpha\delta}\epsilon^{\beta\gamma}\epsilon^{\dot{\alpha}\dot{\gamma}}\epsilon^{\dot{\beta}\dot{\delta}} - \epsilon^{\alpha\gamma}\epsilon^{\beta\delta}\epsilon^{\dot{\alpha}\dot{\delta}}\epsilon^{\dot{\beta}\dot{\gamma}})(\epsilon_{\gamma\delta} F_{\dot{\gamma}\dot{\delta}} + \epsilon_{\dot{\gamma}\dot{\delta}} F_{\gamma\delta}) \\
&= \epsilon^{\dot{\alpha}\dot{\beta}} F^{\alpha\beta} - \epsilon^{\alpha\beta} F^{\dot{\alpha}\dot{\beta}}
\end{aligned}$$

Hence $\epsilon_{\dot{\alpha}\dot{\beta}} F_{\alpha\beta} = 0 \iff F = -\star F$ and $\epsilon_{\alpha\beta} F_{\dot{\alpha}\dot{\beta}} = 0 \iff F = \star F$ □

So we see that the self dual component is purely positive helicity, and the anti-self dual part of the curvature is purely negative helicity.

4.3.2 The Zero Curvature Representation

We shall see that anti-self dual connections have vanishing curvature on null planes. This is a useful way to determine anti-self duality. First, we establish the following complex-2 planes as null:

[6] Note that in Ward's paper, he uses the opposite convention for self/anti-self duality.

Proposition 4.3. The complex 2-planes in $\mathbb{C}^4$ given by the constraints:

$$\lambda_\alpha x^{\alpha\dot{\alpha}} = 0 \tag{4.5}$$

$$\mu_{\dot{\alpha}} x^{\alpha\dot{\alpha}} = 0 \tag{4.6}$$

for some constant spinors λ_α or $\mu_{\dot{\alpha}}$ are null.

Proof of Proposition 4.3. We will show that the plane 4.5 is null. The argument for 4.6 is almost identical. The only way for $x^{\alpha\dot{\alpha}}$ to contract with the constant non-zero spinor λ_α to give zero is if x is proportional to that spinor:

$$x^{\alpha\dot{\alpha}} = \epsilon^{\alpha\beta}\lambda_\beta \mu^{\dot{\alpha}}$$

$\mu^{\dot{1}}, \mu^{\dot{2}}$ are co-ordinates on the plane. Hence we have:

$$[\Delta\mu\Delta\mu] = \epsilon^{\dot{\alpha}\dot{\beta}}\Delta\mu_{\dot{\alpha}}\Delta\mu_{\dot{\beta}} = 0$$

□

Planes of the form 4.5 are referred to as alpha planes in the literature, and planes of the form 4.6 are called beta planes. We would like to say alpha planes are entirely null self dual complex 2-planes; and beta planes entirely null anti-self dual complex 2-planes. The sense in which this is meant is made precise in the following proposition:

Proposition 4.4. Only the self dual component of a generic 2-form $F\,|_{\alpha-plane}$ restricted to an α-plane is non-zero. In particular then, if the curvature $\mathcal{F}$ of some connection $\mathcal{A}$ vanishes on every α-plane 4.5, the connection $\mathcal{A}$ is anti-self dual.

Proof of Proposition 4.4.

$$\begin{aligned}
\mathcal{F} &= \epsilon_{\alpha\beta}\mathcal{F}_{\dot{\alpha}\dot{\beta}}\, dx^{\alpha\dot{\alpha}} \wedge dx^{\beta\dot{\beta}} + \epsilon_{\dot{\alpha}\dot{\beta}}\mathcal{F}_{\alpha\beta}\, dx^{\alpha\dot{\alpha}} \wedge dx^{\beta\dot{\beta}} \\
&= \epsilon_{\alpha\beta}\mathcal{F}_{\dot{\alpha}\dot{\beta}}\, d(\epsilon^{\alpha\gamma}\lambda_\gamma\mu^{\dot{\alpha}}) \wedge d(\epsilon^{\beta\sigma}\lambda_\sigma\mu^{\dot{\beta}}) + \epsilon_{\dot{\alpha}\dot{\beta}}\mathcal{F}_{\alpha\beta}\, d(\epsilon^{\alpha\gamma}\lambda_\gamma\mu^{\dot{\alpha}}) \wedge d(\epsilon^{\beta\sigma}\lambda_\sigma\mu^{\dot{\beta}}) \\
&= \underbrace{\epsilon_{\alpha\beta}\mathcal{F}_{\dot{\alpha}\dot{\beta}}\, \epsilon^{\alpha\gamma}\lambda_\gamma\epsilon^{\beta\sigma}\lambda_\sigma\, d\mu^{\dot{\alpha}} \wedge d\mu^{\dot{\beta}}}_{0} + \epsilon_{\dot{\alpha}\dot{\beta}}\mathcal{F}_{\alpha\beta}\, \epsilon^{\alpha\gamma}\lambda_\gamma\epsilon^{\beta\sigma}\lambda_\sigma\, d\mu^{\dot{\alpha}} \wedge d\mu^{\dot{\beta}}
\end{aligned}$$

Hence if $\mathcal{F}$ vanishes on every such α-plane we have:

$$\mathcal{F}_{\alpha\beta}\lambda^\alpha\lambda^\beta = 0 \;\forall\; \lambda_\alpha \implies \mathcal{F}_{\alpha\beta} = 0$$

□

Of-course an equivalent proposition can be proven in the same way to show that self dual connections have vanishing curvature on every beta plane.

4.3.3 Ward's Theorem

We are now ready to state the following theorem [14] by Ward.

Theorem 4.2 (Ward)**.** The set $\mathscr{E}_n$ of rank n holomorphic vector bundles on twistor space that are trivial on every linearly embedded $\mathbb{CP}^1 \subset \mathbb{PT}$ is isomorphic to the set $\mathscr{A}_n$ of anti-self dual $GL(n,\mathbb{C})$ bundles on $\mathbb{C}^4$:

$$\mathscr{E}_n \cong \mathscr{A}_n$$

4.3.4 The Ward Construction

Starting with a holomorphic bundle E on twistor space, trivial on each twistor line, we'd like to construct an anti-self dual bundle on $\mathbb{C}^4$.

Consider a rank r holomorphic vector bundle $E \xrightarrow{\pi} \mathbb{PT}$ on twistor space, that is trivial on every line $\mathbb{CP}^1 \subset \mathbb{PT}$.

Then $E \mid_{L_x} \cong L_x \times \mathbb{C}^r$ and on L_x the covariant derivative is just $\bar{\partial}_A = \bar{\partial}$. A holomorphic section is simply r holomorphic functions on $\mathbb{CP}^1$. But by Liouville's theorem (see [6]) these functions are constant, since the Riemann sphere is compact. Hence we have the simple result:

$$H^0(L_x, E \mid_{L_x}) = \mathbb{C}^r$$

It is worth highlighting here that we are looking to build a rank r bundle on spacetime whose fibres are then simply $\mathbb{C}^r$, hence we have a candidate.

Note that $H^0(L_x, E \mid_{L_x}) = \ker(\bar{\partial}_A : \Omega^{0,0}(X, E) \to \Omega^{0,1}(X, E))$. We defined a 'holomorphic' section S as one for which $\bar{\partial}_A S = 0$ so indeed $H^0(L_x, E \mid_{L_x})$ is good notation for the space of holomorphic sections $L_x \to E$.

We can build a rank r vector bundle on $\mathbb{C}^4$ via the map $x^{\alpha\dot{\alpha}} \mapsto H^0(L_x, E \mid_{L_x})$. In other words the bundle is given by the disjoint union:

$$\tilde{E} = \bigsqcup_{x \in \mathbb{C}^4} H^0(L_x, E \mid_{L_x})$$

We must still show that the bundle $\tilde{E}$ is anti-self dual.

4.3.5 Holonomy

In order to endow the bundle $\tilde{E} \to \mathbb{C}^4$ with a connection, for a small parameter ϵ, we need isomorphisms[7] of the form:

$$H^0(L_x, E \mid_{L_x}) \to H^0(L_{x+\epsilon y}, E \mid_{L_{x+\epsilon y}})/\epsilon^2$$

Given such maps $H^0(L_{x_i}, E \mid_{L_{x_i}}) \to H^0(L_{x_{i+1}}, E \mid_{L_{x_{i+1}}})$ for ordered points $x_1, \ldots, x_l$, we can determine the holonomy around the closed path given by the convex hull of these points:

$$Hol_\gamma = \mathcal{P} \exp \int_\gamma \mathcal{A}$$

To make explicit the relationship between connections and these path ordered loops, we shall look at holonomy around a square:

$$Hol_{\mathfrak{sq}} = \mathcal{P} \exp \int_{\mathfrak{sq}} \mathcal{A}$$

We will choose the square to be in the x^i, x^j plane: $\mathfrak{sq} = conv\{(0,0), (\epsilon, 0), (\epsilon, \epsilon), (0, \epsilon)\}$.

[7] If we are being careful here, $x \in \mathbb{C}^4$ and $y \in T_x\mathbb{C}^4$ so the addition $x + \epsilon y$ is then well-defined on $\mathbb{C}^4$. We will want to conformally compactify spacetime later however, so it is more proper to think of $x + \epsilon y$ as the exponential map $\exp_x(\epsilon y)$ modulo ϵ^2.

Proposition 4.5. To order $\mathcal{O}(\epsilon^2)$ the holonomy is given by:

$$Hol_{\mathfrak{sq}} = \mathbf{1} + \epsilon^2 \mathcal{F}_{ij}(0) + \mathcal{O}(\epsilon^3)$$

In order to prove this claim we need a couple of short lemmas:

Lemma 4.1. The following two equations hold in a perturbative expansion in ϵ:

$$\int_0^\epsilon \mathcal{A}_i(x^i, x^j) dx^i = \epsilon \mathcal{A}_i(0, x^j) + \frac{1}{2}\epsilon^2 \partial_i \mathcal{A}_i(0, x^j) + \mathcal{O}(\epsilon^3) \tag{4.7}$$

$$\int_0^\epsilon \mathcal{A}_j(x^i, x^j) dx^j = \epsilon \mathcal{A}_j(x^i, 0) + \frac{1}{2}\epsilon^2 \partial_j \mathcal{A}_j(x^i, 0) + \mathcal{O}(\epsilon^3) \tag{4.8}$$

where i is *not* summed over, and $\mathcal{A}_k(x^i, x^j)$ is shorthand for $\mathcal{A}_k(0, \ldots, x^i, \ldots, x^j, \ldots, 0)$

Proof of Lemma 4.1. For equation 4.7 we may use the fact that $x^i \in [0, \epsilon]$ and Taylor expand:

$$\int_0^\epsilon \mathcal{A}_i(x^i, x^j) dx^i = \int_0^\epsilon \mathcal{A}_i(o, x^j) dx^i + \sum_{n=1}^{\infty} \int_0^\epsilon \frac{(x^i)^n}{n!} \partial_i^n \mathcal{A}_i(0, x^j)$$

But integrating by parts one sees for $n \geq 1$:

$$\int_0^\epsilon \frac{(x^i)^n}{n!} \partial_i^n \mathcal{A}_i(0, x^j) = -\int_0^\epsilon \frac{(x^i)^{n-1}}{(n-1)!} \partial_i^{n-1} \mathcal{A}_i(0, x^j) + \frac{1}{n!}\epsilon^n \partial_i^{n-1} \mathcal{A}_i(0, x^j)$$

Hence we have the result 4.7. The proof of 4.8 is identical. □

Lemma 4.2. The pieces of the quadratic term:

$$\int_{x<y \, on \, \mathfrak{sq}} \mathcal{A}(x) \wedge \mathcal{A}(y)$$

in $\mathcal{P} \exp \int_{\mathfrak{sq}} \mathcal{A}$ where x and y lie on the same side or opposite sides of the square, vanish.

Proof of Lemma 4.2. Schematically, such terms either look like:

$$\int \mathcal{A}_i \mathcal{A}_i dx^i \wedge dx^i \qquad \text{or} \qquad \int \mathcal{A}_j \mathcal{A}_j dx^j \wedge dx^j$$

Where the $i's$ and $j's$ are not summed. Consequently in either case they vanish by the antisymmetry of the wedge product. □

Proof of Proposition 4.5.

$$\mathcal{P} \exp \int_{\mathfrak{sq}} \mathcal{A} = \mathbf{1} + \int_{\mathfrak{sq}} \mathcal{A} + \int_{x<y \, on \, \mathfrak{sq}} \mathcal{A}(x) \wedge \mathcal{A}(y) + \mathcal{O}(\epsilon^3)$$

First we look at the linear term:

$$\begin{aligned}\int_{\mathfrak{sq}} \mathcal{A} &= \int_0^\epsilon A_i(x^i, 0) dx^i + \int_0^\epsilon \mathcal{A}_j(\epsilon, x^j) dx^j - \int_0^\epsilon \mathcal{A}_i(x^i, \epsilon) dx^i - \int_0^\epsilon \mathcal{A}_j(0, x^j) dx^j \\ &= \epsilon(\mathcal{A}_i(0,0) + \mathcal{A}_j(\epsilon, 0) - \mathcal{A}_i(0, \epsilon) - \mathcal{A}_j(0,0)) + \frac{1}{2}\epsilon^2(\partial_i \mathcal{A}_i(0,0) + \partial_j \mathcal{A}_j(\epsilon, 0) - \partial_i \mathcal{A}_i(0, \epsilon) - \partial_j \mathcal{A}_j(0,0))\end{aligned}$$

Where we have used equations 4.7 and 4.8 from lemma 4.1. We may now expand to second order in ϵ:

$$\begin{aligned}\int_{\mathfrak{sq}} \mathcal{A} &= \frac{1}{2}\epsilon^2 \left\{\partial_i \mathcal{A}_j(0,0) - \partial_j \mathcal{A}_i(0,0) + \partial_i \mathcal{A}_i(0,0) + \partial_j \mathcal{A}_j(0,0) - \partial_i \mathcal{A}_i(0,0) - \partial_j \mathcal{A}_j(0,0)\right\} \\ &= \partial_i \mathcal{A}_j(0,0) - \partial_j \mathcal{A}_i(0,0)\end{aligned}$$

Next we look at the quadratic term. Utilising lemma 4.2 we have:

$$\begin{aligned}\int_{x<y\, on\, \mathfrak{sq}} \mathcal{A}(x) \wedge \mathcal{A}(y) &= \int_0^\epsilon \mathcal{A}_i(x^i,0)\mathcal{A}_j(\epsilon,x^j)dx^i \wedge dx^j - \int_0^\epsilon \mathcal{A}_i(x^i,0)\mathcal{A}_j(0,x^j)dx^i \wedge dx^j \\ &\quad - \int_0^\epsilon \mathcal{A}_j(\epsilon,x^j)\mathcal{A}_i(x^i,\epsilon)dx^i \wedge dx^j + \int_0^\epsilon \mathcal{A}_i(x^i,\epsilon)\mathcal{A}_j(0,x^j)dx^i \wedge dx^j \\ &= \epsilon^2 \left\{\mathcal{A}_i(0,0)\mathcal{A}_j(\epsilon,0) - \mathcal{A}_i(0,0)\mathcal{A}_j(0,0) - \mathcal{A}_j(\epsilon,0)\mathcal{A}_i(0,\epsilon) + \mathcal{A}_i(0,\epsilon)\mathcal{A}_j(0,0)\right\} + \mathcal{O}(\epsilon^3) \\ &= \epsilon^2 \left\{\mathcal{A}_i(0,0)\mathcal{A}_j(0,0) - \mathcal{A}_j(0,0)\mathcal{A}_i(0,0)\right\} + \mathcal{O}(\epsilon^3)\end{aligned}$$

In the last line we expanded the expression in powers of epsilon and kept only the leading terms. Putting it together then we have:

$$\begin{aligned}\mathcal{P}\exp\int_{\mathfrak{sq}} \mathcal{A} &= \mathbf{1} + \epsilon^2 \left\{\partial_i \mathcal{A}_j(0) - \partial_j \mathcal{A}_i(0) + [\mathcal{A}_i(0), \mathcal{A}_j(0)]\right\} \\ &= \mathbf{1} + \epsilon^2 \mathcal{F}_{ij}(0)\end{aligned}$$

□

Therefore if the holonomy of a loop is simply the identity, curvature of the connection vanishes.

The decomposition $T_x\mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$ of the tangent bundle into spin bundles makes it clear that $\mathbb{C}^4$ is spanned by null vectors. As such we only need build our connection from the parallel transport along null paths. But for null separated points $x^{\alpha\dot{\alpha}}$ and $x'^{\alpha\dot{\alpha}} = x^{\alpha\dot{\alpha}} + \epsilon y^{\alpha\dot{\alpha}}$, the corresponding twistor lines L_x and $L_{x'}$ intersect exactly once at the twistor Z.

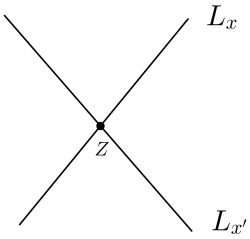


Figure 3: Two twistor lines L_x and $L_{x'}$ intersecting at the twistor Z.

But then since sections of our bundle E restricted to just L_x or $L_{x'}$ are constant, we can simply restrict these constant functions to the intersection point:

$$H^0(L_x, E\mid_{L_x}) \xrightarrow{r_x} H^0(Z, E\mid_Z) \xleftarrow{r_{x'}} H^0(L_{x'}, E\mid_{L_{x'}})$$

By proposition 4.4, in order to show that this connection is anti-self dual we need it's curvature $\mathcal{F}$ to vanish on all alpha planes. Our goal will be to calculate the holonomy of a triangle confined to an alpha plane and show that it's merely the identity. To show this we first we need a small lemma.

Lemma 4.3. If the three twistor lines $L_{x_1}, L_{x_2}, L_{x_3}$ associated to points x_1, x_2, x_3 intersect at a unique twistor $Z = [\mu^{\dot{\alpha}} : \lambda_\alpha]$, the points lie on an alpha plane.

Proof of Lemma 4.3. By translational invariance it suffices to choose $x_3 = 0$ at the origin. Using the incidence relation $\mu^{\dot{\alpha}} = x^{\alpha\dot{\alpha}}\lambda_\alpha$:

$$L_{x_1} \cap L_0 = [\mu^{\dot{\alpha}} : \lambda_\alpha] \implies x_1^{\alpha\dot{\alpha}}\lambda_\alpha = 0$$
$$L_{x_2} \cap L_0 = [\mu^{\dot{\alpha}} : \lambda_\alpha] \implies x_2^{\alpha\dot{\alpha}}\lambda_\alpha = 0$$

Hence both x_1 and x_2 lie on the alpha plane generated by λ_α. □

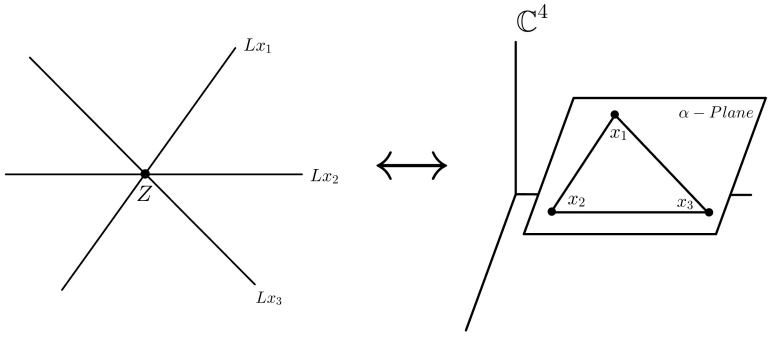


Figure 4: Three points on an alpha plane forming a null triangle in $\mathbb{C}^4$, and their three twistor lines L_{x_1}, L_{x_2} and L_{x_3} intersecting at a unique twistor Z in $\mathbb{PT}$.

Since for three points that lie on an alpha plane we have a triple intersection point in twistor space, we can use our restriction isomorphisms:

$$\begin{array}{ccccc} H^0(L_{x_1}, E|_{L_{x_1}}) & \xrightarrow{\quad r_1 \quad} & H^0(Z, E|_Z) & \xleftarrow{\quad r_2 \quad} & H^0(L_{x_2}, E|_{L_{x_2}}) \\ \downarrow{\scriptstyle r_1} & & & & \downarrow{\scriptstyle r_2} \\ H^0(Z, E|_Z) & \xleftarrow[\quad r_3 \quad]{} & H^0(L_{x_3}, E|_{L_{x_3}}) & \xrightarrow[\quad r_3 \quad]{} & H^0(Z, E|_Z) \end{array}$$

Clearly the parallel transport map that sends $H^0(L_{x_1}, E|_{L_{x_1}}) \to H^0(L_{x_1}, E|_{L_{x_1}})$ is given by:

$$r_1^{-1} \circ r_3 \circ r_3^{-1} \circ r_2 \circ r_2^{-1} \circ r_1 = id$$

Let us summarise what we have so far:

1. We begin with a rank r holomorphic vector bundle $E \to \mathbb{PT}$ trivial on each line $L_x \subset \mathbb{PT}$. Triviality on each line gives us isomorphisms $H^0(L_{x_i}, E\,|_{L_{x_i}}) \to H^0(L_{x_j}, E\,|_{L_{x_j}})$ for any null separated points x_i and x_j, since the lines L_{x_i} and L_{x_j} intersect.

2. We build a $GL(r, \mathbb{C})$ vector bundle $\tilde{E} \to \mathbb{C}^4$ on spacetime by the assignment $x \mapsto H^0(L_x, E\,|_{L_x})$. One can always view this as a $GL(r, \mathbb{C})$ principal bundle by considering the frame bundle.

3. A null polygon confined to an alpha plane has the property that the twistor lines of its vertices all intersect at a unique twistor.

4. This unique intersection gives a trivial identity map $H^0(L_{x_1}, E\mid_{L_{x_1}}) \to \cdots \to H^0(L_{x_i}, E\mid_{L_{x_i}}) \to \cdots \to H^0(L_{x_1}, E\mid_{L_{x_1}})$ and hence holonomy of $\tilde{E}$ is trivial on alpha planes.

5. By the zero curvature representation, holonomy that trivialises on alpha planes arises from an anti-self dual connection.

6. As such the bundle $\tilde{E} \to \mathbb{C}^4$ can be thought of as a $GL(r, \mathbb{C})$ principal bundle with an anti-self dual connection.

One might worry at this point that our bundle is not just anti-self dual but in fact has zero curvature. However the holonomy was only trivial on alpha planes because of the unique intersection twistor. For a general null triangle formed as the convex hull of x_1, x_2, x_3 in $\mathbb{C}^4$, the twistor lines intersect at $L_{x_1} \cap L_{x_2} = Z_3$, $L_{x_2} \cap L_{x_3} = Z_1$, $L_{x_3} \cap L_{x_1} = Z_2$ and we have:

$$\begin{array}{ccccc} H^0(L_{x_1}, E|_{L_{x_1}}) & & \xrightarrow{\quad r_{13} \quad} & & H^0(Z_3, E|_{Z_3}) & & \xleftarrow{\quad r_{23} \quad} & & H^0(L_{x_2}, E|_{L_{x_2}}) \\ \Big\downarrow{\scriptstyle r_{12}} & & & & & & & & \Big\downarrow{\scriptstyle r_{21}} \\ H^0(Z_2, E|_{Z_2}) & & \xleftarrow[\quad r_{32} \quad]{} & & H^0(L_{x_3}, E|_{L_{x_3}}) & & \xrightarrow[\quad r_{31} \quad]{} & & H^0(Z_1, E|_{Z_1}) \end{array}$$

This time the parallel transport map that sends $H^0(L_{x_1}, E|_{L_{x_1}}) \to H^0(L_{x_1}, E|_{L_{x_1}})$ is given by:

$$r_{12}^{-1} \circ r_{32} \circ r_{31}^{-1} \circ r_{21} \circ r_{23}^{-1} \circ r_{13} \neq id$$

5 The ADHM Construction

In the following section we come to the ADHM construction. Again we have closely followed the lectures by Costello [15].

5.1 Building Bundles via Cohomology

Let $E_1 \to X$ and $E_2 \to X$ be rank r_1 and r_2 holomorphic vector bundles. Then $f : E_1 \to E_2$ is a holomorphic bundle map if, in a local frame, f is an $r_2 \times r_1$ matrix of holomorphic functions. If f has maximal rank: $rank(f) = r_2$ then $\ker(f) \subset E_1$ is a holomorphic vector bundle. This essentially follows from the rank-nullity theorem which in this case guarantees the fibres of $\ker(f)$ have a consistent dimension.

Similarly if $E_0 \to X$ and $E_1 \to X$ are rank r_0 and r_1 holomorphic vector bundles and $g : E_0 \to E_1$ is a holomorphic bundle map with $rank(g) = r_0$, then $E_1/\operatorname{im}(g)$ is a holomorphic vector bundle.

Consider the bundles and bundle maps as detailed above such that $f \circ g = 0$. Then we have the following cochain:

$$0 \hookrightarrow E_0 \xrightarrow{g} E_1 \xrightarrow{f} E_2 \to 0$$

and consequently the cohomology:

$$E = \ker(f)/\operatorname{im}(g)$$

Algebraists call a short sequence such as this, with g injective and f surjective as we have; a monad.

Using the fact that $E_0 \cong \mathrm{im}(g)$ and $E_2 \cong \mathrm{im}(f)$ as-well as a simple application of the rank-nullity theorem yields:

$$E_1 = E_0 \oplus E \oplus E_2$$

We see that E is a rank $dim(E_1) - dim(E_0) - dim(E_2)$ holomorphic vector bundle on X.

Letting $k \coloneqq dim(E_1) - dim(E_0) - dim(E_2)$ and taking X to be twistor space, by Ward (4.2) we get a rank k anti-self dual bundle $\tilde{E}$ on spacetime.

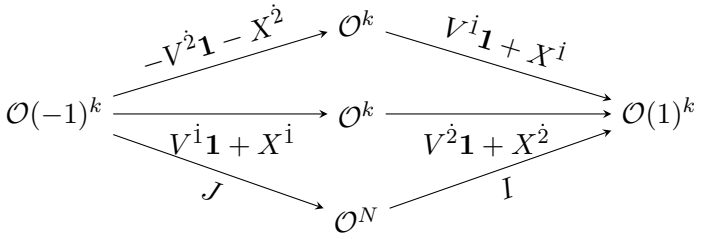
5.2 The General ADHM Construction

We have now seen that given holomorphic bundles E_0, E_1, E_2 on $\mathbb{PT}$ and maps $g : E_0 \to E_1$ and $f : E_1 \to E_2$ with rank $dim(E_0)$ and $dim(E_2)$ respectively; we can take the cohomology $E = \ker(f)/\,\mathrm{im}(g)$ and then by Ward (4.2) get a corresponding anti-self dual bundle on $\mathbb{C}^4$.

5.2.1 The ADHM Monad

The question remains, what bundles and bundle maps should we choose? We shall work in a trivialisation on the standard patch around $z = 0$ on $\mathbb{PT}$ so that specifying f and g is tantamount to giving holomorphic matrices.

The rank N charge k ADHM monad is the following:



Remarks 5.1.

1. Each of the components of g mapping $\mathcal{O}(-1)^k \to \mathcal{O}^k$ are either a $k \times k$ or $k \times N$ matrix worth of sections $s \in H^0(\mathbb{PT}, \mathcal{O}(1))$. Similarly each of the components of f mapping $\mathcal{O}^k \to \mathcal{O}(1)^k$ are either a $k \times k$ or $N \times k$ matrix worth of sections $s \in H^0(\mathbb{PT}, \mathcal{O}(1))$.
2. As discussed in section 2.3 these components of f and g must be at most linear in z.
3. $X^{\dot{\alpha}}$ is a holomorphic function of z, and carries information about the distribution of charge. As indicated above we can write its expansion: $X^{\dot{\alpha}}(z) = X^{1\dot{\alpha}} + zX^{2\dot{\alpha}}$. This data can be then represented by $X^{\alpha\dot{\alpha}}$.
4. I and J are $k \times N$ and $N \times k$ matrices respectively and we can write $I(z) = I^1 + zI^2$ and $J(z) = J^1 + zJ^2$.

We must of course still require that the maps satisfy the rank conditions and ensure $f \circ g = 0$. Furthermore we must yet show that the bundle $\tilde{E}$ is gauge trivial at infinity if we are to ensure the Yang-Mills action is finite.

5.2.2 The Complex ADHM Equation

Proposition 5.1 (ADHM Equation)**.** The ADHM equation:

$$IJ + [X^{\dot{2}}, X^{\dot{1}}] = 0 \tag{5.1}$$

is equivalent to the condition $f \circ g = 0$ for the maps:

$$g = \begin{pmatrix} -V^{\dot{2}}\mathbf{1} - X^{\dot{2}} \\ V^{\dot{1}}\mathbf{1} + X^{\dot{1}} \\ J \end{pmatrix} \qquad \text{and} \qquad f = \begin{pmatrix} V^{\dot{1}}\mathbf{1} + X^{\dot{1}} & V^{\dot{2}}\mathbf{1} + X^{\dot{2}} & I \end{pmatrix}$$

Proof of Proposition 5.1. The proof is a trivial exercise in matrix multiplication. □

It should be noted that the ADHM equation enjoys a $GL(k, \mathbb{C})$ gauge symmetry:

$$\forall\, U \in GL(k, \mathbb{C}) : IJ + [X^{\dot{2}}, X^{\dot{1}}] = 0 \text{ is invariant under:}$$

$$\begin{aligned} X^{\dot{\alpha}} &\mapsto U X^{\dot{\alpha}} U^{-1} \\ I &\mapsto UI \\ J &\mapsto JU^{-1} \end{aligned}$$

Provided I, J, and $X^{\dot{\alpha}}$ satisfy the ADHM equation and are such that f is fibrewise surjective and g fibrewise injective, cohomology of the ADHM monad will yield rank N charge k instantons. We should still ask, is this construction comprehensive? Is it unique?

5.2.3 The ADHM Theorem

Theorem 5.1 (ADHM)**.** Every $GL(N, \mathbb{C})$ instanton on $\mathbb{C}^4$ of charge k, with a gauge at infinity such that $\mathcal{A} = 0$, arises uniquely (up to $GL(k, \mathbb{C})$ transformations) from the ADHM construction.

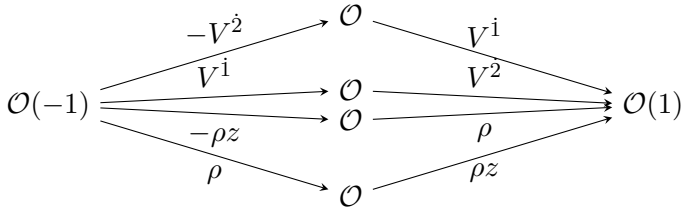
5.3 The Rank Two Charge One Solution

5.3.1 The Rank Two Charge One Monad

The ADHM monad for $N = 2, k = 1$ is:

$$0 \hookrightarrow \mathcal{O}(-1) \xrightarrow{g} \mathcal{O}^4 \xrightarrow{f} \mathcal{O}(1) \to 0$$

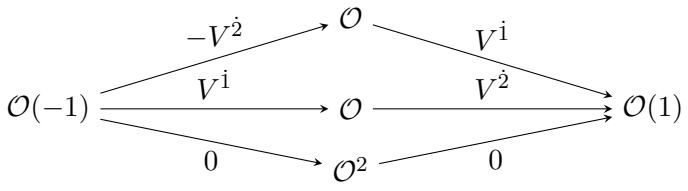
We shall centre our solution on the origin ($X^{\dot{\alpha}} = 0$) and choose I and J to be $\begin{pmatrix} \rho & \rho z \end{pmatrix}$ and $\begin{pmatrix} -\rho z & \rho \end{pmatrix}^T$ respectively. Since $IJ = 0$ the ADHM equation is satisfied. We might better visualise the monad in an expanded form:



Noting we should interpret ρ as a size parameter.

5.3.2 The Zero-Size Limit

Consider the limit $\rho \to 0$. The monad simplifies:



Then for $a, b \in H^0(\mathbb{PT}, \mathcal{O}(1))$:

$$\begin{aligned}
\begin{pmatrix} a \\ b \end{pmatrix} \in \ker \begin{pmatrix} V^{\dot{1}} & V^{\dot{2}} \end{pmatrix} &\implies V^{\dot{1}} a + V^{\dot{2}} b = 0 \\
&\implies a = -\frac{V^{\dot{2}}}{V^{\dot{1}}} b \\
&\implies \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -V^{\dot{2}} \left(\frac{b}{V^{\dot{1}}}\right) \\ V^{\dot{1}} \left(\frac{b}{V^{\dot{1}}}\right) \end{pmatrix} \in \operatorname{im} \begin{pmatrix} -V^{\dot{2}} \\ V^{\dot{1}} \end{pmatrix}
\end{aligned}$$

Hence these maps do not contribute to the cohomology. We then see that:

$$\begin{aligned}
\lim_{\rho \to \infty} E &\cong \ker \left[\begin{pmatrix} V^{\dot{1}} & V^{\dot{2}} & 0 \end{pmatrix} : \mathcal{O}^4 \to \mathcal{O}(1) \right] \Big/ \operatorname{im} \left[\begin{pmatrix} -V^{\dot{2}} \\ V^{\dot{1}} \\ 0 \end{pmatrix} : \mathcal{O}(-1) \to \mathcal{O}^4 \right] \\
&\cong \ker \left[(0) : \mathcal{O}^2 \to \mathcal{O}(1) \right] \Big/ \operatorname{im} \left[(0) : \mathcal{O}(-1) \to \mathcal{O}^2 \right] \\
&\cong \mathcal{O}^2
\end{aligned}$$

So we see that in this limit as the instanton vanishes, the bundle becomes entirely trivial.

5.4 Euclidean Instantons

So far the instantons we have managed to produce have been $GL(N, \mathbb{C})$ bundles on $\mathbb{C}^4$. Now we turn to the job of constructing $U(N)$ instantons on $\mathbb{R}^4$. We shall do this by enforcing the Euclidean reality condition.

Consider the anti-holomorphic map:

$$\begin{aligned}
\sigma : \mathbb{PT} &\to \mathbb{PT} \\
\left(V^{\dot{1}}, V^{\dot{2}}, z \right) &\mapsto \left(\frac{\overline{V}^{\dot{2}}}{\bar{z}}, -\frac{\overline{V}^{\dot{1}}}{\bar{z}}, -\frac{1}{\bar{z}} \right)
\end{aligned} \tag{5.2}$$

We can induce from this an anti-linear map $\sigma : \mathbb{C}^4 \to \mathbb{C}^4$ by requiring $\sigma(L_x) = L_{\sigma(x)}$. Of course one derives such a map from the incidence relations:

$$V^{\dot{1}} = x^{1\dot{1}} + x^{2\dot{1}} z \xrightarrow{\sigma} \frac{\overline{V}^{\dot{2}}}{\bar{z}} = \sigma(x^{1\dot{1}}) - \sigma(x^{2\dot{1}}) \frac{1}{\bar{z}} \implies V^{\dot{2}} = -\overline{\sigma(x^{2\dot{1}})} + \overline{\sigma(x^{1\dot{1}})} z$$

$$V^{\dot{2}} = x^{1\dot{2}} + x^{2\dot{2}} z \xrightarrow{\sigma} -\frac{\overline{V}^{\dot{1}}}{\bar{z}} = \sigma(x^{1\dot{2}}) - \sigma(x^{2\dot{2}}) \frac{1}{\bar{z}} \implies V^{\dot{1}} = \overline{\sigma(x^{2\dot{2}})} - \overline{\sigma(x^{1\dot{2}})} z$$

Hence indeed we preserve the incidence relations if:

$$\begin{gathered}\sigma : \mathbb{C}^4 \to \mathbb{C}^4 \\ \begin{pmatrix} x^{1\dot{1}} & x^{1\dot{2}} \\ x^{2\dot{1}} & x^{2\dot{2}} \end{pmatrix} \mapsto \begin{pmatrix} \overline{x}^{2\dot{2}} & -\overline{x}^{2\dot{1}} \\ -\overline{x}^{1\dot{2}} & \overline{x}^{1\dot{1}} \end{pmatrix}\end{gathered} \tag{5.3}$$

We note that the fixed point of this map:

$$\begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} = \begin{pmatrix} \bar{x}^0 - \bar{x}^3 & -\bar{x}^1 + i\bar{x}^2 \\ -\bar{x}^1 - i\bar{x}^2 & \bar{x}^0 + \bar{x}^3 \end{pmatrix}$$

occurs for x^0 real and x^1, x^2, x^3 imaginary. Since the metric is $(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$ this is exactly the Euclidean slice for which the metric is positive definite.

5.4.1 The Euclidean ADHM Equations

Our job then is to find a condition on the ADHM data that yields a Hermitian structure on the bundle $\tilde{E}$ restricted to the Euclidean slice. We begin with the following proposition:

Proposition 5.2. Holomorphic vector bundles $E \to \mathbb{PT}$ on twistor space (trivial on each $\mathbb{CP}^1 \subset \mathbb{PT}$) satisfying:

$$\overline{\sigma^* E} = E^*$$

for the map σ given by 5.2 induce vector bundles $\tilde{E} \to \mathbb{C}^4$ on spacetime (via Ward) which, when restricted to the Euclidean slice $\{x \in \mathbb{C}^4 \mid \sigma(x) = x\} \cong \mathbb{R}^4$, have a Hermitian structure.

Proof of Proposition 5.2. Suppose the rank N holomorphic vector bundle $E \to \mathbb{PT}$ (trivial on each $\mathbb{CP}^1 \subset \mathbb{PT}$) indeed satisfies $\overline{\sigma^* E} = E^*$. Then restricting to a twistor line L_x:

$$\overline{\sigma^* E} \mid_{L_x} = \overline{E_{L_{\sigma(x)}}} = E^*_{L_x}$$

Then if we further restrict to $x = \sigma(x)$:

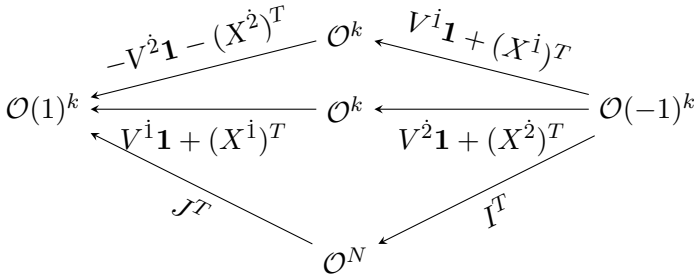
$$\overline{E_{L_x}} = E^*_{L_x}$$

This descends to sections:

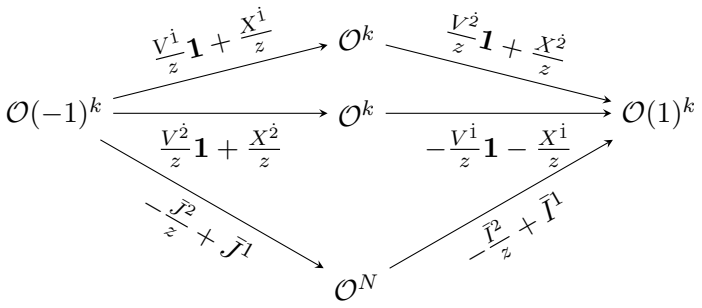
$$\overline{H^0(L_x, E \mid_{L_x})} = H^0(L_x, E \mid_{L_x})^*$$

Therefore by proposition 2.4, the restricted bundle $\tilde{E} \to \mathbb{R}^4$ with typical fibre $\tilde{E}_x = H^0(L_x, E \mid_{L_x})$ has a Hermitian structure. □

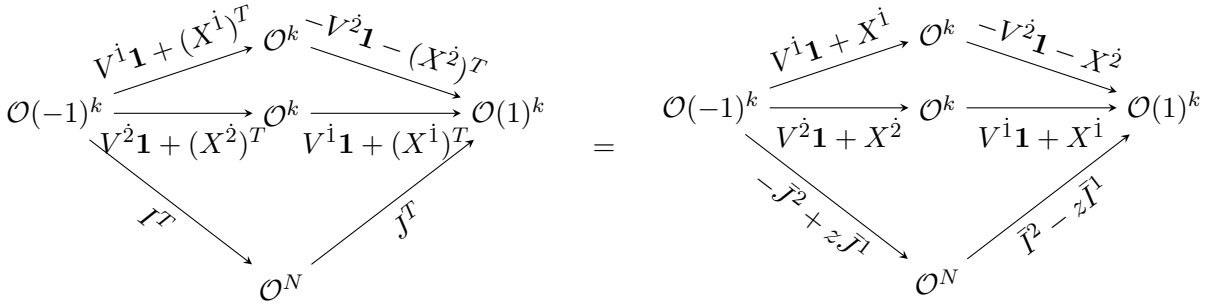
The dual bundle E^* is obtained from the ADHM data simply by choosing the ADHM monad with each bundle replaced by its dual, each map its transpose and with each arrow reversed:



Treating $X^{\alpha\dot{\alpha}}$ as co-ordinates on $\mathbb{C}^4$, we can map them under sigma as in 5.3. The monad for $\overline{\sigma^* E}$ is:



We can change frame (trivialisation) for the bundles $\mathcal{O}(-1)^K$ and $\mathcal{O}(1)^K$ and effectively scale the maps in this monad by $z/-z$. Hence to achieve $U(N)$ instantons on $\mathbb{R}^4$ we require the following equality of the ADHM data:



This equality is guaranteed if we enforce:

$$I^1 = -(J^2)^\dagger \qquad I^2 = (J^1)^\dagger \qquad X^{\dot{\alpha}} = (X^{\dot{\alpha}})^T$$

Hence we can make some redefinitions and remove the synthetic information:

1. Let $I \coloneqq I^1 = -(J^2)^\dagger$
2. Let $J \coloneqq J^1 = (I^2)^\dagger$
3. Let $B_1 \coloneqq X^{1\dot{1}} = \overline{X}^{2\dot{2}} = (X^{2\dot{2}})^\dagger$
4. Let $B_2 \coloneqq -X^{1\dot{2}} = \overline{X}^{2\dot{1}} = (X^{2\dot{1}})^\dagger$

Proposition 5.3 (Euclidean ADHM Equations)**.** The ADHM equation 5.1 in the redefined variables above become:

$$IJ + [B_1, B_2] = 0 \tag{5.4}$$

$$II^\dagger - J^\dagger J + [B_1, B_1^\dagger] + [B_2, B_2^\dagger] = 0 \tag{5.5}$$

$$J^\dagger I^\dagger + [B_2^\dagger, B_1^\dagger] = 0 \tag{5.6}$$

Proof. The ADHM equation is:

$$(I^1 + zI^2)(J^1 + zJ^2) + [(X^{1\dot{2}} + zX^{2\dot{2}}), (X^{1\dot{1}} + zX^{2\dot{1}})]$$

Then each coefficient in the quadratic in z must individually vanish:

$$\begin{aligned} z^0: &\quad I^1 J^1 + [X^{1\dot{2}}, X^{1\dot{1}}] = 0 \longrightarrow IJ + [B_1, B_2] = 0 \\ z^1: &\quad I^2 J^1 + I^1 J^2 + [X^{1\dot{2}}, X^{2\dot{1}}] + [X^{2\dot{2}}, X^{1\dot{1}}] = 0 \longrightarrow II^\dagger - J^\dagger J + [B_1, B_1^\dagger] + [B_2, B_2^\dagger] = 0 \\ z^2: &\quad I^2 J^2 + [X^{2\dot{2}}, X^{2\dot{1}}] = 0 \longrightarrow J^\dagger I^\dagger + [B_2^\dagger, B_1^\dagger] = 0 \end{aligned}$$

□

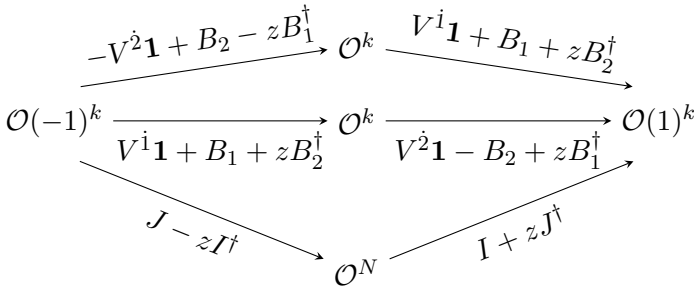
The residual $GL(k, \mathbb{C})$ gauge symmetry in the ADHM data reduces to a $U(k)$ symmetry:

$$\begin{aligned} &\forall\, U \in U(k) : IJ + [B_1, B_2] = 0 \text{ is invariant under:} \\ &\qquad B_\alpha \mapsto U B_\alpha U^{-1} \\ &\qquad I \mapsto UI \\ &\qquad J \mapsto JU^{-1} \end{aligned}$$

5.5 Conformal Compactification

In order for the instantons to have finite action, the connection must approach pure gauge at infinity. We must still show that the ADHM data produce a bundle which is framed (trivialised with a chosen gauge) at infinity. As long as this is the case, we can consistently extend the framed connection over a single point $\{\infty\}$. We are then free to extend the bundle over $\mathbb{R}^4$ to one over S^4.

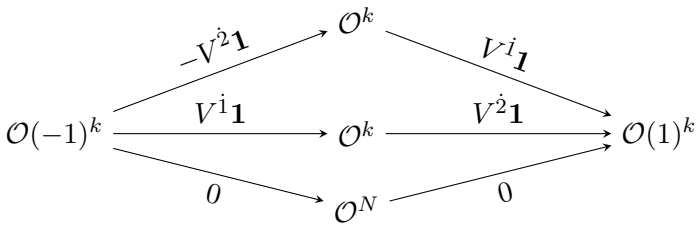
We of course begin with the ADHM monad:



with I, J, B_1, B_2 satisfying the ADHM equations 5.4 and 5.5. Using the incidence relation:

$$V^{\dot{\alpha}} = x^{1\dot{\alpha}} + z x^{2\dot{\alpha}}$$

we can see that in the limit $x^{\alpha\dot{\alpha}} \to \infty$, we also have $V^{\dot{\alpha}} \to \infty$. In this limit the monad simplifies:



Let a and b be $k \times k$ matrices each in $\mathcal{O}^k$. Then entirely analogously to what we did for the $\rho \to 0$ rank two charge one instanton:

$$\begin{aligned}
\begin{pmatrix} a \\ b \end{pmatrix} \in \ker \begin{pmatrix} V^{\dot{1}}\mathbf{1} & V^{\dot{2}}\mathbf{1} \end{pmatrix} &\implies V^{\dot{1}}a + V^{\dot{2}}b = 0 \\
&\implies a = -\frac{V^{\dot{2}}}{V^{\dot{1}}}b \\
&\implies \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -V^{\dot{2}}\left(\frac{b}{V^{\dot{1}}}\right) \\ V^{\dot{1}}\left(\frac{b}{V^{\dot{1}}}\right) \end{pmatrix} \in \operatorname{im} \begin{pmatrix} -V^{\dot{2}}\mathbf{1} \\ V^{\dot{1}}\mathbf{1} \end{pmatrix}
\end{aligned}$$

Hence these maps do not contribute to the cohomology. We then see that:

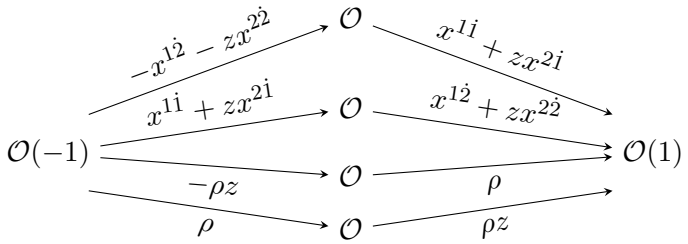
$$\begin{aligned}
E\mid_{L_\infty} &\cong \ker \left[\begin{pmatrix} V^{\dot{1}}\mathbf{1} & V^{\dot{2}}\mathbf{1} & 0 \end{pmatrix} : \mathcal{O}^{2k+N} \to \mathcal{O}(1)^k \right] / \operatorname{im} \left[\begin{pmatrix} -V^{\dot{2}}\mathbf{1} \\ V^{\dot{1}}\mathbf{1} \\ 0 \end{pmatrix} : \mathcal{O}(-1)^k \to \mathcal{O}^{2k+N} \right] \\
&\cong \ker \left[(0) : \mathcal{O}^N \to \mathcal{O}(1)^k \right] / \operatorname{im} \left[(0) : \mathcal{O}(-1)^k \to \mathcal{O}^N \right] \\
&\cong \mathcal{O}^N
\end{aligned}$$

Hence the bundle 'at infinity' is nothing more than the trivial rank N bundle $\mathcal{O}^N$; and consequently the Ward-induced bundle $\tilde{E}\mid_{S^3_\infty}$ is trivial as well.

5.6 The BPST Instanton Revisited

We can now derive the BPST instanton from the ADHM data. We have the rank 2 charge 1 monad as given in section 5.3. The ADHM equations 5.4 and 5.5 are trivially satisfied for this data.

In order to calculate the BPST instanton, we must restrict the monad to a twistor line L_x before calculating the cohomology.



The combination of the map $g : \mathcal{O}(-1) \to \mathcal{O}^4$ being injective, and the fact that $H^0(\mathbb{CP}^1, \mathcal{O}(-1)) = 0$ means that $\operatorname{im}(g) = 0$ and so $E\mid_{L_x} = \ker(f)$. A section of $\mathcal{O}^4$ is simply a constant $\begin{pmatrix} s_1 & s_2 & s_3 & s_4 \end{pmatrix}^T \in \mathbb{C}^4$:

$$E\mid_{L_x} = \left\{ \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} : \begin{pmatrix} x^{1\dot{1}} + zx^{2\dot{1}} & x^{1\dot{1}} + zx^{2\dot{1}} & -\rho z & \rho \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = 0 \ \forall \ z \in L_x \right\}$$

This gives two linear constraints, for z^0 and z^1. This can be repackaged into:

$$E\,|_{L_x} = \ker \Delta^\dagger$$
$$\Delta^\dagger = \begin{pmatrix} x^{1\dot{1}} & x^{1\dot{2}} & \rho & 0 \\ x^{2\dot{1}} & x^{2\dot{2}} & 0 & \rho \end{pmatrix}$$

Where $\Delta^\dagger$ is simply the notation typically used in the literature. We might use the identification $\mathbb{H} \cong \mathbb{R}^4$ used in section 3.3 to write:

$$\Delta^\dagger = \begin{pmatrix} \mathbf{x} & \rho \end{pmatrix}$$

An obvious normalised[8] basis for the kernel of $\Delta^\dagger$ and hence $\tilde{E}_x$ is:

$$W = \frac{1}{\sqrt{\rho^2 + |\mathbf{x}|^2}} \begin{pmatrix} \rho \\ -\mathbf{x} \end{pmatrix}$$

It is now a simple exercise to compute the connection $\mathcal{A} = \mathscr{I}(W^\dagger dW)$ to obtain:

$$\mathcal{A}(\mathbf{x}) = \frac{1}{\rho^2 + |\mathbf{x}|^2} \mathscr{I}(\overline{\mathbf{x}}\mathbf{dx})$$

Choosing $\rho = 1$ recovers the solution we saw in section 3.3.

6 Concluding Remarks

We have examined the ADHM mechanism for generating instantons and, in doing so, characterised the moduli space of $SU(N)$ instantons with integer topological charge k. The construction ultimately rests on the Ward correspondence. Ward showed that there is an equivalence between anti-self dual $GL(N, \mathbb{C})$ bundles on complexified spacetime $\mathbb{C}^4$ and holomorphic rank N vector bundles on twistor space that are trivial on twistor lines. The ADHM system provides a way to classify these holomorphic vector bundles and to systematically relate them to anti-self dual bundles $\tilde{E} \to \mathbb{C}^4$. We then imposed Euclidean reality conditions on spacetime, allowing us to endow $\tilde{E}$ with a Hermitian structure and thereby reduce the structure group $GL(N, \mathbb{C}) \to SU(N)$.

The ADHM data can be stated without reference to twistors, however. In that setting the construction appears somewhat ad hoc. From this perspective, the replacement of the non-linear partial differential equation $\mathcal{F} + \star\mathcal{F} = 0$ by linear data I, J, B_1, B_2 is rather striking. In this essay we have aimed not only to present the ADHM construction, but also, through the lens of the twistor programme, motivate the origin of this linear structure.

An auxiliary aim of this essay was to convey some of the utility of twistors. A striking connection is emerging between algebraic geometry and the study of amplitudes, with twistor theory at its centre. It's now apparent that amplitudes in certain theories can be efficiently computed via geometry in twistor space. For example, Skinner and Mason [10] showed that the full tree amplitude for $\mathcal{N} = 4$ super Yang-Mills can be obtained from Wilson loops in twistor space. Using momentum twistors, Arkani-Hamed and Trynka [16] demonstrated this same amplitude can be computed via the so-called Amplituhedron. For further developments relating Wilson loops and the Amplituhedron, see [17], [18], and [19].

[8] This means we are considering not a U(2) instanton but rather SU(2).

A Appendix: Geometry

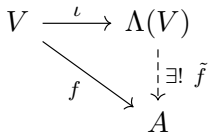
This appendix endeavours to provide a quick but thorough summary of Cartan's language of differential forms; the bundle theoretic machinery of differential geometry and an introduction to complex geometry.

Modern mathematical physics is often written in this language. This is due in part to the simplicity of the formalism. If not simple the reader will surely find it compact. In part the prevalence of these forms is due to their success in unifying apparently distinct physics. Differential forms are a very natural class of objects used to generalise calculus from $\mathbb{R}^n$ to general differentiable manifolds. See [20] and [4] for references of the material.

A.1 Grassmann's Exterior Algebra

Definition A.1. Let V be a linear space over field $\mathbb{F}$. The exterior algebra of V is the pair $(\Lambda(V), \iota)$ where $\Lambda(V)$ is a unital associative algebra over $\mathbb{F}$ and $\iota : V \xrightarrow{\sim} \Lambda(V)$ is a linear map such that:

1. $(\iota(v))^2 = 0 \; \forall \; v \in V$
2. for any unital associative algebra A over $\mathbb{F}$ and linear map $f : V \xrightarrow{\sim} A$ also satisfying $(f(v))^2 = 0 \; \forall \; v \in V$, then $\exists! \tilde{f} : \Lambda(V) \xrightarrow{\sim} A$ extending f.



If the reader is algebraically inclined then this short definition is very clean, and easy to express in a categorical language. For those who prefer a more constructive approach we provide the alternative (still algebraic) definitions.

Definition A.2. The tensor algebra $T(V)$ of a linear space V is given by:

$$T(V) \coloneqq \bigoplus_{k=0}^{\infty} V^{\otimes k}$$

Multiplication in this algebra is given by the tensor product.

The tensor algebra ought to be a familiar construction. It is the starting point for many useful spaces that arise from quotienting a tensor algebra. Some examples are:

1. Fock space:
$$\mathcal{F} = T(\mathcal{H}_n)/\sim$$
where $\mathcal{H}_n$ is the n particle Hilbert space and $\sim$ symmetrises or anti-symmetrises bosonic or fermionic states.
2. Universal Enveloping Algebra:
$$\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/\sim$$
where $[A, B] \sim A \otimes B - B \otimes A$

Definition A.3. Let I be the two-sided ideal generated by $v \otimes v$ for $v \in V$ i.e. all elements $w_1 \otimes \cdots \otimes w_n \otimes v \otimes v \otimes z_1 \otimes \cdots \otimes z_m$ in the tensor algebra $T(V)$ containing the factor $v \otimes v$. The the exterior algebra over V is:

$$\Lambda(V) = T(V)/I$$

Remarks A.1.

1. The exterior algebra is graded:

$$\Lambda(V) = \bigoplus_{k=0}^{\infty} \Lambda^k V$$

   where $\Lambda^k V = span\{v_1 \wedge \cdots \wedge v_k \mid v_i \in V \ \forall \ i \in \{1, \ldots, k\}\}$ and the wedge product is the multiplication on the exterior algebra and arises as the image of the tensor product in the quotient.

2. Quotienting by I guarantees the skew-symmetry of the wedge product:

$$0 = (v+w) \wedge (v+w) = v \wedge v + v \wedge w + w \wedge v + w \wedge w = v \wedge w + w \wedge v \implies v \wedge w = -w \wedge v$$

A.1.1 Differential Forms

Grassmann's exterior algebra has many uses, including for example geometric algebra used to frame Euclidean geometry, and non-commuting variables describing fermionic degrees of freedom. We will use it to define Cartan's theory of differential forms.

We define $\Lambda^k T^*\mathscr{M}$, the k^{th} exterior bundle to be the bundle whose typical fibre over p is $\Lambda^k T_p^*\mathscr{M}$. Smooth sections of this bundle are denoted $\Omega^k(\mathscr{M}) \coloneqq \Gamma(\Lambda^k T^*\mathscr{M})$ and referred to as 'k' forms. These objects are best thought of as totally anti-symmetric covariant (0,k) tensor fields on a manifold.

The wedge product lifts to k forms pointwise and is graded commutative:

$$\alpha \wedge \beta = (-1)^{ab} \beta \wedge \alpha \quad \forall \ \alpha \in \Omega^a(\mathscr{M}) \ \forall \ \beta \in \Omega^b(\mathscr{M})$$

Note: $\Omega^k(\mathscr{M})$ is a vector space[9] (albeit infinite dimensional over $\mathbb{R}$) with the addition inherited pointwise from $T_p^*\mathscr{M}$.

We can work in the graded algebra $\Omega^\bullet(\mathscr{M}) = \bigoplus_k \Omega^k(\mathscr{M})$. We wish to have a means of traversing the grading structure. We first address grading regression via a definition. After we tackle graded ascension by means of a theorem.

Definition A.4. The interior multiplication (contraction) is defined for any exterior algebra $\Lambda(V)$ as follows:

$$\begin{aligned} \iota_w &: \Lambda^k V \to \Lambda^{k-1} V \\ v_1 \wedge \cdots \wedge v_k &\mapsto \sum_i (-1)^{i-1} \langle v_i, w \rangle \ v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_k \ \ \forall \ v_i \in V \ \forall \ w \in V^* \end{aligned}$$

where the i^{th} vector is omitted and $\langle \cdot, \cdot \rangle : V \times V^* \to \mathbb{F}$ is the natural pairing. Specialising to forms we sometimes write $X \lrcorner \omega \coloneqq \iota_X(\omega)$ for $X \in \mathfrak{X}(\mathscr{M})$ and $\omega \in \Omega^k(\mathscr{M})$ where contraction is defined pointwise.

Theorem A.1 (Exterior Derivative)**.**

$$\exists! \ \mathbb{R} - \text{linear operator} \quad d : \Omega^k(\mathscr{U}) \to \Omega^{k+1}(\mathscr{U}) \quad \forall \ k \geq 0 \ \forall \ \mathscr{U} \subseteq \mathscr{M} \text{ open, such that:}$$

[9] Alternatively $\Omega^k(\mathscr{M})$ is finite as a module over the ring $C^\infty(\mathscr{M})$

1. $\forall\ \mathscr{U} \subset \mathscr{M}$ open: $\alpha = \alpha'$ on $\mathscr{U} \implies d\alpha = d\alpha'$ on $\mathscr{U}$
2. $\forall\ f \in C^\infty(\mathscr{M})$, df coincides with the differential of the map
3. $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^a \alpha \wedge d\beta\ \forall\ \alpha \in \Omega^a(\mathscr{M})\ \forall\ \beta \in \Omega^b(\mathscr{M})$
4. $d^2\alpha = 0\ \forall\ \alpha \in \Omega^a(\mathscr{M})$

Proof of Theorem A.1. First we prove uniqueness:
Locality (1.) $\implies$ we can work in charts, hence we choose a co-ordinate basis $\{dx^i\}$ for $\Omega^1(\mathscr{U})$ on the patch $\mathscr{U}$. Graded Leibniz (3.) and nilpotence (4.) $\implies d(f dx^I) = df \wedge dx^I$ for a function f and $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ for the multi-index I. Hence by linearity we have $d(\alpha_I dx^I) = d\alpha_I \wedge dx^I$.

Next we prove existence:
Define the derivative on a chart patch $\mathscr{U}$ by $d(\alpha_I dx^I) = d\alpha_I \wedge dx^I$ for k-forms α and simply df is defined to be the differential on functions. Then locality (1.) and agreement with the differential (2.) are immediate and the graded Leibniz rule (3.) follows from the ordinary Leibniz rule $d(fg) = fdg + gdf\ \ \forall\ f, g \in C^\infty(\mathscr{U})$ and the graded commutativity of the exterior algebra. Nilpotence of d (4.) on functions follows from the commutativity of mixed partial derivatives $d^2 f = d(\frac{\partial f}{\partial x^i} dx^i) = \frac{\partial^2 f}{\partial x^j \partial x^i} dx^j \wedge dx^i$ and nilpotence on forms is an immediate consequence of nilpotence on functions $d^2(\alpha_I dx^I) = d(d\alpha_I \wedge dx^I) = d^2\alpha_I \wedge dx^I - d\alpha_I \wedge d^2 x^I = 0$
□

A.1.2 Integration

Differential forms allow us to integrate on a manifold. Take $\mathscr{M}$ to be a smooth oriented n dimensional manifold throughout.

First consider a top form $\omega \in \Omega^n(\mathscr{M})$ with compact support in a single chart $supp\,\omega \subset \mathscr{U}$. Then letting x^i denote both the co-ordinate maps and co-ordinates in the image, ω can be written as $\omega = f\, dx^1 \wedge \cdots \wedge dx^n$ for smooth function $f \in C^\infty(\mathscr{M})$ and its integral is given by:

$$\int_{\mathscr{M}} \omega = \int_{x(\mathscr{U})} f(x)\, dx^1 \wedge \cdots \wedge dx^n$$

Where $f(x) = f \circ x^{-1}$ is the co-ordinate representation of the function and the RHS is to be interpreted as a standard integral of a function over a domain in $\mathbb{R}^n$, the wedge products are purely to track orientation.

Now to relax the single chart support, we make use of a partition of unity $\{\psi_i\}$ subordinate to an open cover $\{\mathscr{U}_i\}$ of $\mathscr{M}$. Then the integral of a compactly supported top form $\omega \in \Omega^n(\mathscr{M})$ is:

$$\int_{\mathscr{M}} \omega = \sum_i \int_{\mathscr{M}} \psi_i\, \omega$$

Here we will simply state Stoke's theorem but not provide a proof, which can be found in [20].

Theorem A.2 (Stoke's Theorem)**.** For an n-dimensional manifold $\mathscr{M}$ without boundary and an $n-1$ form with compact support $\omega \in \Omega^{n-1}(\mathscr{M})$:

$$\int_{\mathscr{M}} d\omega = 0$$

For an n-dimensional manifold $\mathscr{M}$ with boundary $\partial\mathscr{M}$ and an $n-1$ form $\omega \in \Omega^{n-1}(\mathscr{M})$:

$$\int_{\mathscr{M}} d\omega = \int_{\partial\mathscr{M}} \omega$$

A.1.3 Forms Valued in Vector Bundles

A slightly more elaborate object to consider is a form valued in a vector bundle. For a review of fibre bundles peek ahead at section A.2.

Suppose $E \xrightarrow{\pi} \mathscr{M}$ is a vector bundle with the typical fibre of a rank n vector space V. We say an E valued k form ω is a section of $E \otimes \Lambda^k T^*\mathscr{M}$ and write $\omega \in \Omega^k(\mathscr{M}, E)$. To be clear, at a point $p \in \mathscr{M}$ the E valued k form ω is an element of $\Lambda^k T_p^*\mathscr{M} \otimes V$. Alternatively $\omega \mid_p$ is a totally skew symmetric linear map:

$$\omega \mid_p : \underbrace{T_p\mathscr{M} \times \cdots \times T_p\mathscr{M}}_{k \text{ copies}} \xrightarrow{\sim} V$$

In local co-ordinates one could write $\omega = g(x)\, dx^1 \wedge \cdots \wedge dx^k$ where $g : \mathscr{M} \to V$ is a vector-valued function. Or alternatively $\omega = h \otimes \eta$ where the section $h : \mathscr{M} \to E$ is the bundle factor, $\eta = f(x)\, dx^1 \wedge \cdots \wedge dx^k$ is an honest k form. Here $g(x) = f(x) \cdot h(x)$ where $\cdot$ is the pointwise scalar multiplication on V.

A vector bundle can be adorned with additional structure. In particular if V is an algebra, we can multiply vector valued k forms. For physics applications we are typically interested in forms valued in the Lie algebra $\mathfrak{g}$ of some group. Here we treat only matrix algebras, muttering something about Ado's theorem to justify ourselves[10].

Definition A.5. Let $\omega \in \Omega^k(\mathscr{M}, \mathfrak{g})$ and $\eta \in \Omega^p(\mathscr{M}, \mathfrak{g})$ be k and p forms respectively valued in a matrix Lie algebra $\mathfrak{g}$. We can write $\omega = A \otimes \tilde{\omega}$ with $A \in C^\infty(\mathscr{M}, \mathfrak{g})$ and $\tilde{\omega} \in \Omega^k(\mathscr{M})$. Similarly write $\eta = B \otimes \tilde{\eta}$. Then the wedge product is to be understood as:

$$\omega \wedge \eta = AB \otimes (\tilde{\omega} \wedge \tilde{\eta})$$

Where the product is matrix multiplication pointwise on the $\mathfrak{g}$ factors and the ordinary wedge product on forms.

Definition A.6. Let ω and η be as defined in definition A.5. Then the commutator is to be understood as:

$$[\omega, \eta] = [A, B] \otimes \tilde{\omega} \wedge \tilde{\eta}$$

Proposition A.1. Let ω and η be as defined in definition A.5. Then:

$$[\omega, \eta] = \omega \wedge \eta - (-1)^{kp}\, \eta \wedge \omega$$

Proof of Proposition A.1.

$$\begin{aligned}
[\omega, \eta] &= [A, B] \otimes \tilde{\omega} \wedge \tilde{\eta} \\
&= AB \otimes \tilde{\omega} \wedge \tilde{\eta} - BA \otimes \tilde{\omega} \wedge \tilde{\eta} \\
&= AB \otimes \tilde{\omega} \wedge \tilde{\eta} - (-1)^{kp}\, BA \otimes \tilde{\eta} \wedge \tilde{\omega} \\
&= \omega \wedge \eta - (-1)^{kp}\, \eta \wedge \omega
\end{aligned}$$

□

[10] Ado's theorem states that under fairly mild conditions, Lie algebras are always equivalent to some matrix algebra.

Corollary A.1. The commutator vanishes on even degree forms $\omega \in \Omega^{2k}(\mathscr{M}, \mathfrak{g})$:

$$[\omega, \omega] = 0$$

A.2 Fibre-Bundles

Gauge theories at the classical level are enormously redundant descriptions of dynamical fields propagating on spacetime. That sentence should not be parsed as a scathing indictment of gauge theory, but operationally as a counting of the degrees of freedom. If one wants a rigorous account of such theories that can be extended to topologically interesting spacetimes one arrives at the construction of principal bundles. This is a technology that allows us to make local symmetry transformations independently at each spacetime event, while accounting both for the spacetime structure and for the myriad of ways a collection of fields might transform under said symmetry group.

What follows is a series of definitions that form the backbone of gauge theory. To begin with, the cornerstone of modern geometry, the fibre bundle:

A.2.1 Smooth Fibre Bundles

Definition A.7. A smooth fibre bundle consists of the tuple of data $(E, \mathscr{M}, \pi, \mathscr{F})$; referred to as the total space, base space, projection and typical fibre respectively; such that:

1. $E, \mathscr{M}, \mathscr{F}$ are smooth manifolds
2. $\pi : E \to \mathscr{M}$ is a smooth surjective submersion
3. $\forall\ p \in \mathscr{M}\ \exists\ \mathscr{U} \subset \mathscr{M}$ open with $P \in \mathscr{U}$ and $\exists\ \Phi_{\mathscr{U}} : \pi^{-1}(\mathscr{U}) \to \mathscr{U} \times \mathscr{F}$ a diffeomorphism (called a local trivialisation)
4. $\pi_1 \circ \Phi_{\mathscr{U}} = \pi$ where $\pi_1 : \mathscr{U} \times \mathscr{F} \to \mathscr{U}$ is projection onto the first factor

A conservative notation for the bundle $(E, \mathscr{M}, \pi, \mathscr{F})$ is $E \xrightarrow{\pi} \mathscr{M}$. Local trivialisations express the property that locally fibre bundles are simply a Cartesian product of the base space with the typical fibre. The following commutative diagram illustrates these local trivialisations:

$$\begin{array}{ccc} \pi^{-1}(\mathscr{U}) \subseteq E & \xrightarrow{\Phi_{\mathscr{U}}} & \mathscr{U} \times F \\ \pi \big\downarrow & & \big\downarrow \pi_1 \\ \mathscr{U} \subseteq \mathscr{M} & \xrightarrow{id} & \mathscr{U} \subseteq \mathscr{M} \end{array}$$

Maps $\sigma : \mathscr{M} \to E$ assign to each point in the base space, some element in the fibre over that point. Such maps are called cross sections (or sections for short) as they carve a slice through the bundle. All sections are assumed to be smooth as maps between manifolds. $\Gamma(E)$ denotes the space of smooth sections for any bundle E. For the tangent bundle $T\mathscr{M}$ for some manifold $\mathscr{M}$, the special notation $\mathfrak{X}(\mathscr{M}) \coloneqq \Gamma(T\mathscr{M})$ is used as here sections are to be interpreted as vector fields.

Definition A.8. A vector bundle is a smooth fibre bundle as in A.7 such that the fibre $\mathscr{F}$ has a linear structure.

Next we define the notion of a principal bundle. These are fibre bundles with a group action on the fibres. In fact the fibres themselves are each isomorphic to a copy of the structure group. This is the key to describing gauge transformations- a structure that can handle group transformations that vary over the spacetime.

Definition A.9. A fibre bundle $(\mathscr{P}, \mathscr{M}, \pi, G)$ according to definition A.7 where G is a Lie group is called a principal G bundle if:

1. $\exists\, (\mathscr{P} \curvearrowleft G) : \mathscr{P} \times G \to \mathscr{P},\ \ (p, g) \mapsto p \cdot g$ a smooth right action of G on P
2. $p \cdot g = p \implies g = e$ free G action on fibres $\mathscr{P}_x$
3. $\forall p, q \in P_x\ \exists!\ g \in G : q = p \cdot g$ transitive G action on fibres $\mathscr{P}_x$

Remarks A.2.

1. Properties 2. and 3. in definition A.9 guarantee each fibre is diffeomorphic to G.
2. Sections of the bundle are gauge transformations. To each $x \in \mathscr{M}$, $g : \mathscr{M} \to \mathscr{P}$ assigns a group element $\pi_2 \circ \Phi_{\mathscr{U}} \circ g(x) \in G$ on each local trivialisation.

The following commutative diagram illustrates the local trivialisations and the group action:

$$\begin{array}{ccc}
\pi^{-1}(\mathscr{U}) \subseteq \mathscr{P} & \xrightarrow{\Phi_{\mathscr{U}}} & \mathscr{U} \times G \\
{\scriptstyle \mathscr{P} \curvearrowleft G}\uparrow & & \uparrow{\scriptstyle G} \\
\pi^{-1}(\mathscr{U}) \subseteq \mathscr{P} & \xrightarrow{\Phi_{\mathscr{U}}} & \mathscr{U} \times G \\
{\scriptstyle g}\circlearrowleft\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi_1} \\
\mathscr{U} \subseteq \mathscr{M} & \xrightarrow{id} & \mathscr{U} \subseteq \mathscr{M}
\end{array}$$

This structure serves as the master-key. One imagines dialling a gauge on the Principal bundle and each field in the action transforms accordingly. For this to work we require the notion of an associated bundle.

Definition A.10. Let $\mathscr{P} \xrightarrow{\pi} \mathscr{M}$ be a principal bundle as given in definition A.9. Given a representation $\rho : G \to GL(V)$, the associated vector bundle $E \xrightarrow{\pi_E} \mathscr{M}$ is given by the quotient:

$$E \coloneqq (\mathscr{P} \times V)/\sim$$

where the equivalence relation $\sim$ is given by $(p, v) \sim (p', v') \iff \exists\, g \in G : p' = p \cdot g,\ v' = \rho(g^{-1})(v)$.

It's worth pausing here to appreciate the apparatus we've developed. Given an action consisting of an array of field content, invariant under some local group of symmetries; gauge invariance of said action dictates the representation under which the group linearly transforms each field. After constructing a principal bundle and associated bundles for each representation as specified by the action; one has systematic control of the gauge transformations simultaneously for each field. This is true of flat spacetimes, curved spacetimes, spacetimes with holes and handles and all manner of topological quirks.

An important vector bundle associated to a principal bundle is the adjoint bundle, whose representation space is the Lie algebra.

A.2.2 Connections and Curvature

Yang-Mills theory is in truth a theory of connections on principal bundles. In the following we make this precise:

Definition A.11. The vertical subspace $\mathscr{V}_q\mathscr{P}$ of the tangent space of a principal bundle $\mathscr{P}$ at a point $q \in \mathscr{P}$ is the subspace $\mathscr{V}_q\mathscr{P} \subset T_q\mathscr{P}$ given by the kernel of $\pi_* : T\mathscr{P} \to T\mathscr{M}$ (the differential of the projection map π) at q:

$$\mathscr{V}_q\mathscr{P} \coloneqq ker(\pi_*|_q)$$

Definition A.12. A connection is a pointwise smooth choice of decomposition of the tangent bundle into vertical and so called horizontal sub-bundles $T_q\mathscr{P} = \mathscr{V}_q\mathscr{P} \oplus \mathscr{H}_q\mathscr{P}$ subject to:

1. $(\mathscr{P} \curvearrowleft g)_*(\mathscr{H}_q\mathscr{P}) = \mathscr{H}_{q\cdot g}\mathscr{P} \ \ \forall\, g \in G$
2. $\mathscr{H}\mathscr{P} \subset T\mathscr{P}$ is a smooth sub-bundle.

The first property is simply a group action compatibility condition. If we push the horizontal space tangent to some point in a fibre to another point in the fibre using the group action it ought to agree with the choice of horizontal space we'd made there. This is illustrated in figure 5:

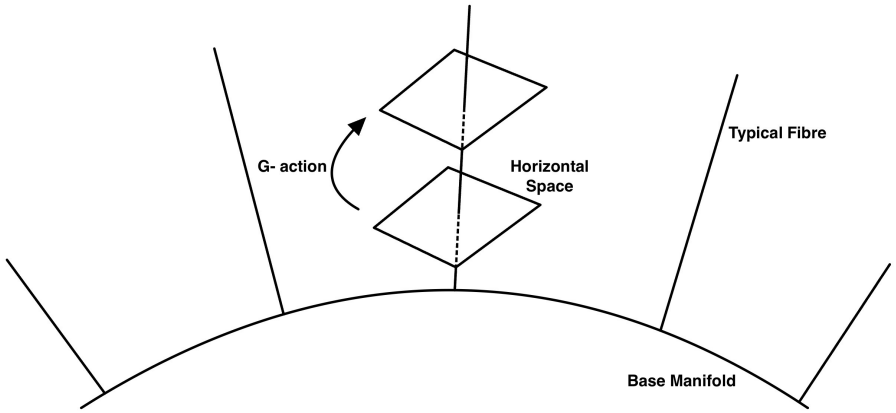


Figure 5: A depiction of the G-invariance of a connection

It turns out there is an equivalent yet much more useful way to characterise the connection. This is done via a global one-form on the principal bundle. This connection one-form is valued however not in functions on the bundle but rather functions on the Lie Algebra. We write $\omega \in \Omega^1(\mathscr{P}, \mathfrak{g})$ to denote the space of such forms.

We also require the notion of a fundamental vector field. These are vector fields generated by a Lie algebra element and tangent to curves lying entirely along a fibre. The vector field generated by X we will denote $X^\#$ and is given by $X^\#_q = \frac{d}{dt}\,|_{t=0}\; q \cdot exp(tX)$ where here $\cdot$ denotes the right action of the group, and exp is the exponential map $\mathfrak{g} \to G$.

Definition A.13 (Connection 1–form)**.** A connection one form on a principal bundle $\mathscr{P}$ is a Lie algebra valued one–form;

$$\omega \in \Omega^1(P, \mathfrak{g})$$

satisfying:

1. $\omega(X^\#) = X \quad \forall\, X \in \mathfrak{g}$,
2. $(\mathscr{P} \curvearrowleft g)^*\omega = \mathrm{Ad}_{g^{-1}}\omega$ for all $g \in G$.

Remarks A.3.

1. The first property in definition A.13 can equivalently be stated by requiring ω restricts to the Maurer-Cartan form on fibres.

2. The second property demands that pulling ω back by a group action is equivalent to instead mapping the Lie algebra factor by the inverse of said group element. This is a compatibility condition that amounts to property 1 in definition A.12

3. The connection is recovered from the connection one form via $\mathscr{H}_q\mathscr{P} = ker(\omega\mid_q)$

As with any connection there is an associated notion of curvature. In General relativity one associates to the Levi-Civita connection on the frame bundle the Riemann curvature 2-form $R = d\omega + \omega \wedge \omega$. It is then common to conflate the $\mathfrak{gl}(4,\mathbb{R})$ valued 2-form with a (1,3) tensor field. Here too we have a $\mathfrak{g}$-valued curvature 2-form on the principal bundle:

$$\Omega = d\omega + \omega \wedge \omega$$

Where the wedge product is to be understood as matrix multiplication on the algebra factors and an ordinary wedge product on the form factors[11].

Given a local choice of gauge, i.e a section $g : \mathscr{U} \to \mathscr{P}$, we can pull back the connection and its curvature:

$$\mathcal{A} \coloneqq g^*\omega \tag{A.1}$$
$$\mathcal{F} \coloneqq g^*\Omega \tag{A.2}$$

These fields are now Lie algebra valued forms on the base manifold; $\mathcal{A} \in \Omega^1(\mathscr{M}, ad\mathscr{P})$, $\mathcal{F} \in \Omega^2(\mathscr{M}, ad\mathscr{P})$ where $ad\mathscr{P}$ is the adjoint bundle whose typical fibre is the Adjoint representation space, $\mathfrak{g}$. This subtly different notation for being 'Lie algebra valued' reflects how these Yang Mills fields transform under a gauge transformation. $\mathcal{A}$ and $\mathcal{F}$ are simply related by:

$$\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$$

For the most part we physicists are concerned with fields that propagate on spacetime. As such from here we will mostly work with $\mathcal{A}$ and $\mathcal{F}$ locally rather that ω and Ω.

Since the group action on a principal bundle is free and transitive on the fibres, locally two section s and s' are related by a function $g : \mathscr{U} \to G$:

$$s'(x) = s(x) \cdot g(x) \;\; \forall x \in \mathscr{U}$$

$g : \mathscr{U} \to G$ is called a gauge transformation.

Consider a tensor field $K \in \Gamma(E)$ for some vector bundle E associated to $\mathscr{P}$ by the representation $\rho : G \to GL(V)$. Then under a gauge transformation $g(x)$, $K \mapsto \rho(g)(K)$.

Hence under such a gauge transformation $\mathcal{A}$ and $\mathcal{F}$ transform as:

$$\mathcal{F} \mapsto g\mathcal{F}g^{-1}$$
$$\mathcal{A} \mapsto g\mathcal{A}g^{-1} - dg \cdot g^{-1}$$

[11] If G is not a matrix group then $\omega \wedge \omega$ is to be understood as $\frac{1}{2}[\omega, \omega]$ where the forms are multiplied with the ordinary wedge product, and $[\cdot, \cdot]$ is the Lie bracket on $\mathfrak{g}$

Where the connection picks up an extra Maurer-Cartan term because it is an affine connection and not a tensor.

A p-form K on the base manifold which is valued in a vector bundle E associated to the principal bundle $\mathscr{P}$ by the representation $\rho : G \to GL(V)$ can be covariantly differentiated:

$$d_{\mathcal{A}}K = dK + \rho_*(\mathcal{A})(K)$$

In practise one simply writes:

$$d_A = d + A$$

And the representation is inferred from the object being differentiated. For our purposes we will only be concerned with the Adjoint bundle and so:

$$d_A = d + [A, \cdot]$$

Proposition A.2 (Bianchi Identity)**.** The local curvature $\mathcal{F}$ satisfies a Bianchi identity:

$$d_{\mathcal{A}}\mathcal{F} = 0$$

Proof of Proposition A.2.

$$\begin{aligned} d_{\mathcal{A}}\mathcal{F} &= d\mathcal{F} + [\mathcal{A}, \mathcal{F}] \\ &= d(d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}) + [\mathcal{A}, d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}] \\ &= d\mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge d\mathcal{A} + \mathcal{A} \wedge d\mathcal{A} - d\mathcal{A} \wedge \mathcal{A} + \mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} \\ &= 0 \end{aligned}$$

□

A.3 Complex Geometry

The purpose of twistors is to connect integrable systems with homological data, and to translate between conformal structure on spacetime with complex structure in twistor space. We must then develop some complex geometry. We begin with complex manifolds.

A.3.1 Complex Manifolds

Definition A.14. Given a manifold $\mathcal{M}$ and $p \in \mathcal{M}$ with tangent space $T_p\mathcal{M}$; a linear map $J_p : T_p\mathcal{M} \xrightarrow{\sim} T_p\mathcal{M}$ such that:

$$J^2 = -id\,|_{T_p\mathcal{M}}$$

Is called an almost complex structure.

The almost complex structure provides at least an analogue of multiplication by i on the tangent bundle.

A complex manifold is a very analogous to a smooth manifold, where local patches are instead homeomorphic to $\mathbb{C}^n$ and the smooth structure is replaced by a holomorphic structure.

Definition A.15. A topological space X which is Hausdorff and second countable is a complex manifold of complex dimension n if:

1. $\forall x \in X$ there exists both $U \subset X$ an open neighbourhood of x and a homeomorphism $\varphi : U \to V \subset \mathbb{C}^n$

2. $\forall\ U, U'$ open in X with $U \cap U' \neq \emptyset$ with $\varphi : U \to V$ and $\varphi' : U' \to V'$, the transition maps $\varphi' \circ \varphi^{-1} \mid_{\varphi(U \cap U')}: \varphi(U \cap U') \to \varphi'(U \cap U')$ are holomorphic as maps from an open subset of $\mathbb{C}^n$ to $\mathbb{C}^n$

Maps between complex manifolds are said to be holomorphic if in some (and hence any) co-ordinates the represented map is holomorphic. The complex analogue of a diffeomorphism is called a biholomorphism; this is a holomorphic map with a holomorphic inverse.

An almost complex structure gives the tangent bundle a pointwise complex structure in the sense of linear algebra. This does not necessarily guarantee the existence of holomorphic co-ordinates.

Proposition A.3. Every complex manifold admits an almost complex structure.

Proof of Proposition A.3. Let $z^i = x^i + iy^i$ be holomorphic co-ordinates on a complex manifold X. It is sufficient to specify a linear map $J : T_pX \xrightarrow{\sim} T_pX$ on basis vectors. Choosing the co-ordinate basis let $J(\partial_x) = \partial_y$ and $J(\partial_y) = -\partial_x$ □

Hence we see that a complex structure is at least as strong as an almost complex structure. In fact it is stronger.

An almost complex structure lends itself to an eigenspace decomposition of the tangent bundle if we first complexify:

$$T_{\mathbb{C}}X = TX \otimes_{\mathbb{R}} \mathbb{C}$$

Then the decomposition is:

$$T_{\mathbb{C}}X = T^{1,0}X \oplus T^{0,1}X$$

Where $J(V) = iV\ \forall\ V \in T_p^{1,0}X\ \forall\ p \in X$, $J(V) = -iV\ \forall\ V \in T_p^{0,1}X\ \forall\ p \in X$ and the direct sum is defined fibre-wise. The sections of $T^{1,0}X$ are called holomorphic vector fields and sections of $T^{0,1}X$ anti-holomorphic.

To see precisely when an almost complex structure on the tangent bundle arises from a complex structure on the manifold, we state but do not provide proof of the following theorem

Theorem A.3 (Newlander-Nirenberg)**.** Let X be a topological space with an almost complex structure J. If the corresponding decomposition of the tangent bundle satisfies:

$$[T_p^{0,1}X, T_p^{0,1}X] \subset T_p^{0,1}X\ \ \forall\, p \in X$$

where $[\cdot, \cdot]$ is the Lie bracket, then J is said to be integrable and arises from a complex structure.

A.3.2 Holomorphic Vector Bundles

Holomorphic vector bundles are a straight forward generalisation of smooth vector bundles (as defined in A.8) to complex geometry. These structures will be of the utmost importance in this essay so we give the full definition, mirroring that of the real case.

Definition A.16. A rank d holomorphic vector bundle consists of the tuple of data (H, X, π, V); referred to as the total space, base space, projection and typical $\mathbb{C}$-linear fibre respectively; such that:

1. H, X are complex manifolds, V is a $\mathbb{C}$-linear d-(complex) dimensional vector space
2. $\pi : E \to X$ is a holomorphic surjective submersion

3. $\forall\, p \in X\ \exists\, U \subset X$ open with $P \in U$ and $\exists\, \Phi_U : \pi^{-1}(U) \to U \times V$ a biholomorphism (called a local trivialisation)

4. $\pi_1 \circ \Phi_U = \pi$ where $\pi_1 : U \times V \to U$ is projection onto the first factor

A useful way to construct a holomorphic vector bundle is by gluing local sheaves[12] together. We give this equivalent definition:

Definition A.17. Let X be a complex manifold. A rank d holomorphic vector bundle over X consists of an open cover $\{U_i\}_{i\in I}$ of X, where I is the index set; and holomorphic transition functions $f_{ij} : U_i \cap U_j \to GL(d, \mathbb{C})$ on every intersection satisfying the following co-cycle conditions:

1. $f_{ii} = id$ on $U_i \forall\ i \in I$

2. $f_{ij} = f_{ji}^{-1}\ \forall\ i, j \in I$

3. $f_{ij} f_{jk} f_{ki} = id$ on $U_i \cap U_j \cap U_k\ \forall\ i, j, k \in I$

We can construct the data in definition A.16 from the data in definition A.17 by:

$$H = \bigsqcup_i U_i \times \mathbb{C}^d / \sim$$
$$(x, v) \in U_i \times \mathbb{C}^d \sim (x, f_{ij}(x)v) \in U_j \times \mathbb{C}^d\ \forall x \in U_i \cap U_j$$
$$\pi([x, v]) = x$$

A.3.3 Complex Projective Space

Next we turn to the subject of projective geometry.

Definition A.18. The nth complex projective space $\mathbb{CP}^n$ is the space of complex one-dimensional subspaces of $\mathbb{C}^{n+1}$:

$$\mathbb{CP}^n = (\mathbb{C}^{n+1} - \{0\}) / \sim$$
$$(z_1, \ldots, z_{n+1}) \sim (\lambda z_1, \ldots, \lambda z_{n+1})\ \ \lambda \in \mathbb{C}^*$$

An extremely useful way to put co-ordinates on $\mathbb{CP}^n$ is to choose 'homogenous co-ordinates' $[z_1 : \cdots : z_{n+1}]$ where the square brackets indicate the equivalence class. In other words these co-ordinates are defined up to an overall scale $[z_1 : \cdots : z_{n+1}] = [\lambda z_1 : \cdots : \lambda z_{n+1}]$ and note: $[z_1 : \cdots : z_{n+1}] \neq [0 : \cdots : 0]$.
It is simple to choose a holomorphic atlas:

$$U_i = \{[z_1 : \cdots : z_{n+1}] \mid z_i \neq 0\} \tag{A.3}$$
$$\phi_i([z_1 : \cdots : z_{n+1}]) = (\frac{z_1}{z_i}, \cdots, \widehat{z_i}, \ldots, \frac{z_{n+1}}{z_i}) \subset \mathbb{C}^n \tag{A.4}$$

Where the hatted co-ordinate is omitted. There are a few things to comment on here.

Remarks A.4.

1. It ought to be noted that 'homogenous co-ordinates' are really elements of the manifold and the charts above give honest co-ordinate charts.

[12] see Appendix B

2. The co-ordinate maps are well-defined. This is because under a change of scale; $z_i \mapsto \lambda z_i, z_j \mapsto \lambda z_i$ the ratio $\frac{z_j}{z_i} \mapsto \frac{\lambda z_j}{\lambda z_i}$ is unchanged.

3. The charts cover the manifold since $[0 : \cdots : 0]$ is not included.

4. The transition functions $\phi_j \circ \phi_i^{-1}$ are holomorphic and as such $\mathbb{CP}^n$ is a complex manifold. This is worked out below:

$$\begin{aligned}\phi_j \circ \phi_i^{-1}(z_1, \ldots, z_n) &= \phi_j([z_1 : \cdots : \underbrace{1}_{i^{th}\text{ slot}} : \cdots : z_n]) \\ &= (\frac{z_1}{z_j}, \ldots \frac{1}{z_j}, \ldots, \frac{z_n}{z_j})\end{aligned}$$

Since the map is a function only of the co-ordinates and not their conjugates it is holomorphic by the Cauchy-Riemann equations. Note on the chart overlap neither z_i nor z_j vanish.

A natural way to produce $\mathbb{CP}^3$ is via the Hopf fibration. Consider the following $U(1)$ action on $S^7 \subset \mathbb{C}^4$:

$$(z_1, z_2, z_3, z_4) \mapsto (e^{i\theta} z_1, e^{i\theta} z_2, e^{i\theta} z_3, e^{i\theta} z_4)$$

Then $\mathbb{CP}^3$ can be realised by the quotient:

$$\mathbb{CP}^3 = S^7 / U(1)$$

The principal $U(1)$ bundle $S^7 \to \mathbb{CP}^3$ is called the Hopf fibration.

The space $\mathbb{CP}^1$ has a very nice interpretation as the 2-sphere. All real and orientable two dimensional manifolds admit a complex structure and are know as Riemann surfaces. $\mathbb{CP}^1$ is the Riemann sphere.

Proposition A.4. Every 2d real orientable manifold is a complex manifold.

Proof of Proposition A.4. Let X be such a real manifold. Choose an almost complex structure J to be 90° rotation. After complexifying the tangent bundle $T^{0,1}X$ is a complex rank 1 vector bundle. As such it is locally generated by a single vector field which we shall call $Z = X + iJX$. We note that $JZ = -iZ$ and as such is indeed a section of $T^{0,1}X$. In fact any such anti-holomorphic vector field Z can be written in this form by taking $X = \frac{1}{2}Z$. Then $\forall\ Z_1, Z_2 \in T^{0,1}X\ \exists\ f, g \in C^\infty(X, \mathbb{C}) : Z_1 = fZ,\ Z_2 = gZ$. Then $[Z_1, Z_2] = (fZ(g) - gZ(f))Z \in \Gamma(T^{0,1}X)$. Hence by the Newlander-Nirenberg theorem A.3 X is a complex manifold. □

B Appendix: Sheaves and Cohomology

The following appendix briefly addresses some cohomology theories. See [6] for a reference of the material provided.

B.1 Sheaves

Cohomology is a tour de force of modern algebraic topology. The idea is rather simple and yet far reaching. Given a topological space, one hopes to associate algebraic structures to the space that are invariant under

some class of topology preserving transformations. For the topologist this class might consist of homeomorphisms or perhaps homotopies. For the physicist who concerns his or herself with smooth spacetimes this class is likely diffeomorphisms. In this manner one might hope to classify certain spaces up to topology.

At the risk of (ever so slightly) reducing the generality of the following section, but in the interest of providing a clear language and algebraic structure we will consider only sheaf cohomology.

Definition B.1. Let X be a topological space. A presheaf of sets $\mathscr{F}$ consists of the following data:

1. $\forall\ U \subseteq X$ open, we have a set $\mathscr{F}(U)$ called the 'sections' over U.
2. For every inclusion $V \subseteq U$ there is a restriction map $\rho_{U,V} : \mathscr{F}(U) \to \mathscr{F}(V)$ such that:
   (a) $\rho_{U,U} = id_{\mathscr{F}(U)}$
   (b) $\rho_{U,W} = \rho_{V,W} \circ \rho_{U,V}\ \forall\ W \subseteq V \subseteq U$.

Definition B.2. Let X be a topological space and $U \subseteq X$ open with $\{U_i\}_{i\in I}$ an open cover of U. A presheaf $\mathscr{F}$ on X is a sheaf if:

1. For $s, t \in \mathscr{F}(U) : \rho_{U,U_i}(s) = \rho_{U,U_i}(t)\ \forall\ i \in I \implies s = t$.
2. $\forall\ \{s_i\}_{i\in I}$ with $s_i \in \mathscr{F}(U_i)$ such that $\rho_{U_i,U_i\cap U_j}(s_i) = \rho_{U_j,U_i\cap U_j}(s_j)\ \forall\ i, j \in I$ there exists a section $S \in \mathscr{F}(U)$ such that $\rho_{U,U_i}(S) = s_i\ \forall\ i \in I$

Remarks B.1.

1. The 'sections' of $\mathscr{F}$ coincide with cross sections of fibre bundles in that context, hence the name.
2. The restriction condition in definition B.1 ensures that the way we restrict sections is well defined on the various ways to 'zoom in' on a patch.
3. The first condition in definition B.2 is known as locality. It tells us that if two sections agree locally at each point in some region, then that is enough to conclude their equality on said region. This ought to sound very reasonable.
4. The second condition in definition B.2 is known as gluing. This gives us the necessary tools to build global sheaves out of local pieces.
5. It should be noted that the first sheaf axiom implies that the global sheaf built from gluing is unique.
6. An abelian sheaf is defined in the same way as definitions B.1 and B.2 but with the additional requirement that $\mathscr{F}(U)$ be an abelian group.

B.2 Cochain Complexes

All sheaf cohomologies considered in this text will be cohomology of Abelian groups. To build such a cohomology, one begins with something called the complex. This is a graded object $C^\bullet$ together with an ascending, nilpotent map $d : C^\bullet \to C^\bullet$ i.e.:

$$C^\bullet = \bigoplus_k C^k \qquad d : C^k \to C^{k+1} \qquad d^2 = 0$$

The cochain complex is then the following sequence:

$$0 \hookrightarrow C^0 \xrightarrow{d} C^1 \xrightarrow{d} C^2 \xrightarrow{d} \dots$$

A step in a sequence is said to be exact if the image of the preceding map equals the kernel of the succeeding map. Cohomology of the cochain complex precisely measures the failure of exactness at each step.

Definition B.3. For a complex $(C^\bullet, d)$ and corresponding cohain:

$$0 \hookrightarrow C^0 \xrightarrow{d} C^1 \xrightarrow{d} C^2 \xrightarrow{d} \dots$$

The k^{th} cohomology group is given by the quotient: $H^k \coloneqq \ker(d : C^k \to C^{k+1}) / \operatorname{im}(d : c^{k-1} \to C^k)$

B.3 de-Rham Cohomology

Definition B.4. The de-Rham complex on a smooth manifold $\mathcal{M}$ is given by:

1. A graded vector space $\Omega^\bullet = \bigoplus_k \Omega^k(\mathcal{M}, \mathbb{R})$
2. $d : \Omega^k(\mathcal{M}, \mathbb{R}) \to \Omega^{k+1}(\mathcal{M}, \mathbb{R})$, Cartan's exterior derivative on forms.

The de-Rham cohomology groups are then:

$$H^k_{dR} = \ker(d : \Omega^k(\mathcal{M}) \to \Omega^{k+1}(\mathcal{M})) / \operatorname{im}(d : \Omega^{k-1}(\mathcal{M}) \to \Omega^k(\mathcal{M}))$$

i.e. closed k-forms modulo exact k-forms.

B.4 Čech Cohomology

Čech cohomology measures the failure of local sheaf data to glue. A fantastic reference for this section is [6].

Let X be a topological space and $\mathscr{F}$ a sheaf of Abelian groups on X. Furthermore fix an open cover $U = \{U_i\}_{i \in I}$. For the multi-index $(i_0, \dots, i_p)$ write $U_{i_0 \cdots i_p} \coloneqq U_{i_0} \cap \cdots \cap U_{i_p}$.

The p^{th} Čech cochain group for this cover is:

$$\check{C}^p(U, \mathscr{F}) = \prod_{(i_0, \dots i_p) \in I^{p+1}} \mathscr{F}(U_{i_0 \dots i_p})$$

To be clear, $\mathscr{F}(U_{i_0 \dots i_p})$ is an abelian group whose elements are sections on $U_{i_0 \cdots i_p}$. An element of the p^{th} cochain group is a family sections on X. It should be noted that for a finite cover, the product of Abelian groups is simply the direct sum.

The 'co-boundary' operator is given by:

$$\delta : \check{C}^p(U, \mathscr{F}) \to \check{C}^{p+1}(U, \mathscr{F})$$
$$(\partial c)_{i_0 \dots i_p} = \sum_{k=0}^{p+1} (-1)^k c_{i_0 \dots \widehat{i_k} \dots i_{p+1}} \mid_{U_{i_0 \cdots i_{p+1}}}$$

Where $\widehat{i_k}$ indicates omission. The easiest way to appreciate this operator is with a basic example. Consider an open cover consisting of two patches U_1 and U_2 and look at $\delta : \check{C}^0(U, \mathscr{F}) \to \check{C}^1(U, \mathscr{F})$. Then for $c = c_1 \oplus c_2$, $(\delta c)_{ij} = c_j \mid_{U_{12}} - c_i \mid_{U_{12}}$ and hence measures the failure of the two sections c_1 and c_2 to glue on the overlap. More generally the co-boundary operator measures how data glues on intersections.

Proposition B.1. The co-boundary operator is nilpotent: $\delta^2 = 0$

Proof of Proposition B.1.

$$(\delta^2 c)_{i_0 \dots i_{p+2}} = \sum_{j=0}^{p+2} (-1)^j (\delta c)_{i_0 \dots \widehat{i_j} \dots i_{p+2}} \, |_{U_{i_0 \cdots i_{p+2}}}$$

$$(\delta c)_{i_0 \dots \widehat{i_j} \dots i_{p+2}} = \sum_{k=0}^{p+1} (-1)^k \begin{cases} c_{i_0 \dots \widehat{i_k} \dots \widehat{i_j} \dots i_{p+2}} & \text{for } k < j \\ c_{i_0 \dots \widehat{i_j} \dots \widehat{i_{k+1}} \dots i_{p+2}} & \text{for } k \geq j \end{cases}$$

$$(\delta^2 c)_{i_0 \dots i_{p+2}} = \sum_{j=0}^{p+2} \sum_{k=0}^{j-1} (-1)^{j+k} c_{i_0 \dots \widehat{i_k} \dots \widehat{i_j} \dots i_{p+2}} + \sum_{j=0}^{p+2} \sum_{k=j}^{p+2} (-1)^{j+k} c_{i_0 \dots \widehat{i_j} \dots \widehat{i_{k+1}} \dots i_{p+2}}$$

let $l = k+1$ in the second sum

$$(\delta^2 c)_{i_0 \dots i_{p+2}} = \sum_{j=0}^{p+2} \sum_{k=0}^{j-1} (-1)^{j+k} c_{i_0 \dots \widehat{i_k} \dots \widehat{i_j} \dots i_{p+2}} - \sum_{j=0}^{p+2} \sum_{l=j+1}^{p+2} (-1)^{j+l} c_{i_0 \dots \widehat{i_j} \dots \widehat{i_l} \dots i_{p+2}}$$

and relabelling indices yields:

$$(\delta^2 c)_{i_0 \dots i_{p+2}} = \sum_{j=0}^{p+2} \sum_{k=0}^{j-1} (-1)^{j+k} c_{i_0 \dots \widehat{i_k} \dots \widehat{i_j} \dots i_{p+2}} - \sum_{k=0}^{p+2} \sum_{j=k+1}^{p+2} (-1)^{k+j} c_{i_0 \dots \widehat{i_k} \dots \widehat{i_j} \dots i_{p+2}}$$

then noting:

$$\sum_{k=0}^{p+2} \sum_{j=k+1}^{p+2} f_{jk} = \sum_{j=0}^{p+2} \sum_{k=0}^{j-1} f_{jk}$$

yields the result

□

As such we have the complex $(\check{C}^\bullet(U, \mathscr{F}), \delta)$ and cochain:

$$0 \hookrightarrow \check{C}^0(U, \mathscr{F}) \xrightarrow{\delta} \check{C}^1(U, \mathscr{F}) \xrightarrow{\delta} \check{C}^2(U, \mathscr{F}) \xrightarrow{\delta} \dots$$

We define the co-cycles and co-boundaries as follows:

$$\check{Z}^p(U, \mathscr{F}) \coloneqq \ker(\delta : \check{C}^p(U, \mathscr{F}) \to \check{C}^{p+1}(U, \mathscr{F})) \leftarrow \text{ co-cycles}$$
$$\check{B}^p(U, \mathscr{F}) \coloneqq \operatorname{im}(\delta : \check{C}^{p-1}(U, \mathscr{F}) \to \check{C}^p(U, \mathscr{F})) \leftarrow \text{ co-boundaries}$$

Then the p^{th} Čech cohomology group with respect to this cover is:

$$\check{H}^p(U, \mathscr{F}) \coloneqq \check{Z}^p(U, \mathscr{F}) / \check{B}^p(U, \mathscr{F})$$

References

[1] Michael F Atiyah, Nigel J Hitchin, Vladimir G Drinfeld, and Yuri I Manin. Construction of instantons. *Physics Letters A*, 65(3):185–187, 1978.

[2] Roger Penrose. Twistor algebra. *Journal of Mathematical physics*, 8(2):345–366, 1967.

[3] Mikio Nakahara. *Geometry, topology and physics*. CRC press, 2018.

[4] John Willard Milnor and James D Stasheff. *Characteristic classes*. Number 76. Princeton university press, 1974.

[5] Shiing-Shen Chern and James Simons. Characteristic forms and geometric invariants. *Annals of Mathematics*, 99(1):48–69, 1974.

[6] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. John Wiley & Sons, 2014.

[7] Robin Hartshorne. *Algebraic geometry*. Springer Science & Business Media, 2013.

[8] Alexander A Belavin, Alexander M Polyakov, Albert S Schwartz, and Yu S Tyupkin. Pseudoparticle solutions of the yang-mills equations. *Physics Letters B*, 59(1):85–87, 1975.

[9] José Figueroa-O'Farrill. Lecture 4: Instantons. Lecture notes, GT 2006, 2006. Accessed: 2026.

[10] Lionel Mason and David Skinner. The complete planar s-matrix of $\mathcal{N} = 4$ sym as a wilson loop in twistor space. *Journal of High Energy Physics*, 2010.

[11] Roger Penrose, Wolfgang Rindler, and Joshua N Goldberg. Spinors and space-time, volume i: Two-spinor calculus and relativistic fields, 1986.

[12] Nima Arkani-Hamed, Jacob L Bourjaily, Freddy Cachazo, Alexander B Goncharov, Alexander Postnikov, and Jaroslav Trnka. Scattering amplitudes and the positive grassmannian. *arXiv preprint arXiv:1212.5605*, 2012.

[13] Bryce S DeWitt. Quantum theory of gravity. ii. the manifestly covariant theory. *Physical Review*, 162(5):1195, 1967.

[14] Richard S Ward. On self-dual gauge fields. *Physics Letters A*, 61(2):81–82, 1977.

[15] Kevin Costello. Lecture – mathematical physics, phys 777. PIRSA lecture series, 2025. Perimeter Institute Recorded Seminar Archive (PIRSA).

[16] Nima Arkani-Hamed and Jaroslav Trnka. The amplituhedron. *Journal of High Energy Physics*, 2014(10):1–33, 2014.

[17] Susama Agarwala and Eloi Marin-Amat. Wilson loop diagrams and positroids. *Communications in Mathematical Physics*, 350(2):569–601, 2017.

[18] Susama Agarwala and Siân Zee Fryer. A study in $g_{\searrow,\geq 0}(2,6)$: from the geometric case book of wilson loop diagrams and $\mathcal{N} = 4$ sym. *Annales de l'Institut Henri Poincaré D*, 10(1), 2023.

[19] Paul Heslop and Alastair Stewart. The twistor wilson loop and the amplituhedron. *Journal of High Energy Physics*, 2018(10):1–19, 2018.

[20] John M Lee. *Smooth manifolds*. Springer, 2003.