Self-dual Yang-Mills and the Nahm transform

Josh Cork*

November/December 2020

These notes are intended as supplementary material for the lectures on self-dual Yang-Mills and the Nahm transform, delivered online as part of the London Mathematical Society lectures on topological solitons. The notes follow the same structure as the video lectures, but contain some clarifications, additional exercises, and references for further reading and study. If there are any comments, corrections, or suggestions, please feel free to email me at the address at the bottom of this page.

1 Lecture 1: Yang-Mills theory, and instantons

1.1 Introduction

The Yang-Mills equations are a particularly nice set of partial-differential-equations, with both mathematical and physical interest. A special case of the Yang-Mills equations are the self-dual Yang-Mills equations, and their finite-energy solutions are called *instantons*. Their physical importance (and original motivation) lies in quantum field theory, where the instantons are seen as pseudo-particles which describe tunnelling trajectories between classically forbidden regions. A good discussion of this physical understanding may be found in [CL84, SS98]. Another physical application is found in holographic QCD, where instantons are seen to approximate nucleons [AM89, SS05, Sut10]. Later, mathematicians realised that there is a natural interpretation of the equations geometrically, enabling a more general understanding of the topic mathematically. For example, mathematically this tunnelling effect may be regarded as saying that instantons are the gradient flows between critical points of the Chern Simons functional. An aim of this lecture is to build the mathematical machinery needed to understand the self-duality equations, and study some qualitative properties. The upshot will be that the space of solutions to the equations is intrinsically linked to the topology of space on which they are defined. For this reason, instantons are often thought of as *topological solitons* [MS04].

*email: joshua.cork@itp.uni-hannover.de

Mathematically, the study of the spaces of instantons has been a catalyst for various major advances in geometry, and continues to be an important area of research combining aspects of geometry, topology, analysis, and physics. One the most famous uses of instantons mathematically was how their spaces were exploited by Donaldson in his pioneering work on invariants for four-manifolds, which led to him being awarded with a Fields Medal. For a complete story of this, a good reference is his book with Kronheimer, *the geometry of four manifolds* [DK97].

Before we proceed, a disclaimer must be given. A lot of the material can be very technical, and although we aim to keep technical detail to a minimum, some assumptions will be made. In particular, we assume a basic understanding of differentiable manifolds, riemannian (and pseudo-riemannian) geometry, topology and homotopy, and differential forms. The interested reader is encouraged to work through the details themselves, and to review any basic definitions where necessary. Some good references for these topics, which also include more details on the preliminary topics discussed in this lecture regarding vector bundles and Chern numbers are [dC92, GS89, Hat02, Spi70, Tau11].

1.2 The Yang-Mills equations and gauge theory

In order to gain a full appreciation of the self-dual Yang-Mills equations, it is worth first starting with a more general set of equations called simply *the Yang-Mills equations*. In short, the Yang-Mills equations are a non-abelian generalisation of the vacuum Maxwell's equations, applicable to any n-dimensional pseudo riemannian manifold (M, η). The equations can be written in a concise form as

$$D^A \star F^A = 0. \tag{1}$$

Now, whenever you write down an equation, it is of course useful to know what all the individual symbols actually mean. A geometric understanding is given in the next section, but it's possible to understand these equations just in terms of relatively simple objects too. The fundamental ingredient here is the object A, which is a matrix-valued 1-form called a *gauge field*. Crucially, the matrix components A_μ (with respect to some local coordinates on $U \subset M$) take values in the Lie algebra $\mathfrak{g}$ of a Lie group, i.e. $A_\mu : U \longrightarrow \mathfrak{g}$. In this way, A is known as a G gauge field. For example, when the Lie group $G = U(n)$ is unitary, then A_μ are anti-hermitian matrices.

The gauge field A defines the objects D^A and F^A. The first is called the *covariant derivative*, which is defined as the ordinary derivative plus the gauge field $D^A = \mathrm{d} + A$. In components this reads $D^A_\mu = \partial_\mu + A_\mu$. The second, F^A, is called the *gauge field strength*, and is a 2-form defined as the Lie bracket (commutator) of the covariant derivative with itself

$$F^A = [D^A, D^A] = \mathrm{d}A + A \wedge A. \tag{2}$$

Note that since A is matrix-valued, it doesn't behave like a normal 1-form, and the term $A \wedge A$ vanishes if and only if G is abelian. This is seen more clearly by viewing F^A in components

Exercise 1. *By evaluating $[D^A, D^A]$ locally on a function ϕ via $D^A_\mu \phi = \partial_\mu \phi + A_\mu \phi$, show that*

$$F^A_{\mu\nu} = [D^A_\mu, D^A_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{3}$$

How does this formula change if you expand A in a non-coordinate basis of 1-forms as $A = \sum_\mu A_\mu\, e^\mu$? [HINT for second part: you will need to consider a dual basis for e^μ and construct a formula for $\mathrm{d}e^\mu$.]

The final ingredient to understand in these equations is the symbol $\star$. This is an operator which sends k-forms to $(n-k)$-forms called the *Hodge star isomorphism*. So far everything in equation (1) has been independent of the metric on the manifold M, but this part explicitly depends on the metric. On a general pseudo-riemannian manifold (M, η) let $\mathrm{Vol}_\eta \in \Lambda^n(T^*M)$ be the volume form induced by the metric η.[1] Then for $u \in \Lambda^k(T^*M)$ its *Hodge dual* $\star u \in \Lambda^{n-k}(T^*M)$ is defined such that for any $v \in \Lambda^k(T^*M)$,

$$v \wedge \star u = \widetilde{\eta}(u, v)\ \mathrm{Vol}_\eta, \tag{4}$$

where $\widetilde{\eta}$ is the induced metric on differential forms.[2] For practical purposes, it is helpful to understand the Hodge star in terms of a choice of coordinates[3] $(x^1, \ldots, x^n)$ for M. Explicitly, if $\alpha \in \Lambda^k(T^*M)$ is given in totally anti-symmetric components as

$$\alpha = \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^{n} \alpha_{i_1 \cdots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}, \tag{5}$$

then $\star\alpha$ has the coordinate representation displayed

$$(\star\alpha)_{i_{k+1}\cdots i_n} = \frac{\sqrt{|\det \eta|}}{(n-k)!} \sum_{i_1, j_1 \ldots, i_k, j_k = 1}^{n} \eta^{i_1 j_1} \cdots \eta^{i_k j_k} \alpha_{j_1 \cdots j_k} \varepsilon_{i_1 \cdots i_n}, \tag{6}$$

where η^{ij} is the inverse of the matrix for η in the coordinate frame, i.e the inverse of the matrix with components $\eta_{ij} = \eta\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$.

[1] In local coordinates $\{x^i\}$ this is simply the standard volume form $\mathrm{d}^n x$ scaled by the root of the absolute determinant of η.

[2] Every 1-form u has an associated vector field X^u defined by the metric so that $u(Y) = \eta(X^u, Y)$ for all Y. Thus the induced metric on 1-forms is simply defined as $\widetilde{\eta}(u, v) = \eta(X^u, X^v)$. This is naturally extended to k-forms by expressing the k-form in terms of wedge products of a basis of 1-forms.

[3] The same formula applies if you express α in terms of a non-coordinate basis of 1-forms, but the correct adjustments need to be made to the metric in the non-coordinate basis.

Exercise 2. *Let s be the parity of the signature of η (that is 0 or 1 if the determinant is positive or negative). Repeated operation of $\star$ gives a well-defined map $\star^2 : \Lambda^k(T^*M) \to \Lambda^k(T^*M)$. Show that this map satisfies $\star^2 = (-1)^{k(n-k)+s}\mathbb{1}$, where $\mathbb{1}$ is the identity map, and $n = \dim(M)$.*

Now that we have a little more understanding of the objects in the Yang-Mills equations $D^A \star F^A = 0$, it is a good exercise to see why these are considered as a 'non-abelian generalisation to Maxwell's equations'.

Exercise 3. *Consider a $U(1)$ gauge field on Minkowski space $\mathbb{R}^{1,3}$, that is a 1-form $A = \imath a$, with $a = \sum_{\mu=0}^{3} a_\mu \, \mathrm{d}x^\mu$, with $a_\mu : \mathbb{R}^{1,3} \to \mathbb{R}$. Here the metric should be viewed as $\eta = \mathrm{diag}\{-1, 1, 1, 1\}$.*

1. *Show that $D^A F^A = 0$, and write down the Yang-Mills equations $D^A \star F^A = 0$ in terms of the fields a_0, a_1, a_2, a_3.*
2. *Denote the electric and magnetic fields $\vec{E}$ and $\vec{B}$ as the vectors $\vec{E} = \mathrm{grad}(a_0) - \partial_t \vec{a}$, and $\vec{B} = \mathrm{curl}(\vec{a})$, where $\vec{a} = (a_1, a_2, a_3)$. Thus show that the equation $D^A F^A = 0$, and the Yang-Mills equations take the form of the vacuum Maxwell's equations*

$$\begin{aligned} \mathrm{div}(\vec{E}) = 0, \quad \mathrm{curl}(\vec{E}) = -\partial_0 \vec{B}, \\ \mathrm{div}(\vec{B}) = 0, \quad \mathrm{curl}(\vec{B}) = \partial_0 \vec{E}. \end{aligned} \tag{7}$$

1.2.1 The geometric description – gauge theory and vector bundles

The Yang-Mills equations appeared first in physics as a way to generalise Maxwell's equations to non-abelian gauge fields, and this has important interpretations in quantum field theory, in particular for describing the strong nuclear force (there $G = SU(3)$). However, mathematicians realised later that these objects have a very natural geometric interpretation, which in turn allows for a deeper understanding of the possible solutions to the Yang-Mills equations. The interpretation relies on the notion of a G-bundle over a manifold M, and F^A is a local description of the *curvature* of this object. I will now give a brief exposition of this geometric interpretation.

Vector bundles. The most convenient way to understand this is via the language of *connections* on *vector bundles* with *structure group* G.[4] A vector bundle E of *rank* p is a smooth assignment $\pi : E \to M$ of a p-dimensional (real or complex) vector space V to

[4] Although convenient for our discussions, it is worth mentioning that for a deeper understanding, it is better to talk about connections as a choice of horizontal distribution on a principal G bundle, but for that I refer to [Tau11], where this, along with many other equivalent notions is discussed. Our description is equivalent to considering the vector bundle E as the associated bundle $P \times_\rho V$ of a principal bundle.

each point $x \in M$, which is locally trivial, by this we mean on a neighbourhood $\mathcal{U}_x$ of x, $\pi^{-1}(\mathcal{U}_x)$ is diffeomorphic to $\mathcal{U}_x \times V$ – this diffeomorphism is called a 'local trivialisation'. The *total space* E of the bundle is then stitched together over all of M by its *transition functions*, which are a family of smooth functions $g_{ij} : U_i \cap U_j \to G$, where G is a group which acts on V called the **structure group**, on each intersection of a cover of M, such that $g_{ii} = \mathbb{1}$, $g_{ji} = g_{ij}^{-1}$,[5] and when restricted to the triple overlaps $U_i \cap U_j \cap U_k$, they satisfy the cocyle condition

$$g_{ik} = g_{ij} g_{jk}. \tag{8}$$

The total space E may be identified as the space of equivalence classes:

$$E = \bigcup_i U_i \times V /_{\sim}, \quad \text{where} \quad (x, v) \sim (x, (\rho \circ g_{ij}(x))(v)), \tag{9}$$

on $U_i \cap U_j$, where $\rho : G \to GL(V)$ is a faithful representation of G, in order for the action above to make sense. Typically G is a matrix group inside the set of $p \times p$ matrices, and this action is then simply left multiplication. The equivalence classes of pairs (x, v) are known as the **fibers** of E, denoted E_x. It is worth noting that this is not the standard way to define a vector bundle – the construction of vector bundles via transition functions does not give all vector bundles in the strict sense of the standard definition (see e.g. [Spi70]). However, there is a natural notion of isomorphism of vector bundles, and this construction gives all vector bundles up to isomorphism, which is sufficient for the purpose of our discussion. Furthermore, the understanding in terms of transition functions allows for a clearer understanding of topics later on.

Sections. The most elementary objects one defines on a vector bundle are *sections*, which are a generalisation of a vector field. These are smooth functions $\psi : M \to E$ such that $\psi(x) \in E_x$ for all $x \in M$. Locally, these are represented as $\psi(x) = (x, v(x))$, where $v : U \to V$, and these local sections are related on overlapping patches via the transition functions. A section is a natural object to study physically, for example, the wave function in quantum mechanics is best thought of as a section of a rank 1 complex vector bundle.

Gauge transformations. For each point $x \in M$ the vector $v(x)$ can be represented with respect to a choice of basis of V by a column vector, and a smooth change of basis corresponds to a choice of smooth function $g : U \to G$ acting via $v(x) \mapsto g(x)v(x)$. This smooth function is called a *gauge transformation*, and the choice of basis for each fiber $E_x \cong \{x\} \times V$ is known as a *gauge choice*. The local gauge transformation may be

[5] This notation may be confusing. Here, and throughout, we use the standard notation that for a map $g : U \to G$, sending x to $g(x) \in G$, g^{-1} is the map $g^{-1} : U \to G$ identified by $g(x)^{-1}$, that is, the inverse element in G of $g(x)$.

understood globally by insisting that for any pair $g_i : U_i \to G$ and $g_j : U_j \to G$, they are related on $U_i \cap U_j$ via the transition functions as

$$g_i = g_{ij} g_j g_{ij}^{-1}. \tag{10}$$

The set of all tuples of local gauge transformations satisfying this rule forms a group $\mathcal{G}$ called the *gauge group*, and these form the (smooth) isomorphisms of the vector bundle. In the case of a trivial bundle, that is, where the transition functions are all identity, the gauge group is simply $\mathcal{G} = C^\infty(M, G)$. Warning: sometimes the structure group G is also called the gauge group, so it is important to be aware of this confusing multiple use of terminology.

Connections. It makes sense both from a mathematical and physical perspective to measure as much as possible independently of a choice of basis. This is known in the present context as *gauge invariance*. In short, a connection provides a gauge-invariant way to differentiate sections, namely via the covariant derivative $D^A = \mathrm{d} + A$ introduced earlier, where A is the gauge field – a 1-form taking values in the Lie algebra $\mathfrak{g}$ of G. This is understood by imposing that $D^A v$ transforms to $gD^A v$.

Exercise 4. *Show that this holds if and only if A transforms as*

$$A \mapsto gAg^{-1} - \mathrm{d}gg^{-1}. \tag{11}$$

Note that this certainly isn't generally the case for the ordinary derivative d, unless your gauge transformations are constant everywhere. The connection may be extended to an action on k-forms, sending k-forms to $(k+1)$-forms via the action

$$D^A \omega = \mathrm{d}\omega + A \wedge \omega + (-1)^{k+1} \omega \wedge A, \tag{12}$$

which allows for an understanding of the Yang-Mills equation, since $\star F^A$ is an $(n-2)$-form.

The connection is defined generally from a collection of local gauge fields $\{A^i\}$ on the chart neighbourhoods U^i of M. These local one-forms are related on overlapping patches by the transition functions:

$$A^i = g_{ij} A^j g_{ij}^{-1} - \mathrm{d}g_{ij} g_{ij}^{-1}. \tag{13}$$

Generally one drops all indices referring to the local set U_i, with the understanding that the gauge field is defined locally. The explicit reference is only made when relating the gauge field on different patches.

Curvature. The final ingredient in understanding the objects in the Yang-Mills equations geometrically is the field strength $F^A = \mathrm{d}A + A \wedge A$. In this context, F^A is known as the *curvature* of the connection. The reason this measures curvature may be explained roughly by noting that it is formed as a second derivative, much like how the curvature of a curve is described by its second derivative[6]. Different connections can have different curvature – one you may be familiar with is the Levi-Civita connection on the tangent bundle, and the Riemann curvature. F^A is a less strict version of this applied to general vector bundles. Again, F^A is not a global object, but is stitched together by the local curvatures F^{A^i} via the transition functions.

Exercise 5. *Show that on $U_i \cap U_j$, the local curvatures F^{A^i} and F^{A^j} are related by the transition function g_{ij} via*

$$F^{A^i} = g_{ij} F^{A^j} g_{ij}^{-1}. \tag{14}$$

1.2.2 The Yang-Mills equations as a variational problem.

So what do the Yang-Mills equations mean in this geometric context? The energy of A, or *Yang-Mills* action, given by the total curvature

$$S_{\mathrm{YM}}[A] = ||F^A||^2_{L^2} = \int_M \mathrm{tr}\,(F^A \wedge \star (F^A)^\dagger). \tag{15}$$

In general the L^2 inner product for $\mathfrak{g}$-valued forms η, ξ on $\mathcal{U} \subset M$ is

$$\langle \eta, \xi \rangle_{L^2} := \int_{\mathcal{U}} \mathrm{tr}\, \left(\eta \wedge \star \xi^\dagger \right). \tag{16}$$

Note that we are implicitly assuming for simplicity that G is a matrix group, and shall from now on restrict further to the case of unitary groups, but this can be understood in further generality. Note also that the action is well-defined on M due to how the connection is patched together by the transition functions. Finally, it is worth mentioning that this inner product is only positive definite in the case of riemannian manifolds, and furthermore in the lorentzian case one replaces this by its negative so as to really think of it as an *action*. Indeed, one can view (15) in riemannian signature as the *static energy* of an object defined on $X = \mathbb{R} \times M$, where the additional 'time' component of X gives X lorentzian signature, and the action for fields on X is minus that given in (15), integrating over all of X. The 'kinetic' term T is then realised as the integral over M of all terms involving time derivatives, and the 'potential' term V is the integral over M of everything else. Then the action really is $S = \int (T - V)\, \mathrm{d}t$, as one expects from newtonian mechanics.

[6]Again, for a different perspective in terms of how a bundle is "curved" by a choice of connection seen as a horizontal subspace of the tangent bundle, see [Tau11].

Exercise 6. *Show that the co-connection $(D^A)^* = (-1)^{n(k-1)+1} \star D^A \star : \Lambda^k \to \Lambda^{k-1}$ is the L^2-adjoint of the connection D^A, that is*

$$\langle D^A u, v\rangle = \langle u, (D^A)^* v\rangle \tag{17}$$

for all compactly supported forms u, v.

The extrema of S_{YM} are determined by calculus of variations, i.e. by solving

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} S_{\mathrm{YM}}[A + \varepsilon\alpha] = 0, \tag{18}$$

for all smooth, compactly-supported variations $\alpha \in \Lambda^1(T^*M, \mathfrak{g})$ (this notation simply means they are $\mathfrak{g}$-valued 1-forms).

Exercise 7. *Show that the equation obtained in this way is*

$$(D^A)^* F^A = 0, \tag{19}$$

and that this is equivalent to the Yang-Mills equation.

In other words, any minima of the Yang-Mills action are necessarily solutions of the Yang-Mills equation! This can be interpreted as saying, for a given bundle E, the closest connections to being flat, namely $F^A = 0$, necessarily satisfy the Yang-Mills equations. However, as we shall see shortly, there are topological obstructions to the existence of flat connections on a given bundle, and this will be seen via the concept of an instanton.

Finally, the space $\mathcal{A}$ of solutions A to the Yang-Mills equations is called the space of *Yang-Mills gauge fields* or *Yang-Mills connections*. However, we don't want to count two gauge fields which differ by gauge transformations as different, thus one typically cares only about the gauge-equivalence classes of Yang-Mills gauge fields $\mathcal{M} = \mathcal{A}/\mathcal{G}$. $\mathcal{M}$ is known as the *moduli space of Yang-Mills connections*.

1.3 Self-duality

Recall that a crucial element of the Yang-Mills equations is the Hodge star isomorphism $\star$, for which so far we've only given a formal definition. Geometrically $\star$ may be thought of as a correspondence between orthogonally complementary subspaces of an n-dimensional inner-product space, where each space has a natural choice of orientation, i.e. an ordered basis – this is the role of the volume form in the definition. For example, on $\mathbb{R}^3$ with the standard 'dot-product' metric, the vector cross product gives a duality between planes and their normal vectors, with orientation dictated by the right-hand-rule. This may be understood as a special case of the Hodge-star duality. Specifically

Exercise 8. *Let $U, V \in \mathbb{R}^3$ be vectors which span a plane, and u and v be their corresponding covectors defined by $u(X) = U \cdot X$ and $v(X) = V \cdot X$ for all $X \in \mathbb{R}^3$. Then show that $\star(u \wedge v)$ is the covector to $U \times V$, i.e.*

$$\star(u \wedge v)(X) = (U \times V) \cdot X, \quad \text{for all } X \in \mathbb{R}^3. \tag{20}$$

This duality between k-forms and $(n-k)$-forms has an important consequence for k-forms on manifolds where $n = 2k$, i.e. forms of degree half the dimension of M. First note that it is always true that $\star^2 = \pm 1$, where the $\pm$ is determined by the degree of the forms $\star$ is acting on, the dimension of M, and the signature of the metric (see exercise 2). When the parity of the signature is 1 ($\det \eta > 0$), and k is even (i.e. the manifold has dimension a multiple of 4), $\star^2 = 1$, and thus is an involution of $\Lambda^k(T^*M)$ with eigenvalues ± 1. This means $\Lambda^k(T^*M)$ has the orthogonal decomposition

$$\Lambda^k(T^*M) = \Lambda^k_+(T^*M) \oplus \Lambda^k_-(T^*M). \tag{21}$$

The subspaces are known as the self- and anti-self-dual k-forms, i.e. every k-form u on a $2k$-dimensional manifold, with k even, can be written as

$$u = u_+ + u_-, \quad \text{with} \quad u_\pm = \frac{1}{2}(u \pm \star u), \tag{22}$$

where $\star u_\pm = \pm u_\pm$, i.e. are self/anti-self-dual.

Exercise 9. *Show that these are real linear subspaces of $\Lambda^k(T^*M)$.*

The first instance where this can occur is on four-manifolds (M, η) with even parity, where the 2-forms split into self-dual and anti-self-dual parts. But this case is interesting for the Yang-Mills equations. Recall that the curvature F^A is a 2-form. Therefore here one can talk about *self-* or *anti-self-dual* connections, that is, where the curvature 2-form satisfies

$$\star F^A = \pm F^A, \tag{23}$$

or equivalently $F^A_\pm = 0$ for a choice of $\pm$.

This is important for studying the Yang-Mills equations because they are automatically satisfied by self- or anti-self-dual connections. This is because in this sector, the Yang-Mills equation becomes

$$D^A F^A = 0. \tag{24}$$

Exercise 10. *Show that equation (24) holds for all A.*

This equation is therefore an intrinsic property of connections, known as the *Bianchi identity*. The equations $\star F^A = \pm F^A$ are known as the *self-dual Yang-Mills equations*. In fact, the choice of $\pm$ is ultimately irrelevant, corresponding to a change of orientation on M, but the modern preference is to talk about *anti-self-dual* gauge fields.

1.3.1 Example

To get a feeling for the self-duality equations, let's consider an example on $\mathbb{R}^4$ with the standard flat euclidean metric. We also need a structure group, and the simplest non-abelian example is the special unitary group $SU(2)$. It's also the first non-trivial case:

Exercise 11. *Show that there are no solutions to the self-duality equations on euclidean $\mathbb{R}^4$ with structure $G = U(1)$.*

The Lie algebra $\mathfrak{su}(2)$ of $SU(2)$ is represented by the set of traceless, anti-hermitian 2×2 complex matrices, however it is best thought of as the set of imaginary quarternions, i.e. if $X \in \mathbb{H}$, then $\mathrm{Im}\,(X) \in \mathfrak{su}(2)$ takes the form

$$\mathrm{Im}\,(X) = X^1\boldsymbol{i} + X^2\boldsymbol{j} + X^3\boldsymbol{k}, \quad X^j \in \mathbb{R}. \tag{25}$$

The idea here is that we have identified the generators τ^j of $\mathfrak{su}(2)$, given by the matrices

$$\tau^1 = \begin{pmatrix} 0 & -\imath \\ -\imath & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} -\imath & 0 \\ 0 & \imath \end{pmatrix}, \tag{26}$$

as $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$ respectively.

Exercise 12. *Verify that these satisfy the quarternionic relations*

$$\tau^i\tau^j = -\delta^{ij} + \varepsilon_{ijk}\tau^k$$

for all i, j.

From this, we can consider the following canonical 1-forms on $\mathbb{H}$

$$\mathrm{d}x = \mathrm{d}x^1\boldsymbol{i} + \mathrm{d}x^2\boldsymbol{j} + \mathrm{d}x^3\boldsymbol{k} + \mathrm{d}x^4, \tag{27}$$
$$\mathrm{d}\overline{x} = -\mathrm{d}x^1\boldsymbol{i} - \mathrm{d}x^2\boldsymbol{j} - \mathrm{d}x^3\boldsymbol{k} + \mathrm{d}x^4, \tag{28}$$

whose imaginary parts are $\mathfrak{su}(2)$-valued. From these we can obtain examples of self- and anti-self-dual 2-forms.

Exercise 13. *Show that the two-forms*

$$\mathrm{d}x \wedge \mathrm{d}\overline{x}, \quad \text{and} \quad \mathrm{d}\overline{x} \wedge \mathrm{d}x \tag{29}$$

are both completely imaginary, i.e. $\mathfrak{su}(2)$-valued, and in particular, are anti-self-dual and self-dual respectively.

Thus, any 2-form proportional to $\mathrm{d}x \wedge \mathrm{d}\overline{x}$ is anti-self-dual. However, note that all we have found is a sufficient condition for the *curvature* 2-form of a connection to be self or anti-self-dual. What we really want is how to represent a gauge field whose curvature is of this form. Although, due to linearity, it is sufficient to find something where the curvature is proportional to $\mathrm{d}x \wedge \mathrm{d}\overline{x}$ by some real-valued function $\phi : \mathbb{R}^4 \to \mathbb{R}$.

To see how to obtain this, first note that any $x \in \mathbb{R}^4$ can be thought of equivalently as an element of the quarternions by the canonical isomorphism

$$x = x^1 \boldsymbol{i} + x^2 \boldsymbol{j} + x^3 \boldsymbol{k} + x^4. \tag{30}$$

So then any $\mathfrak{su}(2)$-valued 1-form on $\mathbb{R}^4$ can be represented in the form

$$A = \mathrm{Im}\,(f(x)\,\mathrm{d}\overline{x}), \tag{31}$$

where $f : \mathbb{H} \to \mathbb{H}$, and Im takes the imaginary part of the resulting $\mathbb{H}$-valued 1-form. The choice of $\mathrm{d}\overline{x}$ is just for convenience here, and you can equivalently choose $\mathrm{d}x$ and replace f by $-\overline{f}$. This has curvature

$$F^A = \mathrm{Im}\,(\mathrm{d}f \wedge \mathrm{d}\overline{x} + f(x)\,\mathrm{d}\overline{x} \wedge f(x)\,\mathrm{d}\overline{x}). \tag{32}$$

Note that the second term does not necessarily vanish as the quarternions are non-commutative. It is not so easy to derive an explicit expression for a general f so that one obtains something anti-self-dual, so in order to get some solutions, one can impose a much stricter condition on the function f. For example, one may ask that the gauge field is symmetric with respect to some isometry of $\mathbb{R}^4$. A simple example is to impose $SO(4)$-invariance, and a suitable $SO(4)$-symmetric ansatz is given by

$$f(x) = g(r)x, \tag{33}$$

where $r = |x| = \sqrt{x\overline{x}}$ is the length of $x \in \mathbb{R}^4$, and $g : (0, \infty) \to \mathbb{R}$ only depends on this radial coordinate.

Exercise 14. *Show that the curvature of this connection takes the form*

$$F^A = \left(g + \frac{rg'}{2}\right) \mathrm{d}x \wedge \mathrm{d}\overline{x} + \left(\frac{g'}{2r} + g^2\right) \mathrm{Im}\,(x\,\mathrm{d}\overline{x} \wedge x\,\mathrm{d}\overline{x}).$$

[HINT: You will need to represent $\mathrm{d}r$ in terms of $x, \overline{x}, \mathrm{d}x$, and $\mathrm{d}\overline{x}$.]

Imposing that F^A is proportional to $\mathrm{d}x \wedge \mathrm{d}\overline{x}$, and thus anti-self-dual, therefore requires that g satisfies the differential equation

$$g' + 2rg^2 = 0. \tag{34}$$

This is separable, with the full family of solutions given by

$$g(r) = \frac{1}{r^2 + \rho}, \tag{35}$$

where $\rho \in \mathbb{R}$ is a constant.

Exercise 15. *Show that up to gauge equivalence it suffices to choose $\rho > 0$, thus removing the apparent singularity at $x = 0$ for the choice $\rho = 0$.*

Thus we may represent ρ by λ^2, for some $\lambda > 0$. Note that the same argument would have followed if we replace x by an arbitrary translation, $x - a$, for some $a \in \mathbb{H} \cong \mathbb{R}^4$. This leads to a five-parameter family of anti-self-dual connections, with gauge field represented by the $\mathfrak{su}(2)$-valued 1-form

$$A = \operatorname{Im}\left(\frac{x-a}{|x-a|^2+\lambda^2}\,\mathrm{d}\overline{x}\right). \tag{36}$$

1.4 Instantons

The $SU(2)$ gauge field (36) was the first example of what is now known as an **instanton**, and it was first discovered by Belavin, Polyakov, Schwartz, and Tyupkin, hence is known as the *BPST instanton* [BPST75]. An instanton is simply an anti-self-dual connection for which the Yang-Mills action is finite. Note that the condition of finite action is not trivial – it is possible for the action to diverge for non-compact manifolds like $\mathbb{R}^4$, so suitable boundary conditions often need to be imposed. In the case of the BPST instanton, the curvature is

$$F^A = \frac{\lambda^2}{(|x-a|^2+\lambda^2)^2}\,\mathrm{d}x \wedge \mathrm{d}\overline{x}, \tag{37}$$

which behaves like $|x|^{-4}$ as $|x| \to \infty$, which is a necessary condition for finite action. In turn, this connection does in fact exhibit finite action.

Exercise 16. *Show that for the BPST instanton, the action is given by $S_{\mathrm{YM}}[A] = 8\pi^2$ for all choices of $a \in \mathbb{H}$ and $\lambda > 0$.*

In general on $\mathbb{R}^4$, stricter boundary conditions need to be imposed, but this we shall examine a little bit later. However, it turns out that the value obtained for the action is due to something topological, which in turn is related to the condition of finite action.

For the time-being let's talk more generally (although in turn, more restrictively) and consider the anti-self-dual connections on closed riemannian manifolds M, that is, compact manifolds without boundary, with positive metric signature. In this case finite action is guaranteed, so all ASD connections are instantons. More importantly, instantons are exactly the global minima[7] of the Yang-Mills action! To see why, consider the following

Exercise 17. *Show that $F^A_\pm$ are orthogonal to each other with respect to the L^2 inner product.*

[7] It makes sense to talk about minima here as the Yang-Mills action is positive.

Thus

$$S_{\rm YM}[A] = ||F_+^A||_{L^2}^2 + ||F_-^A||_{L^2}^2. \tag{38}$$

Moreover,

$$||F_+^A||_{L^2}^2 - ||F_-^A||_{L^2}^2 = \int_M \mathrm{tr}\,(F^A \wedge F^A), \tag{39}$$

and therefore we may write

$$S_{\rm YM}[A] = 2||F_\pm^A||_{L^2}^2 \mp \int_M \mathrm{tr}\,(F^A \wedge F^A). \tag{40}$$

The quantity

$$Q = \frac{1}{8\pi^2}\int_M \mathrm{tr}\,(F^A \wedge F^A) \tag{41}$$

is an invariant of the vector bundle, known as the Yang-Mills *topological charge*, and crucially is independent of the choice of connection A. Thus, the global minima are precisely the connections for which one of $||F_\pm^A||^2$ vanishes, i.e. the instantons. Note that it is not necessarily true that all *local* minima are instantons, although it is true in certain restrictive cases, for example when the structure group is $SU(2)$, and M is a homogeneous space (e.g. a sphere) [Ste10]. It is still an important open problem to classify all local minima of the Yang-Mills action in full generality. See [Bec20] for recent developments.

1.4.1 The topological charge in detail – Chern and Chern-Simons forms

So why is Q an invariant of the vector bundle? The key to this lies in Chern-Weil theory, which says that the 4-form

$$C_2(F^A) = \frac{1}{8\pi^2}\mathrm{tr}\,(F^A \wedge F^A) \tag{42}$$

is closed, and the difference $C_2(F^A) - C_2(F^{A'})$ is globally exact, for any two gauge fields A and A'. The latter of these means that, for every pair of gauge fields A and A', there exist globally defined 3-form $Y_3(A, A')$, called the *Chern-Simons form*, such that $C_2(F^A) - C_2(F^{A'}) = \mathrm{d}Y_3(A, A')$. Thus, for closed manifolds M, by Stokes' theorem we have

$$\int_M C_2(F^A) - \int_M C_2(F^{A'}) = \int_M \mathrm{d}Y_3(A, A') = \int_{\partial M} Y_3(A, A') = 0, \tag{43}$$

where the last integral is zero as $\partial M = \emptyset$. Thus Q, which is the integral of C_2, is the same regardless of the choice of gauge field, and so provides an invariant for the bundle. As may be evident from the notation, this C_2 is one of a whole sequence of $2i$-forms called the *Chern forms*, and the integrals of C_k are likewise invariants of the bundle, called the Chern numbers. Far more remarkably, the Chern numbers are always an integer! The topological charge Q provides a topological invariant for the bundle, but since instantons are global minima of the Yang-Mills action, this invariant is often thought of as an invariant for the instanton too.

Since the Chern forms are closed, they must be locally exact.

Exercise 18. *Show that for C_2, locally we have*

$$C_2(F^A) = \frac{1}{8\pi^2}\mathrm{d}\mathrm{tr}\left(\mathrm{d}A \wedge A + \frac{2}{3}A \wedge A \wedge A\right), \tag{44}$$

i.e. the Chern-Simons 3-form takes the above local form.

One only obtains a globally exact object by taking the difference for two connections.

In general, $SU(n)$ instantons on any closed manifold M, have 'topological charge' Q exactly determined by (minus) the second Chern number. In this context, the negative second Chern number is known as the *instanton number*, and furthermore, the value of the action is given by $S_{\mathrm{YM}}[A] = 8\pi^2|Q|$.

1.5 Instantons on $\mathbb{R}^4$

So how do we reconcile the result about Chern numbers above to the case of $\mathbb{R}^4$. In general, for a non-closed manifold all of this is not true, so one has to impose boundary conditions on the manifold in order to make it look like one which is closed. Let's consider this in detail in the case of $\mathbb{R}^4$.

The usual way to ensure finite action is to impose that $F^A \to 0$ as $|x| \to \infty$ for $x \in \mathbb{R}^4$. One can make sense of this by viewing $\mathbb{R}^4$ as the interior of a 4-ball B^4, and considering the "at infinity" condition to mean restricting to the boundary S^3_∞ – known as the 3-sphere at infinity. The condition $F^A \to 0$ means the connection is flat on S^3_∞, which due to some sophisticated results in differential geometry, implies $A_\infty = -\mathrm{d}g_\infty g_\infty^{-1}$, for some function $g_\infty : S^3_\infty \to SU(n)$. This "gauge transformation at infinity" is classified topologically by its representative in the third homotopy group $\pi_3(SU(n))$, which is isomorphic to the group of integers. To see why this is true, first think of the case $n = 2$. In this case, the manifold $SU(2)$ is diffeomorphic to S^3, so g_∞ is a map between 3-spheres. Any map $\phi : S^3 \to S^3$ has a 'winding number' which can be understood analogously to the winding number of a loop to higher dimensions, essentially measuring the number of times one wraps the 3-sphere around itself.[8] The case $n > 2$ follows from a powerful theorem that

[8] An alternative way to understand this is by counting zeros. Specifically, every map $g : S^3 \to S^3$ can be extended to a map $\widetilde{g} : \mathbb{R}^4 \to \mathbb{R}^4$, and the degree is the number of zeros of $\widetilde{g}$ inside S^3.

every map $g : S^3 \to G$ for a Lie group G containing $SU(2)$ is homotopic to a map of 3-spheres.

So this boundary condition gives instantons on $\mathbb{R}^4$ a topological classification, again by an integer. But it turns out that this integer is also related to a Chern number, but to see why, we need to look at things in a different way. The key is to impose that the instanton on $\mathbb{R}^4$ extends to an instanton on S^4. The reason this is possible relies on two important ideas. The first is that the one-point compactification $\mathbb{R}^4 \cup \{\infty\}$ is diffeomorphic to S^4. Specifically this is given by the (inverse) stereographic projection:

$$\begin{array}{cccc} \phi : & \mathbb{R}^4 \cup \{\infty\} & \longrightarrow & S^4 \\ & x = (x^1, \dots, x^4) & \mapsto & \dfrac{1}{|x|^2+1}\left(2x^1, \dots, 2x^4, |x|^2 - 1\right), \\ & \infty & \mapsto & (0,0,0,0,1), \end{array} \tag{45}$$

the idea being that every point in $\mathbb{R}^4$ may be uniquely determined by every point bar one on the sphere. To see this, consider a light shining from the North pole onto flat space below. In this way one identifies $\mathbb{R}^4$ as the southern hemisphere, and the point at ∞ as the north pole N. If we remove this point, then we can relate the metric on $\mathbb{R}^4$ to the metric on $S^4 \setminus \{N\}$.

Exercise 19. *Let $\eta_{\mathbb{R}^4}$ denote the standard metric on $\mathbb{R}^4$, and η_{S^4} denote the round metric on S^4 induced by the flat metric on $\mathbb{R}^5$ via its embedding into $\mathbb{R}^5$. Consider the (inverse) stereographic projection ϕ restricted to $\mathbb{R}$. Show that $\eta_{\mathbb{R}^4}$ is conformally equivalent to $\phi^* \eta_{S^4 \setminus \{N\}}$.*

This leads to the second important point, which is that the Hodge star behaves nicely under conformal rescalings of a metric. Indeed,

Exercise 20. *Show that if two metrics η_1 and η_2 over M are conformally equivalent, that is they are related by $\eta_1 = e^{\lambda}\eta_2$, then*

$$\star_{\eta_1} = e^{\frac{1}{2}(n-2k)\lambda} \star_{\eta_2}, \tag{46}$$

where k is the degree of the forms $\star$ acts on, and $n = \dim M$.

Thus, for $k = 2$ and $n = 4$, which is our case study, $\star$ is conformally invariant, meaning every instanton on S^4 gives rise to an instanton on $\mathbb{R}^4$. As it happens, every instanton on $\mathbb{R}^4$ arises in this way, but this relies on a sophisticated analytical result of Uhlenbeck from the 1980s [Uhl82]. Uhlenbeck's theorem also shows that the condition of finite action on $\mathbb{R}^4$ is exactly equivalent to the condition that the instanton extends to S^4 in this way. Thus, instantons on $\mathbb{R}^4$ are in one-to-one correspondence with instantons on S^4.

So how do we pick out the degree of the map g_∞ as the Chern number of a bundle over S^4? Well, firstly, g_∞ is related to a transition function on S^4. Vector bundles over S^4

are defined by a single transition function between the northern and southern hemispheres. Specifically, S^4 is covered by two charts U_N and U_S defined by removing a small closed neighbourhood of the south and north poles, and the transition function is a map

$$g_{NS} : U_N \cap U_S \longrightarrow SU(n). \tag{47}$$

The intersection $U_N \cap U_S$ has the homotopy type of S^3, and so g_{NS} is thus classified by an integer as before. The corresponding vector bundle over $\mathbb{R}^4 \cup \{\infty\}$ is the *pullback* of a bundle over S^4 via the inverse stereographic projection, defined by the transition function $\widetilde{g} : \mathbb{R}^4 \setminus \{0\} \to SU(n)$

$$\widetilde{g} = g_{NS} \circ \phi|_{\mathbb{R}^4\setminus\{0\}}. \tag{48}$$

The transition function $\widetilde{g}$ restricted to S^3_∞ is identified with g_∞ from above.

Now, to see how this relates to the Chern number requires a bit of calculation. Importantly, there is a nice formula for the 'degree' k of a map $g : S^3 \to G$ in terms of an integral:

$$k = \frac{1}{24\pi^2} \int_{S^3} \mathrm{tr}\,(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g). \tag{49}$$

Next note that the local charts U_N and U_S are oppositely oriented, with U_N positively oriented with respect to the standard volume form on S^4. Also, they both have boundaries diffeomorphic to S^3. So, using the local formula (44) from before, and Stokes' theorem, the contributions of $\mathrm{tr}\,(F^A \wedge F^A)$ on the northern and southern hemisphere give opposite signs in the integral, leading to

$$c_2(E, M) = \frac{1}{8\pi^2} \int_{S^3} \mathrm{tr}\,\left(A_N \wedge \mathrm{d}A_N - A_S \wedge \mathrm{d}A_S + \frac{2}{3}(A_N \wedge A_N \wedge A_N - A_S \wedge A_S \wedge A_S)\right),$$

where A_* is the gauge field on U_*. Next, recall that the gauge fields are related on $S^3 \sim U_N \cap U_S$ by $A_N = gA_S g^{-1} - \mathrm{d}g g^{-1}$, where $g = g_{NS}$ is the transition function.

Exercise 21. *Show that this gives*

$$Q = -\frac{1}{24\pi^2} \int_{S^3} \mathrm{tr}\,\left(g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g \wedge g^{-1}\mathrm{d}g\right) = -k, \tag{50}$$

which is (minus) the degree of g, the topological invariant introduced above.

[9] So now we can see that the boundary condition that the instanton on $\mathbb{R}^4$ has F^A zero at infinity is equivalent choosing a gauge with $A_S = 0$ in the above discussion, leaving $A_N = -\mathrm{d}\widetilde{g}\widetilde{g}^{-1}$ as before. Since the topological charge doesn't depend on the choice of connection, these connections on S^4 again are classified by the instanton number k.

Now, recall our concrete example, the BPST instanton. Because of what we just talked about, this must also have a topological charge, and this must be determined by instanton number $k = \pm 1$.

[9] Note that the sign here is really an artifact of a choice of orientation, so irrelevant.

Exercise 22. *Show that the BPST instanton (that is, the anti-self-dual solution) has instanton number* $k = 1$*, and the self-dual solution (found by sending* $x \mapsto \overline{x}$*) has instanton number* $k = -1$*.*

In general, writing down examples of arbitrary charge k instantons is not straightforward. One approach is to again assume certain symmetry; for instance there are charge k examples of instantons invariant under an action of $SO(3)$ on $\mathbb{R}^4$, described first by Witten [Wit77]. However, Witten's instantons, and the BPST instanton may be described within a more general formulation known as the Corrigan-Fairlie-'t Hooft [CF77, tH] ansatz, which generalises relatively easily to other manifolds beyond $\mathbb{R}^4$.[10] The rough idea is to rescale the metric on M via $\eta \mapsto \phi^2\eta$, and to locally think of the Levi-Civita connection on T^*M for the rescaled metric as a 1-form A with values in the Lie-algebra $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$. This construction gives rise to two $SU(2)$ gauge fields $A^{\pm}$ on $(M, \phi^2\eta)$, defined by this splitting, and the condition of anti-self-duality of A^+ (or self-duality of A^-) may be shown to be equivalent to the vanishing of the scalar curvature of $(M, \phi^2\eta)$. This is the same as requiring that ϕ is a harmonic function on M. See [Lan05] for more details.

1.5.1 Moduli spaces of instantons

Recall the moduli space $\mathcal{M}$ of Yang-Mills connections, namely the gauge-equivalence classes of solutions to the Yang-Mills equations. Restricting the moduli space of $SU(n)$ Yang-Mills connections on $\mathbb{R}^4$ to instantons yields, for each instanton number $k \in \mathbb{Z}$, the **moduli space of instantons on** $\mathbb{R}^4$. It is often convenient, however, to refer to what is known as the *framed moduli space*. The boundary conditions for instantons on $\mathbb{R}^4$ involve extending the gauge field to $\mathbb{R}^4 \cup \{\infty\}$, which in the bundle world means determining a choice of trivialisation at infinity. This choice of trivialisation is known as a *framing* for the bundle over $\mathbb{R}^4 \cup \{\infty\}$ at infinity.

In general a **framing** for an hermitian bundle E over M is an isomorphism[11]

$$f : E|_N \longrightarrow N \times \mathbb{C}^n, \tag{51}$$

where $N \subset M$ is a submanifold, i.e. a choice of trivial bundle over the submanifold N. In our case the framing is a choice of isomorphism

$$f_\infty : E_\infty \longrightarrow \mathbb{C}^n. \tag{52}$$

The group $\mathcal{G} = \{g : \mathbb{R}^4 \cup \{\infty\} \to SU(n)\}$ of gauge transformations acts on the framing via $f_\infty \mapsto g(\infty) f_\infty$. The ordinary moduli space doesn't see this change, since it doesn't

[10] Specifically, this works in the case of *anti-self-dual* manifolds, which are spin manifolds which anti-self-dual Weyl tensor.

[11] That is, for local sets $U \subset N$, a smooth map $F : U \times \mathbb{C}^n \to U \times \mathbb{C}^n$ determined by $F(x, v) = (x, \rho \circ g_F(x) v)$ for some $g_F : U \to U(n)$.

consider the framing as part of the instanton data. However, by considering instead the space to be the space of pairs (A, f) where A is an instanton, and f a framing at infinity, then two framings are equivalent if and only if they are fixed by the action of gauge transformations. Thus, the *framed moduli space*, which we denote by $\mathcal{I}_{k,n}(\mathbb{R}^4)$, is the space of all instantons modulo **framed gauge transformations**, namely $g \in \mathcal{G}$ such that $g \to \mathbb{1}$ as $|x| \to \infty$, which are those which fix the framing.

A theorem of Atiyah, Hitchin, and Singer (first formulated in the case of $G = SU(2)$, but easily generalised) [AHS78] states that the framed moduli space $\mathcal{I}_{k,n}(\mathbb{R}^4)$ of $SU(n)$ instantons has the structure of a connected complex $2kn$-dimensional manifold. Moreover, this manifold has a natural riemannian metric, which turns out to be hyperkähler, meaning that the *holonomy group* of the Levi-Civita connection on $T^*\mathcal{I}_{k,n}$ is contained in the symplectic group $\mathrm{Sp}(N)$ for some $N > 0$. By holonomy group here, we mean the group formed by the parallel transport around all loops based at a point $x \in X$ (which are all equivalent since the moduli space is connected).

For the BPST instanton, there are 5 real unframed moduli space parameters (four positions for $a \in \mathbb{R}^4$ and one scale $\lambda > 0$), and by performing a general 'unframed' gauge transformation, corresponding (up to gauge equivalence) to a constant element $q \in SU(2)$, one obtains the additional 3 real parameters needed to obtain the full moduli space, which has real dimension 8. For higher charges, the real dimension for $SU(2)$ instantons is $8k$. Unfortunately, the best one can achieve in the CF'tH ansatz is a $(5k+7)$-dimensional subspace of the total moduli space (note that there are some redundant parameters for $k = 1$ and 2, which may be accounted for by gauge transformations), but this falls short of the mark in general for $k > 2$. A complete ansatz for instantons on $\mathbb{R}^4$ is found via the ADHM construction, which falls under the category of a more general process called the Nahm transform, which is discussed in the next lecture.

1.6 Instantons beyond $\mathbb{R}^4$

Although everything has been set up relatively generally, we have so far only really discussed instantons in detail in the case of $M = S^4$ and $M = \mathbb{R}^4$, which are equivalent! There are of course significantly more four-manifolds (M, η) to choose from. For closed manifolds, the instantons are classified by the second Chern number of the bundle. However, as we have already experienced with the case of euclidean $\mathbb{R}^4$, instantons on non-closed manifolds require the imposition of boundary conditions in order to ensure finite action, and it is certainly not known what these boundary conditions look like for a general non-closed four-manifold (M, η), and this can also vary dramatically for different choices of non-flat metric.

1.6.1 Preview for the next lecture

An interesting class of instantons are those on the manifolds $T^m \times \mathbb{R}^{4-m} \cong \mathbb{R}^4/\mathbb{Z}^m$, with T^m the m-torus, $T^m = S^1 \times \cdots \times S^1$, for $m = 1, \ldots, 4$, equipped with the flat product metric induced from $\mathbb{R}^4$. These types of instantons can be loosely referred to as **periodic instantons**. A nice thing about periodic instantons is that they may be constructed explicitly via a process called the Nahm transform. However, all of these except for T^m are non-compact, thus they require asymptotic conditions to obtain finite action. In all of these cases, the asymptotic conditions can be very subtle, and are not as straightforward to describe in full generality as was the case for $\mathbb{R}^4$. These will be looked at in some detail in the case of $m = 1$, namely instantons on $S^1 \times \mathbb{R}^3$. In fact, we shall also see that it is possible to derive analogous notions of "self-duality" in lower dimensions by imposing translation invariance on the self-duality equations on $\mathbb{R}^4$.

1.7 Conclusion

In this lecture we have looked at some depth at gauge fields on four-manifolds, with an emphasis on those with anti-self-dual curvature, and their topology. The motivation for studying these gauge fields has been to understand the finite-action solutions to the self-dual Yang-Mills equations, however we have not yet gone into much detail about how to solve these PDEs. In the next lecture, we shall see a powerful process for constructing solutions to the self-duality equations.

2 Lecture 2: The Nahm transform

In the last lecture we looked at the geometry and topology of gauge fields on four-manifolds, motivated by the self-dual Yang-Mills equations. As a reminder, these are a set of non-linear partial-differential equations for the curvature 2-form, namely

$$\star F^A = -F^A, \tag{53}$$

and their finite-action solutions, that is, with

$$||F_A||_{L^2}^2 = \int_M \mathrm{tr}\,(F^A \wedge \star(F^A)^\dagger) < \infty \tag{54}$$

are called instantons. We have already studied in some detail what it means to be an instanton on flat euclidean $\mathbb{R}^4$, and have seen that instantons are classified by a topological integer k called the instanton number. Furthermore, we saw a large family of solutions in the case $k = 1$. In this lecture we shall turn our focus to constructing solutions to these equations in more generality, and to looking at a special family of manifolds, namely those of the form $\mathbb{R}^4/\Gamma$, where $\Gamma \subset \mathbb{R}^4$ is a group of translations acting on $\mathbb{R}^4$. These manifolds are known as *generalised tori*. They include examples such as the four torus T^4 (when $\Gamma = \mathbb{Z}^4$), the cylinder $S^1 \times \mathbb{R}^3$ (when $\Gamma = \mathbb{Z}$), and also lower-dimensional spaces such as $\mathbb{R}^3$ (when $\Gamma = \mathbb{R}$). Of course, the notion of anti-self-dual doesn't make sense on manifolds of dimension not 4, so to understand this here, the idea is that you really do have a connection defined on $\mathbb{R}^4$, but it is completely invariant in 1 (or more) of the coordinate directions. For example, on $\mathbb{R}^3$, one would consider the 4D connection

$$A = \sum_j A_j \,\mathrm{d}x^j + \Phi \,\mathrm{d}x^4, \tag{55}$$

where $A_j, \Phi : \mathbb{R}^3 \to \mathfrak{g}$, i.e. are invariant under changes in the direction x^4. We shall see how to perform this dimensional reduction in practice when we look at specific examples.

The process behind their construction is known generally as the Nahm transform, due to its description by Nahm [Nah83, Nah84] but this idea first appeared in the context of constructing instantons on $\mathbb{R}^4$, where it is known as the Atiyah-Drinfeld-Hitchin-Manin construction (ADHM for short) [ADHM78]. Since we have some understanding of instantons on $\mathbb{R}^4$, the first explicit example we shall see is this.

2.1 The Nahm transform heuristic

Before we study explicit examples, it is worth building a picture of the general process of the Nahm transform. There is no precise recipe, but there is a convenient heuristic treatment, which was originally reviewed in [Jar04]. The heuristic is simple: anti-self-dual connections on the generalised torus $\mathbb{R}^4/_{\Gamma}$ are in one-to-one correspondence with anti-self-dual connections on the *dual* generalised torus $(\mathbb{R}^4)^*/_{\Gamma^*}$, where here $\mathbb{R}^4$ is the dual space (the space of linear maps from $\mathbb{R}^4$ to $\mathbb{R}$), and Γ^* is the dual group, namely

$$\Gamma^* = \{\varphi \in (\mathbb{R}^4)^* \ : \ \varphi(g) \in \mathbb{Z}, \text{ for all } g \in \Gamma\}. \tag{56}$$

This can be understood by looking at table 1, which relates the spaces via transposition of the table. Recently, Charbonneau and Hurtubise filled in the last piece of the puzzle by establishing the Nahm transform between "spatially periodic instantons" i.e. on $T^3 \times \mathbb{R}$, and completely periodic *singular monopoles*[12], i.e. on T^3 [CH19]. One analogy that is

T^4	$T^3 \times \mathbb{R}$	$T^2 \times \mathbb{R}^2$	$S^1 \times \mathbb{R}^3$	$\mathbb{R}^4$
T^3	$T^2 \times \mathbb{R}$	$S^1 \times \mathbb{R}^2$	$\mathbb{R}^3$	
T^2	$S^1 \times \mathbb{R}$	$\mathbb{R}^2$		
S^1	$\mathbb{R}$			
$\{p\}$				

Table 1: The possible correspondences (up to diffeomorphism) between anti-self-dual connections via a Nahm transform on flat space are given by transposing this table along the forwards diagonal.

often given is that the Nahm transform is a generalised Fourier or Laplace transform, whereby it reverses the roles of differential and algebraic operators. For example, on $\mathbb{R}^4$, the instanton connection D^A consists of four differential operators $D^A_j = \partial_j + A_j$ for each of the four coordinates of $\mathbb{R}^4$, and its transformed data consists of four algebraic operators Φ_j, $j = 1, 2, 3, 4$, satisfying an associated 'self-duality' equation.

Alongside this neat correspondence between spaces and dual spaces, the process of the Nahm transform also follows a relatively common recipe. Although this needs to be taken with a pinch of salt – the process outlined below is not precise, and only a flavour of what occurs in general.

- The starting point is an anti-self-dual connection on $M = \mathbb{R}^4/_{\Gamma}$ (that is, on $\mathbb{R}^4$ with either some periodic dimensions, or completely invariant dimensions), on a hermitian bundle E, i.e. with fibers $\mathbb{C}^N$ and structure $U(N)$. We assume suitable additional boundary conditions are given.

- We construct what is known as the *Dirac operators* for that connection. To understand what this is explicitly relies on understanding spin structures and spin bundles, which we won't go into, but a good reference on spinors is [Har90]. The basic idea in 4-dimensions is that you have a bundle S equipped with an action of the quarternions. For $\mathbb{R}^4$, S decomposes as $S = S^+ \oplus S^-$,[13] whose fibers can each be

[12] The notion of a monopole will be introduced later in this lecture.

[13] This is due to the fact that $\mathrm{Spin}(4) = SU(2) \times SU(2)$.

viewed as just $\mathbb{C}^2$ with the obvious action, and the Dirac operators for a connection A are operators which act on sections of the tensor products $E \otimes S^{\pm}$. A section s here can simply be viewed as an element of $C^\infty(M, \mathbb{C}^N \otimes \mathbb{C}^2)$. Explicitly the Dirac operators are

$$\not{D}_A^+ s = \left(\sum_{a=0}^{3} D_a^A \otimes \tau^a \right) s, \quad \not{D}_A^- s = -\left(\sum_{a=0}^{3} D_a^A \otimes (\tau^a)^\dagger \right), \tag{57}$$

where $\tau^a : \mathbb{C}^2 \to \mathbb{C}^2$ satisfy $(\tau^a)^\dagger \tau^b + (\tau^b)^\dagger \tau^a = 2\delta^{ab}$. These we shall think of as $\tau^0 = \mathbb{1}_2$, and τ^1, τ^2, τ^3 as the Pauli matrices (26). Sometimes we shall also interchange the index 0 and 4. These operators may then be 'dimensionally reduced' to the connection on $\mathbb{R}^4/\Gamma$.

- For a point x in the dual space M^*, one then forms the related "Nahm operators"

  $$\Delta_x = \not{D}_A^+ - \imath x^a \mathbb{1}_N \otimes \tau^a, \quad \Delta_x^\dagger = \not{D}_A^- + \imath x^a \mathbb{1}_N \otimes (\tau_a)^\dagger. \tag{58}$$

  If any coordinate is completely invariant, we set it to 0.

- There is a natural hermitian inner-product on the space of sections of $E \otimes S^{\pm}$, given by the L^2-inner-product:

  $$\langle s_1, s_2 \rangle_{L^2} = \int_M s_1^\dagger s_2 \, \mathrm{Vol}_M, \tag{59}$$

  and one can show that the Nahm operators are adjoints of each other, i.e. $\langle \Delta_x^\dagger s_1, s_2 \rangle = \langle s_1, \Delta_x s_2 \rangle$, hence the notation.

- The condition of anti-self-duality of A, and the boundary conditions are chosen in such a way to ensure $\Delta_x^\dagger \Delta_x$ is invertible, and in particular $\ker \Delta_x = 0$, and $\ker \Delta_x^\dagger$ is finite-dimensional, say k. This space $\ker \Delta_x^\dagger$ is identified as the fibre $\widetilde{E}_x$ of a rank k vector bundle over M^* with structure $U(k)$. This bundle we denote as $\widetilde{E}$

- One can form an orthonormal basis $\{e^i\}_{i=1}^k$ for $\widetilde{E}_x$. A connection $\widetilde{A}$ on $\widetilde{E}$ is then formed from the induced connection on $\widetilde{E}$ from the trivial bundle[14] $\mathbb{C}^\infty \times M^*$. By this we mean for sections $\psi : M^* \to E$, the connection is given by the orthognal projection

  $$D^{\widetilde{A}} \psi(x) = \pi_{\widetilde{E}_x} \mathrm{d}\psi(x). \tag{60}$$

  The gauge field $\widetilde{A}$ is a $k \times k$ matrix of 1-forms with matrix components

  $$\widetilde{A}^{ij} = \langle e^i, \mathrm{d}e^j \rangle. \tag{61}$$

  See exercise 23 to explore this more.

[14] The space of all sections is infinite-dimensional.

- This induced connection provides and anti-self-dual connection on $\widetilde{E} \longrightarrow M^*$.

The important point to make here is that this is purely a heuristic, and in general additional structures and conditions need to be imposed – this idea is somehow covered by the statement that the connection has "suitable boundary conditions". Often this entails in the Nahm operator taking a slightly more general form than above. Furthermore, sometimes the transformed bundle is really some composition of bundles with different ranks, and the connection is also allowed to have singularities. The situation can become extremely complicated.

Nevertheless, the powerful feature of these transforms is that they are reversible (i.e. there is a transform both ways), and in most cases the reverse transform has been proven to be the true inverse, that is, if you apply the transform twice, you obtain the same connection (up to gauge equivalence).

One of the major features of the Nahm transform that we haven't stressed yet is that it typically interchanges the role of vector bundle rank with topological data, and vice versa. This is seen explicitly in the case of the Nahm transform on T^4 (which is a correspondence between ASD connections on T^4 and $(T^4)^* \cong T^4$) where one finds

$$\mathrm{rank}(E) = c_2(\widetilde{E}), \quad \mathrm{rank}(\widetilde{E}) = c_2(E). \tag{62}$$

This relationship is again by no means a recipe,[15] and we shall see that it manifests itself in various different ways when we look at different examples.

Exercise 23. *Consider the rank p trivial bundle $C^p \times \mathbb{R}^m \to \mathbb{R}^m$ equipped with the trivial (flat) connection d, and an hermitian inner product $\langle \cdot, \cdot \rangle$. Let E be a rank N subbundle. The* ***induced connection*** *∇ is the covariant derivative on E defined by its action on sections $s : \mathbb{R}^m \to E$ via*

$$\nabla s(x) = \pi_{E_x} \mathrm{d}s(x), \tag{63}$$

where $\pi_{E_x} : \mathbb{C}^m \to E_x$ is the orthogonal projection. With respect to an orthonormal basis $\{e^i\}_{i=1}^N$ of E_x, this is simply $\pi_{E_x}(v) = \sum_i \langle e^i, v \rangle e^i$.

1. *Show that the gauge field A of ∇ has matrix components (61).*
2. *Show that the gauge field is made up of anti-hermitian matrices, i.e. it is a $U(N)$ gauge field.*
3. *Show that a change of basis $\{e^i\} \mapsto \{\widetilde{e}^i\}$ is equivalent to a $U(N)$ gauge transformation of ∇.*

[15] Indeed, in many cases it doesn't make sense to talk about second Chern numbers.

2.2 The ADHM construction of instantons on $\mathbb{R}^4$

The first example of a "Nahm transform" was the Atiyah-Drinfeld-Manin-Hitchin construction of instantons on $\mathbb{R}^4$, known in short as the ADHM construction. It appeared long before the Nahm transform heuristic was ever thought of! The remarkable fact about it is that every $SU(N)$ instanton with instanton number k can be obtained by solving a system of purely algebraic equations, corresponding the the completely dimensionally reduced self-duality equations.

Exercise 24. *By writing a connection 1-form on $\mathbb{R}^4$ as $A = \sum_\mu A_\mu \, \mathrm{d}x^\mu$, show that the anti-self-duality equations reduce, in the case that each $A_\mu \equiv \Phi_\mu$ is constant, to the equations*

$$[\Phi_1, \Phi_i] + [\Phi_j, \Phi_k] = 0, \tag{64}$$

for cyclic permutations of the indices $(i, j, k) = (2, 3, 4)$.

As it happens, the real picture is slightly more complicated, and the true Nahm transform of an instanton on $\mathbb{R}^4$ is something for which the right-hand-side of these equations is non-zero. This seemingly surprising fact arises due to the non-compactness of $\mathbb{R}^4$.

The correct picture is as follows. The **ADHM data** for an $SU(N)$ instanton of instanton number k on $\mathbb{R}^4$ is determined by a tuple (V, W, M_1, M_2, I, J) where

- $V \cong \mathbb{C}^k$ and $W \cong \mathbb{C}^N$;
- $M_1, M_2 : V \to V$ are linear, i.e. are represented by complex $k \times k$ matrices;
- $I : W \to V$, $J : V \to W$ are linear, i.e. are represented by complex $k \times N$ and $N \times k$ matrices respectively.

These are to satisfy:

1. The **ADHM equations**

$$[M_1, M_2] + IJ = 0, \tag{65}$$

$$[M_1, M_1^\dagger] + [M_2, M_2^\dagger] + II^\dagger - J^\dagger J = 0. \tag{66}$$

2. The following "**full-rank conditions**", namely, for all $x \in \mathbb{R}^4$, here viewed as an element $x = (z_1, z_2) \in \mathbb{C}^2$, the maps

$$\alpha_x : V \longrightarrow W \oplus V \otimes \mathbb{C}^2 \tag{67}$$

$$\beta_x : W \oplus V \otimes \mathbb{C}^2 \longrightarrow V \tag{68}$$

given by

$$\alpha_x = \begin{pmatrix} J \\ -M_1 - z_1 \mathbb{1} \\ M_2 + z_2 \mathbb{1} \end{pmatrix}, \quad \beta_x = \begin{pmatrix} I & M_2 + z_2 \mathbb{1} & M_1 + z_1 \mathbb{1} \end{pmatrix}, \tag{69}$$

are full-rank, i.e. injective and surjective respectively. Here $\mathbb{1}$ is the $k \times k$ identity matrix.

Note that there is some ambiguity in the choice of complex structure which transitions $\mathbb{R}^4$ to $\mathbb{C}^2$ – in fact there is a whole family of isomorphisms one could choose, and this is quite important in an understanding of the bigger picture (see [Ati78]). For our purposes it doesn't matter, and so here we view $(z_1, z_2) \in \mathbb{C}^2$ via

$$z_1 = x_2 + \imath x_1, \quad z_2 = x_4 + \imath x_3. \tag{70}$$

Exercise 25. *Show that the ADHM equations (65)-(66) are equivalent to (64) in the case*[16] $M_1 = \Phi_1 + \imath\Phi_2$, $M_2 = \Phi_3 + \imath\Phi_4$, *and* $I = J = 0$.

So at first glance, these data look nothing like the condition of anti-self-duality, it's also not obvious how it fits into the Nahm transform heuristic given above. The details of where this comes from can be found in [Ati78], but the remainder of this section is dedicated to forming the "ADHM operator", which is the analogue of the Nahm operator, and then deriving an $SU(N)$ instanton of instanton number k.

The first step in understanding this process is as follows.

Exercise 26. *A sequence of vector spaces*

$$V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} V_3 \tag{71}$$

is called a **monad** *if it is a chain complex, i.e.* $f_2 \circ f_1 = 0$, *with* f_1 *injective, and* f_2 *surjective.*

1. *Let* f_1 *and* f_2 *define a monad. Calculate the dimension of the homology*

$$E = \ker(f_2) \big/ \mathrm{im}(f_1),$$

in terms of the dimensions of the V_i.

[16] Note there are various identifications that could be made to make this work, but it doesn't matter because this is simply an illustration.

2. *Show that sequence*

$$V \xrightarrow{\alpha_x} W \oplus V \otimes \mathbb{C}^2 \xrightarrow{\beta_x} V \tag{72}$$

where α_x, β_x are defined in (69) with the correct properties, is a monad if they are formed from ADHM data.

The consequence of this is that the ADHM data define a monad, whose homologies

$$E_x = \ker(\beta_x) \Big/ \mathrm{im}(\alpha_x), \tag{73}$$

which are vector spaces parameterised by $x \in \mathbb{C}^2$, form the fibres for a rank N vector bundle over $\mathbb{C}^2 \cong \mathbb{R}^4$. These vector spaces can be calculated explicitly at each $x \in \mathbb{R}^4$ by considering the kernel of the map $\Delta_x^\dagger : W \oplus V \otimes \mathbb{C}^2 \longrightarrow V \times V$, which acts on $\psi \in W \oplus V \otimes \mathbb{C}^2$ via

$$\Delta_x^\dagger \psi = (\beta_x \psi, \alpha_x^\dagger \psi). \tag{74}$$

This should be understood by thinking of E_x as the orthogonal complement of $\mathrm{im}(\alpha_x)$ inside $\ker(\beta_x)$ with respect to the natural hermitian inner product $\langle v, w \rangle = v^\dagger w$ on complex vector spaces.

Note that the map $\Delta_x^\dagger$ may be represented by the $(2k) \times (2k + N)$-matrix

$$\Delta_x^\dagger = \begin{pmatrix} I & M_2 & M_1 \\ J^\dagger & -M_1^\dagger & M_2^\dagger \end{pmatrix} + (\overline{x} \otimes \mathbb{1}_k) \begin{pmatrix} 0 & \mathbb{1}_k & 0 \\ 0 & 0 & \mathbb{1}_k \end{pmatrix}, \tag{75}$$

where $x = x^1\tau^1 + x^2\tau^2 + x^3\tau^3 + x^4\mathbb{1}_2$ is a quarternion represented as a 2×2 matrix via Pauli matrices (26), i.e.

$$x = \begin{pmatrix} \overline{z_2} & -z_1 \\ \overline{z_1} & z_2 \end{pmatrix}. \tag{76}$$

This map is precisely the analogue of the (adjoint) Nahm operator introduced before in the general heuristic recipe.

Constructing the $SU(N)$ gauge field. E_x is equipped with the natural hermitian inner product induced by the inner product on $W \oplus V \otimes \mathbb{C}^2$, and we can impose that the basis $\{e_i\}_{i=1}^n$ is orthonormal with respect to this inner product. Varying over $x \in \mathbb{R}^4$, this basis defines an orthonormal basis of smooth sections of the bundle E. From this, we define the $N \times N$ matrix $A = (A^{ij})$ as the induced connection (61).

From exercise 23, we know that A is a $U(N)$ gauge field. The important claim however is that this connection is an $SU(N)$ instanton with instanton number k. The subtlety is that the gauge field is in $\mathfrak{su}(N)$, which means it is traceless, and the key to this comes from anti-self-duality and that fact that there are no $U(1)$ connections with anti-self-dual curvature.

Note that at each point x we can form a $(2k+N)\times N$ matrix B_x whose columns form a basis for E_x. In this way the orthonormality condition is just $B_x^\dagger B_x = \mathbb{1}_N$, and each of these satisfy $\Delta_x^\dagger B_x = 0$. The connection is thus conceived as the gauge field with components

$$A_\mu(x) = B_x^\dagger \partial_\mu B_x. \tag{77}$$

Anti-self-duality. The process to show A has anti-self-dual curvature is essentially a long calculation, which is left as a guided exercise for the reader. Now, the matrix $\Delta_x^\dagger$ has an hermitian conjugate Δ_x, which must be full-rank by the full-rank conditions, and hence the $(2k)\times(2k)$ matrix $R(x) = \Delta_x^\dagger \Delta_x$ is invertible by surjectivity of $\Delta_x^\dagger$. This and the ADHM equations then tell us it takes the form

$$R(x) = \mathbb{1}_2 \otimes f(x), \tag{78}$$

for some invertible $k\times k$ matrix $f(x) = \alpha_x^\dagger \alpha_x = \beta_x \beta_x^\dagger$. The result of anti-self-duality hence follows by calculation.

Exercise 27. *1. Show that at any given point $x\in\mathbb{R}^4$, the curvature of (77) has components*

$$F_{\mu\nu} = B^\dagger \left(\partial_{[\mu}\Delta\right) R^{-1} \left(\partial_{\nu]}\Delta^\dagger\right) B, \tag{79}$$

where the explicit dependency on x is dropped for brevity, and we have introduced the notation $X_{[\mu}Y_{\nu]} = X_\mu Y_\nu - X_\nu Y_\mu$.

2. Show that (79) reduces to

$$F_{\mu\nu} = B^\dagger \tau^{[\mu}(\tau^{\nu]})^\dagger \otimes f^{-1} B, \tag{80}$$

where $\tau^4 \equiv \mathbb{1}_2$, and additional tensor products with matrices only containing zeros or identity matrices are omitted for simplicity.

3. By showing that $\eta_{\mu\nu} = \tau^{[\mu}(\tau^{\nu]})^\dagger$ is ASD, which amounts to showing

$$\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\eta_{\rho\sigma} = -\eta_{\mu\nu}, \tag{81}$$

conclude that the induced connection on $E = \ker(\beta) \big/ \mathrm{im}(\alpha)$ for the ADHM complex is anti-self-dual.

Instanton number The fact that this construction yields an instanton with instanton number k is not so clear. In brief, the idea is that this "monad construction" is part of a more general object, whereby the maps α_x and β_x may be seen as a special case of more general maps defined for a general point $(x, y) \in \mathbb{C}^2 \times \mathbb{C}^2 \setminus \{(0, 0)\}$, where the extra element of $\mathbb{C}^2$ acts by multiplication on the first term of $\Delta^\dagger$ (see exercise 28). One can show that the fibres $E_{(x,y)}$ and $E_{(xq,yq)}$ are identical, for $q \in \mathbb{H}$ non-zero, thus this defines a bundle over the quarternionic projective line $\mathbb{HP}^1$, which is identified as the four-sphere S^4, and thus the instanton number is (minus) the second Chern number of this bundle. Now one still has an operator $\Delta^\dagger_{(x,y)}$ which acts on $\mathbb{C}^{2k+N}$ analogously, and E is the bundle defined by the kernel of this operator. The complement to E in the trivial bundle $\mathbb{HP}^1 \times \mathbb{C}^{2k+N}$ is therefore a (complex) rank $2k$ bundle, and one can show that this can be identified as the direct sum of k tautological quarternionic bundles Σ, namely the bundle whose fiber is just the quarternions $\mathbb{H}$. These have second Chern number 1, whereas c_2 of a trivial bundle is 0. Thus as $E \oplus (\Sigma^k) \cong \mathbb{HP}^1 \times \mathbb{C}^{2k+N}$, their Chern numbers coincide, and so since $c_2(E \oplus F) = c_2(E) + c_2(F)$, one gets

$$c_2(E) + k = 0. \tag{82}$$

Exercise 28. *Let $x, y \in \mathbb{H}$ be identified by pairs of complex numbers $(z_1, z_2), (w_1, w_2)$ as in (76). For $(x, y) \neq (0, 0)$, define the operator*

$$\Delta^\dagger_{x,y} = (\bar{y} \otimes \mathbb{1}_k) \begin{pmatrix} I & M_2 & M_1 \\ J^\dagger & -M_1^\dagger & M_2^\dagger \end{pmatrix} + (\bar{x} \otimes \mathbb{1}_k) \begin{pmatrix} 0 & \mathbb{1}_k & 0 \\ 0 & 0 & \mathbb{1}_k \end{pmatrix}. \tag{83}$$

Notice that (75) is the special case of this with $y = (0, 1)$.

1. *Derive the corresponding full-rank conditions analogous to (69) for $(z_1, z_2), (w_1, w_2)$ non zero that, when coupled with the ADHM equations, ensure that $\Delta^\dagger_{x,y}\Delta_{x,y}$ is invertible.*

2. *Consider the relation on $\mathbb{H}^2$ given by $(x, y) \sim (qx, qy)$ for some $q \in \mathbb{H} \setminus \{0\}$. Show that this is an equivalence relation, and that the quotient space $\mathbb{H}^2 /_\sim$ may be identified as the four sphere S^4.*

3. *Show that $\ker \Delta^\dagger_{x,y}$ is invariant on the equivalence classes, i.e. $\ker \Delta^\dagger_{qx,qy} = \ker \Delta^\dagger_{x,y}$.*

2.2.1 Moduli

The ADHM construction is a relatively abstract recipe for generating instantons on $\mathbb{R}^4$, but the magical fact is that one no longer needs to solve any partial differential equations.

The even more important point, which we won't prove, is that every instanton on $\mathbb{R}^4$ arises uniquely in this way, thus the problem has been reduced to a system of algebraic conditions. Now, of course, as was the case with the instanton, there are redundant degrees of freedom here too, namely gauge transformations. In this context, gauge transformations are simply isomorphisms of the vector spaces V and W which preserve the hermitian inner product, and act on the linear maps as changes of basis, namely for $g \in U(k)$ and $h \in SU(N)$, the ADHM data are invariant under

$$(M_1, M_2, I, J) \mapsto (gM_1g^{-1}, gM_2g^{-1}, gIh^{-1}, hJg^{-1}). \tag{84}$$

The moduli space of ADHM data is then the set of all (M_1, M_2, I, J) modulo this action. Again, there is a notion of framing here, which is identified by a fixed choice of basis for the vector space W, which then views the $SU(N)$ action as non-trivial, and is not quotiented out. It is now easy to check by a simple counting argument what the dimension is. The matrices contribute $2k^2 + 2Nk$ complex moduli. The ADHM equations are $3k^2$ real equations, and the action of $U(k)$ forms k^2 real equations, thus contributing a total reduction of $2k^2$ complex dimensions. The full-rank conditions are open conditions, thus do not change the dimension, so the full complex dimension of the moduli space is $2Nk$, which is exactly what is required for an instanton on $\mathbb{R}^4$.

Exercise 29. *Show that any two instanton gauge fields (77) derived from gauge-equivalent ADHM data are gauge equivalent.*

Exercise 30. *Show that all solutions to the ADHM equations modulo the $U(k)$ gauge transformations (84) may be realised uniquely by solutions to the complex equation (65) modulo $GL(k, \mathbb{C})$ gauge transformations, that is, via (84) where g is replaced by the more general $g \in GL(k, \mathbb{C})$ (and h is ignored).*

2.2.2 The BPST instanton via ADHM data

Recall in the previous lecture we derived the BPST instanton via a symmetric ansatz, and by exploiting isometries of $\mathbb{R}^4$ we arrived at a general instanton with instanton number 1. To illustrate how the ADHM construction works, we can calculate the BPST instanton from its ADHM data, which are found in the case $k = 1$ and $N = 2$ above. The important point here is that the process is significantly more straightforward, with no need to solve a differential equation.

In this case, M_1, M_2 are just complex numbers, which we could call $M_1 = a_2 + \imath a_1$, and $M_2 = a_4 + \imath a_3$, with $a \in \mathbb{R}^4$. In particular their commutators are zero, leaving the ADHM equations (65)-(66) to be simply

$$IJ = 0, \quad \text{and} \quad II^\dagger = J^\dagger J, \tag{85}$$

for I, J a 2-component complex column and row vector respectively. This is equivalent to choosing two elements of $\mathbb{C}^2$ which are orthogonal with respect to the real inner product on $\mathbb{R}^4$, and have the same length with respect to the standard hermitian inner product. The full-rank conditions insist that these are non-zero. Up to unframed unitary gauge transformations in $SU(2)$ we can choose w.l.o.g. $I = (\lambda, 0)$, $J = (0, \lambda)^T$, and we may fix $\lambda \geq 0$ via a $U(1)$ transformation. In fact, $\lambda \neq 0$ due to the full-rank conditions.These parameters a and λ are precisely the moduli parameters we met for the BPST instanton in the previous lecture. Three extra parameters can be introduced by varying the choice of $SU(2)$ orientation, i.e. by replacing I, J by Iq^{-1} and qJ for $q \in SU(2)$, giving a general element of the framed moduli space.

Exercise 31. *Show that the gauge field (77) derived from the unframed ADHM data is the BPST instanton*

$$A = \mathrm{Im}\left(\frac{x-a}{\lambda^2 + |x-a|^2}\,\mathrm{d}\overline{x}\right). \tag{86}$$

2.2.3 Higher charge examples

In principal, the ADHM construction allows for the construction of examples of instantons with arbitrary instanton number k, and indeed the full space of solutions has been understood in the cases $k = 1, 2$, and 3. Unfortunately, despite being more straightforward than solving the anti-self-duality equations, the ADHM equations become more difficult to pin down for $k > 3$, simply due to the number of degrees of freedom involved. Nevertheless, solutions are known for all k, for example with high degrees of symmetry about a fixed axis [FH90], or special cases with polytope symmetries [AS13, Sut04], and importantly, these solutions have not been realised via any other more direct ansatz for the gauge field itself.

2.3 Monopoles

The next example we shall consider is the original Nahm transform, namely the Nahm transform for *monopoles* [Nah83]. Monopoles are defined as finite-energy solutions to a set of PDEs on $\mathbb{R}^3$. These equations (and finite-energy condition) may be realised by dimensional reduction of the ASD equations on $\mathbb{R}^4$ by 1 dimension. To see what is meant by this, fix global coordinates $(x^1, \ldots, x^4)$ for $\mathbb{R}^4$, and assume that the gauge field A is constant in the x^4-direction. In this way we obtain a gauge field B on $\mathbb{R}^3$, defined by the components A_1, A_2, and A_3, and in the invariant dimension x^4 the component A_4 of the gauge field A is simply an endomorphism of a vector bundle over $\mathbb{R}^3$ – in the unitary case, this may be viewed simply as a map $A_4 : \mathbb{R}^3 \to \mathfrak{u}(n)$. This endomorphism, typically denoted by Φ, is called the *Higgs field*.

Exercise 32. *Let $A = B + \Phi \, \mathrm{d}x^4$, where B is a connection on $\mathbb{R}^3$, and $\Phi : \mathbb{R}^3 \to \mathfrak{u}(n)$. Show that (up to a choice of orientation on $\mathbb{R}^3$) the anti-self-duality of the connection A is equivalent to the Bogomolny equations*

$$\star D^B \Phi = -F^B, \tag{87}$$

where here $\star$ is the Hodge-star on $\mathbb{R}^3$.

Finite energy solutions to the Bogomolny equations are called **monopoles**. But what does "finite energy" mean? By introducing this dimensional reduction from $\mathbb{R}^4$, having an invariant dimension automatically disqualifies the connection from yielding finite Yang-Mills action, since the action will have an infinite contribution from one of the dimensions. However, this notion can be understood by compactifying the invariant dimension to a circle. In this case the Yang-Mills action takes the form

$$S_{YM} = \beta E[B, \Phi], \tag{88}$$

where β is the period of the circle (which can be anything you want), and E is the contribution of the remaining gauge field and Higgs field in the Yang-Mills action, and is an integral over $\mathbb{R}^3$.

Exercise 33. *Show that for the connection $A = B + \Phi \, \mathrm{d}t$, with t denoting a coordinate on the invariant circle dimension, one obtains the Yang-Mills-Higgs energy*

$$E = ||F^B||_{L^2} + ||D^B \Phi||_{L^2}. \tag{89}$$

Exercise 34. *Let (B, Φ) be an $SU(N)$ gauge field B and $\mathfrak{su}(n)$ Higgs field satisfying the Bogomolny equations (87) on $\mathbb{R}^3$.*

1. *Show that the Yang-Mills-Higgs energy (89) takes the form*

$$E = -\int_{\mathbb{R}^3} \mathrm{dtr}\,(\Phi \star D^B \Phi). \tag{90}$$

2. *Show that this reduces to*

$$E = 2 \int_{\mathbb{R}^3} \nabla^2 ||\Phi||^2, \tag{91}$$

where $||\Phi||^2 = -\frac{1}{2}\mathrm{tr}\,(\Phi^2)$is the norm on $\mathfrak{su}(N)$, and ∇^2 is the laplacian on $\mathbb{R}^3$.

In order to obtain finite energy, as we know from the case of instantons on $\mathbb{R}^4$, boundary conditions must be imposed. However, these turn out to be more complicated! Naïvely we may want to draw inspiration from instantons on $\mathbb{R}^4$ and ask for the monopole (viewed as something defined on $S^1 \times \mathbb{R}^3$) to extend to the one-point-compactification $S^1 \times \mathbb{R}^3 \cup \{\infty\} \cong S^1 \times S^3$, and then the monopole would be classified by a second Chern number like with instantons. However, since we ask for S^1-invariance, this is topologically uninteresting as this means all transition functions can be viewed as S^1-invariant too, i.e. a transition function for a bundle over S^3, which is a map $g_{NS} : S^2 \to SU(N)$. But $\pi_2(SU(N)) = 0$, so the Chern number would be zero. Luckily, finite action does not require such a condition here, and there are less strict conditions one can impose which also guarantee finite-action, and these in turn define well-defined topological charges. The key is to think of $\mathbb{R}^3$ as being the interior of a 3-ball B^3, and a suitable extension to the boundary S^2_∞ is given.

These conditions require going into a lot of divergent details in general, so for simplicity we shall consider the case $SU(2)$. In this case, the boundary conditions may be summarised by the imposition that

$$||\Phi||_{S^2_\infty} = \nu > 0. \tag{92}$$

Note that the level sets of the norm in $\mathfrak{su}(2)$ are 2-spheres, hence this condition implies Φ extends to a map $\Phi_\infty : S^2_\infty \to S^2_\nu$. These are classified by their representative in the second homotopy group $\pi_2(S^2) = \mathbb{Z}$. Thus, monopoles are classified, like instantons, by an integer. But this time the charge is something rather different which depends on the topology of the boundary rather than the global space. The integer charge is called in this context the **monopole number** or *magnetic charge* m of the monopole. It's worth remarking that this is actually a Chern number – the boundary condition says Φ_∞ has eigenvalues $\pm\imath\nu$, and its eigenvectors span two rank 1 sub-bundles whose *first Chern numbers* are $\pm m$ respectively. This also helps to understand why they're called *magnetic charges* – the first Chern number of a unitary connection on a rank 1 bundle over S^2 is calculated via the formula

$$c_1 = \frac{\imath}{2\pi} \int_{S^2} \operatorname{tr}(F^A), \tag{93}$$

which may be interpreted as something proportional to the total magnetic flux induced by the connection through S^2. The case of $SU(N)$ is similar, but more complicated, with N eigenvalues, and N eigenbundles contributing $N-1$ integer magnetic charges.

Again, like with instantons, the space of all monopoles is unnecessarily large. We instead consider the **moduli spaces** of monopoles

$$\mathcal{M}_{m,\nu} = \{\text{monopoles of charge } m \text{ and size } \nu\} /_{\mathcal{G}}, \tag{94}$$

where $\mathcal{G}$ is the gauge group, that is all smooth $g : \mathbb{R}^3 \to SU(2)$ acting on (B, Φ) via

$$B \mapsto gBg^{-1} - \mathrm{d}gg^{-1}, \tag{95}$$

$$\Phi \mapsto g\Phi g^{-1}. \tag{96}$$

There is also a notion of framing here, where now one fixes a choice of trivial bundle over the whole 2-sphere S^2_∞, specifically the one which is formed from the eigenbundles $\mathcal{O}(\pm m)$, and asks for gauge transformations to be identity there. Note that this is important since a general action of $\mathcal{G}$ can alter the eigenspaces of Φ_∞.

2.3.1 The Nahm transform for monopoles

So $SU(2)$ monopoles are solutions to the self-duality equations invariant in one of the dimensions, which are classified by an integer m, and a size ν corresponding to the asymptotic length of the Higgs field. According to the Nahm transform heuristic, these should be in correspondence to rank m solutions of the anti-self-duality equations in which three of the dimensions are invariant, plus some boundary conditions.

Exercise 35. *Let $A = T^0\,\mathrm{d}s + \sum_{j=1}^{3} T^j\,\mathrm{d}x^j$ be an $\mathbb{R}^3$-invariant connection on $\mathbb{R}^4$, that is where $T^\mu : \mathbb{R} \to \mathfrak{u}(m)$ are just functions of $s \in \mathbb{R}$. Show that the anti-self-duality of A reduces to* **Nahm's equations***, that is the equations*

$$\frac{\mathrm{d}T^j}{\mathrm{d}s} + [T^0, T^j] + \frac{1}{2}\varepsilon_{jkl}[T^k, T^l] = 0, \tag{97}$$

for $j = 1, 2, 3$

As was the case for instantons on $\mathbb{R}^4$ the situation is not simply that all solutions to Nahm's equations give monopoles, and other conditions need to be imposed. In general, the **Nahm data** for a monopole of charge m and size ν corresponds to a tuple (K, T^0, T^1, T^2, T^3) such that

- $K \longrightarrow [-\nu, \nu]$ is a rank k hermitian vector bundle, equipped with a connection $\nabla = \mathrm{d} + T^0$ and three endomorphisms T^j;
- $T^\mu : [-\nu, \nu] \longrightarrow \mathfrak{u}(m)$ satisfy Nahm's equations (97);
- T^0 is analytic everywhere, and T^j are analytic on the interior $(-\nu, \nu)$, and have simple poles at $\pm\nu$;
- The residues R^j of the poles form an irreducible m-dimensional representation of $\mathfrak{su}(2)$.

Remark 36. *In many cases one also imposes the reality condition, which is that $T^\mu(-s) = T^\mu(s)^t$, where ^{t} denotes matrix transposition. This is not necessary, but merely a gauge choice.*

There is a group of gauge transformations acting on Nahm data, namely smooth functions $g : [-\mu, \mu] \to U(m)$ which act via

$$T^0 \mapsto gT^0g^{-1} - \mathrm{d}gg^{-1}, \tag{98}$$

$$T^j \mapsto gT^jg^{-1}. \tag{99}$$

This is exactly what one should expect if we think of the gauge group as being the restriction of the group of gauge transformations on $\mathbb{R}^4$ which are invariant on $\mathbb{R}^3$, and this is the action on the invariant connection defined by the T^μ. The remarkable fact is that there is a one-to-one correspondence between moduli of Nahm data and moduli of monopoles.

The Nahm operator and transform. The process to obtain a monopole from Nahm data is similar to the ADHM construction. The process again requires, for each $x \in \mathbb{R}^3$ forming an adjoint Nahm operator $\Delta_x^\dagger$ from the Nahm data, whose kernel provides the fibres for a rank 2 vector bundle over $\mathbb{R}^3$, and then the connection and Higgs field can be formed from the induced connection on this bundle. In this case though, the analysis is less straightforward than the case of the ADHM construction, since now, rather than just being a linear operator on a finite-dimensional vector space, the Nahm operator is defined on an infinite-dimensional function space. In this sense, this example is closer to the heuristic sketched earlier.

Explicitly, consider the space $V = C^\infty([-\mu, \mu], \mathbb{C}^m \otimes \mathbb{C}^2)$. V has a natural inner-product on it, namely

$$\langle \psi_1, \psi_2 \rangle_V = \int_{-\mu}^{\mu} \psi_1^\dagger(s)\psi_2(s)\ \mathrm{d}s. \tag{100}$$

We then construct the Nahm operator $\Delta_x : V \to V$ for $x = (x^1, x^2, x^3) \in \mathbb{R}^3$, whose adjoint is given via

$$\Delta_x^\dagger = -\left(\frac{\mathrm{d}}{\mathrm{d}s} + T^0\right) \otimes \mathbb{1}_2 - \sum_{j=1}^{3} (T^j - \imath x^j) \otimes \tau^j, \tag{101}$$

Unlike the case of the ADHM operator, it is not clear that the kernel of this operator is 2-dimensional. However, given this is true, again we can find an orthonormal basis $\{\psi^1, \psi^2\}$ of $\ker \Delta_x^\dagger$, from which the connection B and Higgs field Φ arise as the induced connection. This is understood by considering them as before as a connection $A = B + \Phi\ \mathrm{d}t$, for some invariant coordinate $t \in \mathbb{R}$, and noting that $\ker(\Delta_x^\dagger + \imath t \otimes \mathbb{1}) = e^{\imath ts} \ker \Delta_x^\dagger$. Then one may form similarly an orthonormal basis $\{e^1, e^2\}$ of $\ker(\Delta_x^\dagger + \imath t \otimes \mathbb{1})$, from which the induced connection with matrix components

$$A^{ij} = \langle e^i, \mathrm{d}e^j \rangle \tag{102}$$

gives the connection and Higgs field via

$$B_a^{ij} = \langle e^i, \partial_a e^j \rangle = \langle \psi^i, \partial_a \psi^j \rangle, \quad a = 1, 2, 3, \tag{103}$$

$$\Phi^{ij} = \langle e^i, \partial_t e^j \rangle = \langle \psi^i, \imath s \psi^j \rangle. \tag{104}$$

These turn out to define a monopole with monopole number m and size ν, and every monopole arises in this way. The details of the proof of this are relatively involved, but the main points are as follows:

- The Nahm equations are equivalent to the condition that $\Delta_x^\dagger \Delta_x$ is invertible (See exercise 37). The invertibility of $\Delta_x^\dagger \Delta_x$ is key to the proof that B and Φ satisfy the Bogomolny equation, just as was the case of the ADHM construction. This is a long calculation involving constructing formulae for the Green's function of $\Delta_x^\dagger \Delta_x$.

- One also has by invertibility of $\Delta_x^\dagger \Delta_x$ that $\ker \Delta_x = 0$. The boundary conditions on the Nahm data imply that the operator Δ_x has index -2, so that by the index formula

$$\operatorname{index}(\Delta_x) = \dim(\ker \Delta_x) - \dim(\ker \Delta_x^\dagger), \tag{105}$$

  the dimension result follows. See [Hit83] for details.

- The result about the size is relatively obvious: rescaling the interval $[-\nu, \nu]$ to $[-1, 1]$ rescales the Higgs field via $\Phi \mapsto \nu\Phi$. The result about the monopole number requires analysis of the Nahm operator for large x, and the Chern numbers of the associated eigenbundles which arise.

Exercise 37. *Consider the operator*

$$\Delta_x = \left(\frac{\mathrm{d}}{\mathrm{d}s} + T^0\right) \otimes \mathbb{1} - \sum_{j=1}^{3} (T^j - \imath x^j) \otimes \tau^j. \tag{106}$$

1. *Show that $\Delta_x^\dagger$ given by (101) is the adjoint of Δ_x with respect to the inner product (100).*

2. *Show that the condition $\Delta_x^\dagger \Delta_x$ is a real operator, i.e. has no component tensored with τ^j, is equivalent to Nahm's equations.*

3. *Use this to argue that $\ker \Delta_x^\dagger \Delta_x = 0$ and hence $\ker \Delta_x = 0$.*

2.3.2 Examples

Firstly, it's a relatively simple exercise in differential equations to show that there is always a gauge transformation which sends T^0 to 0 – it is the so-called 'path-ordered exponential' of $-T^0$, or inverse parallel transport operator [GS89]. Thus, Nahm's equations really only amount to finding the three functions T^j with the proper behaviour at $\pm\nu$.

The case $m = 1$. The simplest case $m = 1$ shows the power of the Nahm transform very clearly. In this case, the Nahm data are functions $T^j : [-\mu, \mu] \to \imath\mathbb{R}$, so Nahm's equations are simply

$$\frac{\mathrm{d}T^j}{\mathrm{d}s} = 0. \tag{107}$$

i.e. they're constants. These can be interpreted as having poles with residue 0, and clearly the representation $\tau^a \mapsto 0$ is the only 1-dimensional irreducible representation of $\mathfrak{su}(2)$. For simplicity, we can set all $T^j = 0$, and one can see that this simply corresponds to a translation $\vec{x} \mapsto \vec{x} + \vec{T}$ in $\mathbb{R}^3$. To obtain the full moduli space, we can translate back later. The fact we can get something non-trivial from functions which are identically 0 is one of the most curious things about the Nahm transform. It is left as an exercise for the reader to show that the resulting monopole, up to gauge equivalence, is given by

$$\Phi = \sum_{i=1}^{3} \frac{1}{r}\left(\frac{1}{2r} - \nu\coth(2\nu r)\right) x^i \tau^i, \tag{108}$$

$$B = \sum_{i,j,k=1}^{3} \frac{1}{2r^2}\left(1 - \frac{2\nu r}{\sinh(2\nu r)}\right) \varepsilon_{ijk} x^i \tau^j \, \mathrm{d}x^k, \tag{109}$$

with $r = |\vec{x}|$, and τ^j are the Pauli matrices (26).

Exercise 38. *Consider the rank 1 Nahm data $T^\mu = 0$ for all $\mu = 0, 1, 2, 3$.*

1. *Construct the adjoint Nahm operator $\Delta_x^\dagger$ for T^μ, and show that a general element of $\ker \Delta_x^\dagger$ is given by*

$$\psi_x(s) = \sqrt{\frac{r}{\sinh(\nu r)}} \exp\left(\imath \sum_j x^j \tau^j s\right) V, \tag{110}$$

   where $r = |\vec{x}|$, and V is a constant vector in $\mathbb{C}^2$.

2. *Choose an orthonormal basis for the Kernel so that the resulting connection 1-form and Higgs field are (108)-(109).*

3. *Verify that these satisfy the Bogomolny equations* (87).

As it happens, like with the BPST instanton, this monopole was discovered without the sophistication of the Nahm transform by imposing a spherically-symmetric ansatz [PS75]. And again we can construct the full moduli space by acting with isometries. However, as is clear from this construction, ultimately all of the data for this monopole is encoded in the simple Nahm data $T^\mu = 0$, which is significantly easier to understand than the above solutions to a 3-variable PDE.

The case $m > 1$. Unlike instantons on $\mathbb{R}^4$ where one has the Corrigan-Fairlie-'t Hooft ansatz, there have been no examples of explicit solutions for higher charge constructed simply at the level of Higgs field and connection – the only success in constructing higher charge examples is via the mechanism of the Nahm transform, and even in these cases only recently were the connection and Higgs field constructed explicitly for $m = 2$ [BE18]. The solutions for charge 2 and higher all involve special functions, for example, the general solution for charge 2 requires the Jacobi elliptic functions, and other higher charge solutions make use of the Weierstrass elliptic function $\wp$. Furthermore, in the higher charge cases, the only examples found have a high degree of symmetry imposed, and all require numerical methods in order to reconstruct the gauge and Higgs field. For a good review of these, see [MS04].

2.4 Calorons

The final example of a Nahm transform we will look at is that of **calorons**, which are instantons on $S^1 \times \mathbb{R}^3$. This is also the final case from the table 1 where the Nahm transform reduces a PDE into an ODE (or better, algebraic equations). In this case, there is no dimensional reduction needed in the anti-self-duality equations. The only adjustment is that one of the coordinate directions is periodic. Calorons may be seen as a generalisation of monopoles on $\mathbb{R}^3$ – indeed, we saw that monopoles are realised as instantons on $S^1 \times \mathbb{R}^3$ which are totally independent of the periodic direction. For this reason, the topology of calorons is similar to monopoles, but more general. Again, for simplicity we shall stick to the case of $SU(2)$, but this can be understood in general for structure $SU(N)$.

Recall that in the case of monopoles, we discarded viewing the instanton as something which extends to $S^1 \times S^3$, as for monopoles this was topologically trivial. Well, in this case, if we set all magnetic charges to be 0 instead, then we can define such an extension, and there is a well-defined Chern number, which in this context can be called the **instanton number** of the caloron. So there are two clear distinct cases of charges: calorons with instanton number k, and no magnetic charge, and calorons with magnetic charge m, and no instanton number, i.e. monopoles. In fact, there is a way to introduce the instanton number in more generality which allows non-zero magnetic charges, but this is beyond the scope of this lecture. See [Nye01] for more details. So really, the cases are not just distinct, but one can have a mixture of instanton and monopole charges, where calorons are classified topologically by two integer charges k and m, an instanton number, and a monopole number.

In addition to these charges, there is a parameter μ, which acts similarly to the "size" ν of the monopole from before, corresponding to an eigenvalue, but this time, μ is found in the eigenvalues $\exp(\pm\imath\mu\beta)$ of the asymptotic holonomy of the caloron in the S^1 direction, where here β is the period. This parameter μ is imposed to be constant. When $k = 0$, this is precisely (up to a choice of sign) the size ν. For uniqueness, the parameter μ is restricted to $0 \leq \mu < \pi/\beta$. It turns out that this *holonomy parameter* is extremely important for distinguishing caloron types. Using similar methods to instantons on $\mathbb{R}^4$, one can easily construct examples of instanton number k calorons [HS78]. However, all of these examples have holonomy parameter $\mu = 0$, and are thus classed as having *trivial holonomy*. To obtain calorons with non-trivial holonomy ($\mu \neq 0$), even for $k = 1$, the only known method requires the Nahm transform. In fact calorons were the first example where the construction of an unknown charge 1 example required the sophistication of the Nahm transform.

2.4.1 Nahm's equations on a circle

Inspecting the table 1, we see that calorons should be in correspondence with solutions to the "self-duality" equations on a circle. This is a 1-dimensional space, and we know already that that means the equations are Nahm's equations (97).

As was the case for instantons on $\mathbb{R}^4$, and for monopoles on $\mathbb{R}^3$, there are more technicalities that need to be addressed to paint the full picture. A brief description of the Nahm data for a (k, m) caloron is as follows, but for a stricter treatment see [CH07, Nye01]. One has:

- Two intervals $I_1 = [-\mu, \mu]$ and $I_2 = [\mu, -\mu + \mu_0]$, where $\mu_0 = 2\pi/\beta$. These are seen to cover the circle $\mathbb{R}/_{\mu_0\mathbb{Z}}$ by imposing periodicity in μ_0.
- Bundles $K_p \to I_p$ of rank $k+m$ and k respectively, each equipped with a connection $\nabla_p = \mathrm{d} + T_p^0$ and endomorphisms T_p^j, analytic on the interior, satisfying Nahm's equations (97) on the interior of each interval.
- Projections and inclusions from the higher rank to the lower rank and vice versa respectively on the fibers above the boundaries of the intervals (again identifying $-\mu$ with $-\mu + \mu_0$).

The Nahm matrices T_p^μ are then subject to boundary conditions at the end-points of each interval, which differs depending on rank.

- If $m > 0$, then the boundary conditions are essentially the same as the monopole conditions. The projection of T_1^μ should be identified with T_2^μ for all μ. In the rank m complement, T_1^j have simple poles at $\pm\mu$ with residues forming an irreducible representation of $\mathfrak{su}(2)$. A similar story holds if $m < 0$, but just the other way around.

- If $m = 0$, the fibres at each boundary are identified by bundle isomorphisms between K_1 and K_2, and at these points the matrices are to solve the **matching conditions**

$$\begin{aligned} \mathcal{A}_1(\zeta,\mu) - \mathcal{A}_2(\zeta,\mu) &= (u_+ - w_+\zeta)\left(w_+^\dagger + u_+^\dagger\zeta\right), \\ \mathcal{A}_2(\zeta,\mu_0-\mu) - \mathcal{A}_1(\zeta,-\mu) &= (u_- - w_-\zeta)\left(w_-^\dagger + u_-^\dagger\zeta\right), \end{aligned} \tag{111}$$

  where $(u_\pm, w_\pm) \in \mathbb{C}^k \setminus \{(0,0)\}$, and

$$\mathcal{A}_p(s,\zeta) = T_p^2(s) + \imath T_p^3(s) + 2\imath T_p^1(s)\zeta + (T_p^2(s) - \imath T_p^3(s))\zeta^2, \quad s \in I_p,\ \zeta \in \mathbb{C}. \tag{112}$$

Note if $k = 0$ then the data on I_2 is rank 0, and one is left with monopole Nahm data on $I_1 = [-\mu,\mu]$, satisfying the same boundary conditions as we met previously.[17]

- Finally, the data are imposed to be **irreducible**, that is, there are no sections of $K = K_1 \cup K_2$, parallel with respect to ∇_p, which are invariant under the T_p^j.[18]

Again, there is a notion of gauge transformation, but the situation is slightly more complicated. Assuming $m \geq 0$, these are smooth maps $g_p : I_p \to U(m_p)$ (where $m_1 = k+m$ and $m_2 = k$) such that

$$g_2(-\mu+\mu_0) = \mathrm{diag}\{g_1(-\mu), \mathbb{1}_m\}, \quad g_2(\mu) = \mathrm{diag}\{g_1(\mu), \mathbb{1}_m\}, \tag{113}$$

which act on the Nahm data in the standard way on each interval, and when $m = 0$, act on the vectors $(u_\pm, w_\pm)$ via the standard action on vectors in the fibres.

Exercise 39. *Show that one can always choose a gauge such that T_p^0 is constant on each interval.*

The importance of the above exercise is that the "gauging the connection to 0" idea only works on an interval, i.e. the imposition of periodicity for the gauge transformations makes it no longer possible to do this in general.

[17] Technically, this still needs to be adjusted since the gauge group acting on this data requires a periodicity condition which does not exist in the monopole case.

[18] A section ϕ is parallel to a connection if $\nabla\phi = 0$, and invariant under endomorphisms means simply $T^j\phi = \phi$.

Exercise 40. *Solve Nahm's equations and the matching conditions in the case $(k, m) = (1, 0)$, for arbitrary μ and μ_0. Show that the moduli space is parameterised by 8 real parameters.*

The Nahm transform has been performed analytically for the Nahm data in the case of the previous exercise, and the associated caloron/s are called KvBLL calorons [KvB98, LL98]. This case, although somehow similar to the monopole case, proves to be a far more complicated calculation.

2.5 Concluding comments

This lecture has only shown the tip of the iceberg when it comes to Nahm transforms. In flat space, there are cases for all of the correspondences in table 1, but many of them are only recently developed, and can be far more involved than those we have discussed here.

2.5.1 Completing the list of reductions of $\mathbb{R}^4$

There's one case of dimensional reduction of the anti-self-duality equations on $\mathbb{R}^4$ that we have not seen yet. This applies in the case where there are 2 invariant dimensions, which occurs in three of the examples in table 1. Anti-self-duality of a connection which is invariant in 2 dimensions implies *Hitchin's equations*. Here the equations are more easily understood by identifying a complex structure on $\mathbb{R}^2$. Then the covariant derivative D^B forms two operators ∂^B and $\overline{\partial}^B$ given by

$$\partial^B = \frac{1}{2}\left(D_1^B - \imath D_2^B\right), \quad \overline{\partial}^B = \frac{1}{2}\left(D_1^B + \imath D_2^B\right). \tag{114}$$

One also then thinks of Φ_1 and Φ_2 as one complex 1-form $\Phi = (\Phi_1 + \imath\Phi_2)\mathrm{d}z$, with $z = x^3 + \imath x^4$. Hitchin's equations are then given by

$$\begin{aligned} F^B + [\Phi, \Phi^*] &= 0, \\ \overline{\partial}^B \Phi &= 0. \end{aligned} \tag{115}$$

Solutions (B, Φ) to these equations are called **Higgs pairs**, and can be understood on any Riemann surface, and bundles with Higgs pairs are called **Higgs bundles**.

From the Nahm transform heuristic, one finds that solutions to Hitchin's equations on the torus T^2 and on the cyclinder $S^1 \times \mathbb{R}$ correspond to periodic monopoles (i.e. monopoles on $S^1 \times \mathbb{R}^2$), and doubly periodic instantons (i.e. instantons on $T^2 \times \mathbb{R}^2$). One of the interesting points about the latter objects is that there is no concept of a finite-action solution anymore (the upshot is that the Green's function for the laplacian on $\mathbb{R}^2$ is $\log(r)$). See [CK01, Jar01] for more details.

2.5.2 Nahm transforms beyond dimensional reductions of $\mathbb{R}^4$

There has recently been much interest in developing Nahm transforms for instantons on less trivial backgrounds. The closest of these to what we have discussed are the examples of instantons on *gravitational instantons*, which are manifolds with self/anti-self-dual Levi-Civita connections. The Nahm transforms in these cases appear as a hybrid between the ADHM construction, and the Nahm transform on a circle, and have interesting relationships to the purely algebraic notion of quivers. See [Che11, Che10, Nak94] for more details.

References

[ADHM78] M F Atiyah, V G Drinfeld, N J Hitchin, and Y I Manin. Construction of instantons. *Phys. Lett. A*, 65(3):185–187, 1978.

[AHS78] M F Atiyah, N Js Hitchin, and I M Singer. Self-duality in four-dimensional riemannian geometry. *Proc. R. Soc. Lond. A*, 362(1711):425–461, 1978.

[AM89] M F Atiyah and N S Manton. Skyrmions from instantons. *Phys. Lett. B*, 222(3):438–442, 1989.

[AS13] J P Allen and P M Sutcliffe. ADHM polytopes. *J. High Energ. Phys.*, 2013(5):1–36, 2013.

[Ati78] M F Atiyah. Geometry of yang-mills fields. In *Mathematical problems in theoretical physics*, pages 216–221. Springer, 1978.

[BE18] H W Braden and V Z Enolski. The construction of monopoles. *Commun. Math. Phys.*, 362(2):547–570, 2018.

[Bec20] M Beckett. *Equivariant Nahm Transforms and Minimal Yang–Mills Connections*. PhD thesis, Duke University, 2020.

[BPST75] A Ao Belavin, A M Polyakov, A S Schwartz, and Yu S Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Phys. Lett. B*, 59(1):85–87, 1975.

[CF77] E Corrigan and D B Fairlie. Scalar field theory and exact solutions to a classical $SU(2)$ gauge theory. *Phys. Lett. B*, 67(1):69–71, 1977.

[CH07] B Charbonneau and J Hurtubise. The Nahm transform for calorons. *The many facets of geometry: a tribute to Nigel Hitchin (2010)*, 2007.

[CH19] B Charbonneau and J Hurtubise. Spatially periodic instantons: Nahm transform and moduli. *Commun. Math. Phys.*, 365(1):255–293, 2019.

[Che10] S A Cherkis. Instantons on the Taub-NUT space. *Adv. in Th. and Math. Phys.*, 14(2):609–642, 2010.

[Che11] S A Cherkis. Instantons on gravitons. *Commun. Math. Phys.*, 306(2):449–483, 2011.

[CK01] S Cherkis and A Kapustin. Nahm transform for periodic monopoles and $N = 2$ super Yang–Mills theory. *Commun. Math. Phys.*, 218(2):333–371, 2001.

[CL84] T-P Cheng and L-F Li. *Gauge theory of elementary particle physics*. Clarendon press Oxford, 1984.

[dC92] M P do Carmo. *Riemannian geometry*. Birkhäuser, 1992.

[DK97] S K Donaldson and P B Kronheimer. *The geometry of four-manifolds*. Oxford University Press, 1997.

[FH90] M Furuta and Y Hashimoto. Invariant instantons on S^4. *J. Fac. Sci. Univ. Tokyo*, 37:585–600, 1990.

[GS89] M Göckeler and T Schücker. *Differential geometry, gauge theories, and gravity*. Cambridge University Press, 1989.

[Har90] F R Harvey. *Spinors and calibrations*, volume 9. Elsevier, 1990.

[Hat02] A Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[Hit83] N J Hitchin. On the construction of monopoles. *Commun. Math. Phys.*, 89(2):145–190, 1983.

[HS78] B J Harrington and H K Shepard. Periodic euclidean solutions and the finite-temperature Yang-Mills gas. *Phys. Rev. D*, 17(8):2122, 1978.

[Jar01] M Jardim. Construction of doubly-periodic instantons. *Commun. Math. Phys.*, 216(1):1–15, 2001.

[Jar04] M Jardim. A survey on Nahm transform. *J. Geom. Phys.*, 52(3):313–327, 2004.

[KvB98] T C Kraan and P van Baal. Periodic instantons with non-trivial holonomy. *Nucl. Phys. B*, 533(1):627–659, 1998.

[Lan05] G D Landweber. Singular instantons with $SO(3)$ symmetry. *arXiv preprint math/0503611*, 2005.

[LL98] K Lee and C Lu. $SU(2)$ calorons and magnetic monopoles. *Phys. Rev. D*, 58(2):025011, 1998.

[MS04] N S Manton and P M Sutcliffe. *Topological solitons*. Cambridge University Press, 2004.

[Nah83] W Nahm. All self-dual multimonopoles for arbitrary gauge groups. In *Structural elements in particle physics and statistical mechanics*, pages 301–310. Springer, 1983.

[Nah84] W Nahm. Self-dual monopoles and calorons. In *Group theoretical methods in physics*, pages 189–200. Springer, 1984.

[Nak94] H Nakajima. Instantons on ALE spaces, quiver varieties, and Kac-Moody algebras. *Duke Mathematical Journal*, 76(2):365–416, 1994.

[Nye01] T M W Nye. *The geometry of calorons*. PhD thesis, The University of Edinburgh, 2001.

[PS75] M K Prasad and C M Sommerfield. Exact classical solution for the 't Hooft monopole and the Julia-Zee dyon. *Phys. Rev. Lett.*, 35(12):760, 1975.

[Spi70] M D Spivak. *A comprehensive introduction to differential geometry*. Publish or perish, 1970.

[SS98] T Schäfer and E V Shuryak. Instantons in QCD. *Rev. Mod. Phys.*, 70(2):323, 1998.

[SS05] T Sakai and S Sugimoto. Low energy hadron physics in holographic QCD. *Prog. Th. Phys.*, 113(4):843–882, 2005.

[Ste10] M Stern. Geometry of minimal energy yang-mills connections. *J. Diff. Geom.*, 86(1):163–188, 2010.

[Sut04] P M Sutcliffe. Instantons and the buckyball. In *Proc. R. Soc. Lon. A: Mathematical, Physical and Engineering Sciences*, volume 460, pages 2903–2912. The Royal Society, 2004.

[Sut10] P Sutcliffe. Skyrmions, instantons and holography. *J. High Energ. Phys.*, 2010(8):19, 2010.

[Tau11] C H Taubes. *Differential geometry: Bundles, connections, metrics and curvature*, volume 23. Oxford University Press, 2011.

[tH] G 't Hooft. unpublished.

[Uhl82] K K Uhlenbeck. Removable singularities in Yang-Mills fields. *Commun. Math. Phys.*, 83(1):11–29, 1982.

[Wit77] E Witten. Some exact multipseudoparticle solutions of classical Yang-Mills theory. *Phys. Rev. Lett.*, 38(3):121, 1977.