Preprint typeset in JHEP style - PAPER VERSION

June 2005

TASI Lectures on Solitons

Instantons, Monopoles, Vortices and Kinks

David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OBA, UK
d.tong@damtp.cam.ac.uk

Recommended Books and Resources These lectures cover aspects of solitons with focus on applications to the quantum dynamics of supersymmetric gauge theories and string theory. The lectures consist of four sections, each dealing with a different soliton. We start with instantons and work down in co-dimension to monopoles, vortices and, eventually, domain walls. Emphasis is placed on the moduli space of solitons and, in particular, on the web of connections that links solitons of different types. The D-brane realization of the ADHM and Nahm construction for instantons and monopoles is reviewed, together with related constructions for vortices and domain walls. Each lecture ends with a series of vignettes detailing the roles solitons play in the quantum dynamics of supersymmetric gauge theories in various dimensions. This includes applications to the AdS/CFT correspondence, little string theory, S-duality, cosmic strings, and the quantitative correspondence between 2d sigma models and 4d gauge theories.

Contents

0. Introduction

170 years ago, a Scotsman on horseback watched a wave travelling down Edinburgh's Union canal. He was so impressed that he followed the wave for several miles, described the day of observation as the happiest of his life, and later attempted to recreate the experience in his own garden. The man's name was John Scott Russell and he is generally credited as the first person to develop an unhealthy obsession with the "singular and beautiful phenomenon" that we now call a soliton.

Russell was ahead of his time. The features of stability and persistence that so impressed him were not appreciated by his contemporaries, with Airy arguing that the "great primary wave" was neither great nor primary[1]. It wasn't until the following century that solitons were understood to play an important role in areas ranging from engineering to biology, from condensed matter to cosmology.

The purpose of these lectures is to explore the properties of solitons in gauge theories. There are four leading characters: the instanton, the monopole, the vortex, and the domain wall (also known as the kink). Most reviews of solitons start with kinks and work their way up to the more complicated instantons. Here we're going to do things backwards and follow the natural path: instantons are great and primary, other solitons follow. A major theme of these lectures is to flesh out this claim by describing the web of inter-relationships connecting our four solitonic characters.

Each lecture will follow a similar pattern. We start by deriving the soliton equations and examining the basic features of the simplest solution. We then move on to discuss the interactions of multiple solitons, phrased in terms of the moduli space. For each type of soliton, D-brane techniques are employed to gain a better understanding of the relevant geometry. Along the way, we shall discuss various issues including fermionic zero modes, dyonic excitations and non-commutative solitons. We shall also see the earlier solitons reappearing in surprising places, often nestling within the worldvolume of a larger soliton, with interesting consequences. Each lecture concludes with a few brief descriptions of the roles solitons play in supersymmetric gauge theories in various dimensions.

These notes are aimed at advanced graduate students who have some previous awareness of solitons. The basics will be covered, but only very briefly. A useful primer on solitons can be found in most modern field theory textbooks (see for example [1]). More

[1]More background on Russell and his wave can be found at http://www.ma.hw.ac.uk/~chris/scott_russell.html and http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Russell_Scott.html.

details are contained in the recent book by Manton and Sutcliffe [2]. There are also a number of good reviews dedicated to solitons of a particular type and these will be mentioned at the beginning of the relevant lecture. Other background material that will be required for certain sections includes a basic knowledge of the structure of supersymmetric gauge theories and D-brane dynamics. Good reviews of these subjects can be found in [3, 4, 5].

1. Instantons

30 years after the discovery of Yang-Mills instantons [6], they continue to fascinate both physicists and mathematicians alike. They have lead to new insights into a wide range of phenomena, from the structure of the Yang-Mills vacuum [7, 8, 9] to the classification of four-manifolds [10]. One of the most powerful uses of instantons in recent years is in the analysis of supersymmetric gauge dynamics where they play a key role in unravelling the plexus of entangled dualities that relates different theories. The purpose of this lecture is to review the classical properties of instantons, ending with some applications to the quantum dynamics of supersymmetric gauge theories.

There exist many good reviews on the subject of instantons. The canonical reference for basics of the subject remains the beautiful lecture by Coleman [11]. More recent applications to supersymmetric theories are covered in detail in reviews by Shifman and Vainshtein [12] and by Dorey, Hollowood, Khoze and Mattis [13]. This latter review describes the ADHM construction of instantons and overlaps with the current lecture.

1.1 The Basics

The starting point for our journey is four-dimensional, pure $SU(N)$ Yang-Mills theory with action[2]

$$S = \frac{1}{2e^2} \int d^4x \ \mathrm{Tr}\, F_{\mu\nu} F^{\mu\nu} \tag{1.1}$$

Motivated by the semi-classical evaluation of the path integral, we search for finite action solutions to the Euclidean equations of motion,

$$\mathcal{D}_\mu F^{\mu\nu} = 0 \tag{1.2}$$

which, in the imaginary time formulation of the theory, have the interpretation of mediating quantum mechanical tunnelling events.

The requirement of finite action means that the potential A_μ must become pure gauge as we head towards the boundary $r \to \infty$ of spatial $\mathbf{R}^4$,

$$A_\mu \to i g^{-1} \partial_\mu g \tag{1.3}$$

[2]Conventions: We pick Hemitian generators T^m with Killing form $\mathrm{Tr}\, T^m T^n = \frac{1}{2}\delta^{mn}$. We write $A_\mu = A_\mu^m T^m$ and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]$. Adjoint covariant derivatives are $\mathcal{D}_\mu X = \partial_\mu X - i[A_\mu, X]$. In this section alone we work with Euclidean signature and indices will wander from top to bottom with impunity; in the following sections we will return to Minkowski space with signature $(+,-,-,-)$.

with $g(x) = e^{iT(x)} \in SU(N)$. In this way, any finite action configuration provides a map from $\partial\mathbf{R}^4 \cong \mathbf{S}^3_\infty$ into the group $SU(N)$. As is well known, such maps are classified by homotopy theory. Two maps are said to lie in the same homotopy class if they can be continuously deformed into each other, with different classes labelled by the third homotopy group,

$$\Pi_3(SU(N)) \cong \mathbf{Z} \tag{1.4}$$

The integer $k \in \mathbf{Z}$ counts how many times the group wraps itself around spatial $\mathbf{S}^3_\infty$ and is known as the Pontryagin number, or second Chern class. We will sometimes speak simply of the "charge" k of the instanton. It is measured by the surface integral

$$k = \frac{1}{24\pi^2}\int_{\mathbf{S}^3_\infty} d^3S_\mu \ \mathrm{Tr}\,(\partial_\nu g)g^{-1}\,(\partial_\rho g)g^{-1}\,(\partial_\sigma g)g^{-1}\,\epsilon^{\mu\nu\rho\sigma} \tag{1.5}$$

The charge k splits the space of field configurations into different sectors. Viewing $\mathbf{R}^4$ as a foliation of concentric $\mathbf{S}^3$'s, the homotopy classification tells us that we cannot transform a configuration with non-trivial winding $k \neq 0$ at infinity into one with trivial winding on an interior $\mathbf{S}^3$ while remaining in the pure gauge ansatz (1.3). Yet, at the origin, obviously the gauge field must be single valued, independent of the direction from which we approach. To reconcile these two facts, a configuration with $k \neq 0$ cannot remain in the pure gauge form (1.3) throughout all of $\mathbf{R}^4$: it must have non-zero action.

An Example: $SU(2)$

The simplest case to discuss is the gauge group $SU(2)$ since, as a manifold, $SU(2) \cong \mathbf{S}^3$ and it's almost possible to visualize the fact that $\Pi_3(\mathbf{S}^3) \cong \mathbf{Z}$. (Ok, maybe $\mathbf{S}^3$ is a bit of a stretch, but it is possible to visualize $\Pi_1(\mathbf{S}^1) \cong \mathbf{Z}$ and $\Pi_2(\mathbf{S}^2) \cong \mathbf{Z}$ and it's not the greatest leap to accept that, in general, $\Pi_n(\mathbf{S}^n) \cong \mathbf{Z}$). Examples of maps in the different sectors are

- $g^{(0)} = 1$, the identity map has winding $k = 0$
- $g^{(1)} = (x_4 + ix_i\sigma^i)/r$ has winding number $k = 1$. Here $i = 1, 2, 3$, and the σ^i are the Pauli matrices
- $g^{(k)} = [g^{(1)}]^k$ has winding number k.

To create a non-trivial configuration in $SU(N)$, we could try to embed the maps above into a suitable $SU(2)$ subgroup, say the upper left-hand corner of the $N \times N$ matrix. It's not obvious that if we do this they continue to be a maps with non-trivial winding since one could envisage that they now have space to slip off. However, it turns out that this doesn't happen and the above maps retain their winding number when embedded in higher rank gauge groups.

1.1.1 The Instanton Equations

We have learnt that the space of configurations splits into different sectors, labelled by their winding $k \in \mathbf{Z}$ at infinity. The next question we want to ask is whether solutions actually exist for different k. Obviously for $k = 0$ the usual vacuum $A_\mu = 0$ (or gauge transformations thereof) is a solution. But what about higher winding with $k \neq 0$? The first step to constructing solutions is to derive a new set of equations that the instantons will obey, equations that are first order rather than second order as in (1.2). The trick for doing this is usually referred to as the Bogomoln'yi bound [14] although, in the case of instantons, it was actually introduced in the original paper [6]. From the above considerations, we have seen that any configuration with $k \neq 0$ must have some non-zero action. The Bogomoln'yi bound quantifies this. We rewrite the action by completing the square,

$$\begin{aligned} S_{\rm inst} &= \frac{1}{2e^2}\int d^4x\ {\rm Tr}\, F_{\mu\nu}F^{\mu\nu} \\ &= \frac{1}{4e^2}\int d^4x\ {\rm Tr}\,(F_{\mu\nu} \mp {}^\star F^{\mu\nu})^2 \pm 2{\rm Tr}\ F_{\mu\nu}{}^\star F^{\mu\nu} \\ &\geq \pm\frac{1}{2e^2}\int d^4x\ \partial_\mu\left(A_\nu F_{\rho\sigma} + \tfrac{2i}{3}A_\nu A_\rho A_\sigma\right)\epsilon^{\mu\nu\rho\sigma} \end{aligned} \tag{1.6}$$

where the dual field strength is defined as ${}^\star F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$ and, in the final line, we've used the fact that $F_{\mu\nu}{}^\star F^{\mu\nu}$ can be expressed as a total derivative. The final expression is a surface term which measures some property of the field configuration on the boundary $\mathbf{S}^3_\infty$. Inserting the asymptotic form $A_\nu \to ig^{-1}\partial_\nu g$ into the above expression and comparing with (1.5), we learn that the action of the instanton in a topological sector k is bounded by

$$S_{\rm inst} \geq \frac{8\pi^2}{e^2}\,|k| \tag{1.7}$$

with equality if and only if

$$\begin{aligned} F_{\mu\nu} &= {}^\star F_{\mu\nu} \qquad\qquad (k>0) \\ F_{\mu\nu} &= -{}^\star F_{\mu\nu} \qquad\qquad (k<0) \end{aligned}$$

Since parity maps $k \to -k$, we can focus on the self-dual equations $F = {}^\star F$. The Bogomoln'yi argument (which we shall see several more times in later sections) says that a solution to the self-duality equations must necessarily solve the full equations of motion since it minimizes the action in a given topological sector. In fact, in the case of instantons, it's trivial to see that this is the case since we have

$$\mathcal{D}_\mu F^{\mu\nu} = \mathcal{D}_\mu {}^\star F^{\mu\nu} = 0 \tag{1.8}$$

by the Bianchi identity.

1.1.2 Collective Coordinates

So we now know the equations we should be solving to minimize the action. But do solutions exist? The answer, of course, is yes! Let's start by giving an example, before we move on to examine some of its properties, deferring discussion of the general solutions to the next subsection.

The simplest solution is the $k = 1$ instanton in $SU(2)$ gauge theory. In singular gauge, the connection is given by

$$A_\mu = \frac{\rho^2 (x-X)_\nu}{(x-X)^2((x-X)^2+\rho^2)}\,\bar{\eta}^i_{\mu\nu}\,(g\sigma^i g^{-1}) \tag{1.9}$$

The σ^i, $i = 1, 2, 3$ are the Pauli matrices and carry the $su(2)$ Lie algebra indices of A_μ. The $\bar{\eta}^i$ are three 4×4 anti-self-dual 't Hooft matrices which intertwine the group structure of the index i with the spacetime structure of the indices μ, ν. They are given by

$$\bar{\eta}^1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \;,\quad \bar{\eta}^2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \;,\quad \bar{\eta}^3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.10}$$

It's a useful exercise to compute the field strength to see how it inherits its self-duality from the anti-self-duality of the $\bar{\eta}$ matrices. To build an anti-self-dual field strength, we need to simply exchange the $\bar{\eta}$ matrices in (1.9) for their self-dual counterparts,

$$\eta^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \;,\quad \eta^2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \;,\quad \eta^3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{1.11}$$

For our immediate purposes, the most important feature of the solution (1.9) is that it is not unique: it contains a number of parameters. In the context of solitons, these are known as *collective coordinates*. The solution (1.9) has eight such parameters. They are of three different types:

i) 4 translations X_μ: The instanton is an object localized in $\mathbf{R}^4$, centered around the point $x_\mu = X_\mu$.

ii) 1 scale size ρ: The interpretation of ρ as the size of the instanton can be seen by rescaling x and X in the above solution to demote ρ to an overall constant.

iii) 3 global gauge transformations $g \in SU(2)$: This determines how the instanton is embedded in the gauge group.

At this point it's worth making several comments about the solution and its collective coordinates.

- For the $k = 1$ instanton, each of the collective coordinates described above is a Goldstone mode, arising because the instanton configuration breaks a symmetry of the Lagrangian (1.1). In the case of X_μ and g it is clear that the symmetry is translational invariance and $SU(2)$ gauge invariance respectively. The parameter ρ arises from broken conformal invariance. It's rather common that all the collective coordinates of a single soliton are Goldstone modes. It's not true for higher k.

- The apparent singularity at $x_\mu = X_\mu$ is merely a gauge artifact (hence the name "singular gauge"). A plot of a gauge invariant quantity, such as the action density, reveals a smooth solution. The exception is when the instanton shrinks to zero size $\rho \to 0$. This singular configuration is known as the small instanton. Despite its singular nature, it plays an important role in computing the contribution to correlation functions in supersymmetric theories. The small instanton lies at finite distance in the space of classical field configurations (in a way which will be made precise in Section 1.2).

- You may be surprised that we are counting the gauge modes g as physical parameters of the solution. The key point is that they arise from the *global* part of the gauge symmetry, meaning transformations that don't die off asymptotically. These are physical symmetries of the system rather than redundancies. In the early days of studying instantons the 3 gauge modes weren't included, but it soon became apparent that many of the nicer mathematical properties of instantons (for example, hyperKählerity of the moduli space) require us to include them, as do certain physical properties (for example, dyonic instantons in five dimensions)

The $SU(2)$ solution (1.9) has 8 collective coordinates. What about $SU(N)$ solutions? Of course, we should keep the $4+1$ translational and scale parameters but we would expect more orientation parameters telling us how the instanton sits in the larger

$SU(N)$ gauge group. How many? Suppose we embed the above $SU(2)$ solution in the upper left-hand corner of an $N \times N$ matrix. We can then rotate this into other embeddings by acting with $SU(N)$, modulo the stabilizer which leaves the configuration untouched. We have

$$SU(N)/S[U(N-2) \times U(2)] \tag{1.12}$$

where the $U(N-2)$ hits the lower-right-hand corner and doesn't see our solution, while the $U(2)$ is included in the denominator since it acts like g in the original solution (1.9) and we don't want to overcount. Finally, the notation $S[U(p) \times U(q)]$ means that we lose the overall central $U(1) \subset U(p) \times U(q)$. The coset space above has dimension $4N - 8$. So, within the ansatz (1.9) embedded in $SU(N)$, we see that the $k = 1$ solution has $4N$ collective coordinates. In fact, it turns out that this is all of them and the solution (1.9), suitably embedded, is the most general $k = 1$ solution in an $SU(N)$ gauge group. But what about solutions with higher k? To discuss this, it's useful to introduce the idea of the moduli space.

1.2 The Moduli Space

We now come to one of the most important concepts of these lectures: the *moduli space*. This is defined to be the space of all solutions to $F = {}^\star F$, modulo gauge transformations, in a given winding sector k and gauge group $SU(N)$. Let's denote this space as $\mathcal{I}_{k,N}$. We will define similar moduli spaces for the other solitons and much of these lectures will be devoted to understanding the different roles these moduli spaces play and the relationships between them.

Coordinates on $\mathcal{I}_{k,N}$ are given by the collective coordinates of the solution. We've seen above that the $k = 1$ solution has $4N$ collective coordinates or, in other words, $\dim(\mathcal{I}_{1,N}) = 4N$. For higher k, the number of collective coordinates can be determined by index theorem techniques. I won't give all the details, but will instead simply tell you the answer.

$$\dim(\mathcal{I}_{k,N}) = 4kN \tag{1.13}$$

This has a very simple interpretation. The charge k instanton can be thought of as k charge 1 instantons, each with its own position, scale, and gauge orientation. When the instantons are well separated, the solution does indeed look like this. But when instantons start to overlap, the interpretation of the collective coordinates can become more subtle.

Strictly speaking, the index theorem which tells us the result (1.13) doesn't count the number of collective coordinates, but rather related quantities known as *zero modes*. It works as follows. Suppose we have a solution A_μ satisfying $F = {}^\star F$. Then we can perturb this solution $A_\mu \rightarrow A_\mu + \delta A_\mu$ and ask how many other solutions are nearby. We require the perturbation δA_μ to satisfy the linearized self-duality equations,

$$\mathcal{D}_\mu \delta A_\nu - \mathcal{D}_\nu \delta A_\mu = \epsilon_{\mu\nu\rho\sigma} \mathcal{D}^\rho \delta A^\sigma \tag{1.14}$$

where the covariant derivative $\mathcal{D}_\mu$ is evaluated on the background solution. Solutions to (1.14) are called zero modes. The idea of zero modes is that if we have a general solution $A_\mu = A_\mu(x_\mu, X^\alpha)$, where X^α denote all the collective coordinates, then for each collective coordinate we can define the zero mode $\delta_\alpha A_\mu = \partial A_\mu / \partial X^\alpha$ which will satisfy (1.14). In general however, it is not guaranteed that any zero mode can be successfully integrated to give a corresponding collective coordinate. But it will turn out that all the solitons discussed in these lectures do have this property (at least this is true for bosonic collective coordinates; there is a subtlety with the Grassmannian collective coordinates arising from fermions which we'll come to shortly).

Of course, any local gauge transformation will also solve the linearized equations (1.14) so we require a suitable gauge fixing condition. We'll write each zero mode to include an infinitesimal gauge transformation Ω_α,

$$\delta_\alpha A_\mu = \frac{\partial A_\mu}{\partial X^\alpha} + \mathcal{D}_\mu \Omega_\alpha \tag{1.15}$$

and choose Ω_α so that $\delta_\alpha A_\mu$ is orthogonal to any other gauge transformation, meaning

$$\int d^4x \ \mathrm{Tr}\,(\delta_\alpha A_\mu)\,\mathcal{D}_\mu \eta = 0 \qquad \forall\ \eta \tag{1.16}$$

which, integrating by parts, gives us our gauge fixing condition

$$\mathcal{D}_\mu\,(\delta_\alpha A_\mu) = 0 \tag{1.17}$$

This gauge fixing condition does not eliminate the collective coordinates arising from global gauge transformations which, on an operational level, gives perhaps the clearest reason why we must include them. The Atiyah-Singer index theorem counts the number of solutions to (1.14) and (1.17) and gives the answer (1.13).

So what does the most general solution, with its $4kN$ parameters, look like? The general explicit form of the solution is not known. However, there are rather clever ansätze which give rise to various subsets of the solutions. Details can be found in the original literature [15, 16] but, for now, we head in a different, and ultimately more important, direction and study the geometry of the moduli space.

1.2.1 The Moduli Space Metric

A priori, it is not obvious that $\mathcal{I}_{k,N}$ is a manifold. In fact, it does turn out to be a smooth space apart from certain localized singularities corresponding to small instantons at $\rho \to 0$ where the field configuration itself also becomes singular.

The moduli space $\mathcal{I}_{k,N}$ inherits a natural metric from the field theory, defined by the overlap of zero modes. In the coordinates X^α, $\alpha = 1, \ldots, 4kN$, the metric is given by

$$g_{\alpha\beta} = \frac{1}{2e^2} \int d^4x \; \mathrm{Tr} \; (\delta_\alpha A_\mu) \; (\delta_\beta A_\mu) \tag{1.18}$$

It's hard to overstate the importance of this metric. It distills the information contained in the solutions to $F = {}^\star F$ into a more manageable geometric form. It turns out that for many applications, everything we need to know about the instantons is contained in the metric $g_{\alpha\beta}$, and this remains true of similar metrics that we will define for other solitons. Moreover, it is often much simpler to determine the metric (1.18) than it is to determine the explicit solutions.

The metric has a few rather special properties. Firstly, it inherits certain isometries from the symmetries of the field theory. For example, both the $SO(4)$ rotation symmetry of spacetime and the $SU(N)$ gauge action will descend to give corresponding isometries of the metric $g_{\alpha\beta}$ on $\mathcal{I}_{k,N}$.

Another important property of the metric (1.18) is that it is *hyperKähler*, meaning that the manifold has reduced holonomy $Sp(kN) \subset SO(4kN)$. Heuristically, this means that the manifold admits something akin to a quaternionic structure[3]. More precisely, a hyperKähler manifold admits three complex structures J^i, $i = 1, 2, 3$ which obey the relation

$$J^i \, J^j = -\delta^{ij} + \epsilon^{ijk} \, J^k \tag{1.19}$$

The simplest example of a hyperKähler manifold is $\mathbf{R}^4$, viewed as the quaternions. The three complex structures can be taken to be the anti-self-dual 't Hooft matrices $\bar{\eta}^i$ that we defined in (1.10), each of which gives a different complex pairing of $\mathbf{R}^4$. For example, from $\bar{\eta}^3$ we get $z^1 = x^1 + ix^2$ and $z^2 = x^3 - ix^4$.

[3]Warning: there is also something called a quaternionic manifold which arises in $\mathcal{N} = 2$ supergravity theories [17] and is different from a hyperKähler manifold. For a discussion on the relationship see [18].

The instanton moduli space $\mathcal{I}_{k,N}$ inherits its complex structures J^i from those of $\mathbf{R}^4$. To see this, note if δA_μ is a zero mode, then we may immediately write down three other zero modes $\bar{\eta}^i_{\nu\mu}\,\delta A_\mu$, each of which satisfy the equations (1.14) and (1.17). It must be possible to express these three new zero modes as a linear combination of the original ones, allowing us to define three matrices J^i,

$$\bar{\eta}^i_{\mu\nu}\;\delta_\beta A_\nu = (J^i)^\alpha_{\ \beta}\;[\delta_\alpha A_\mu] \tag{1.20}$$

These matrices J^i then descend to three complex structures on the moduli space $\mathcal{I}_{k,N}$ itself which are given by

$$(J^i)^\alpha_{\ \beta} = g^{\alpha\gamma}\;\int d^4x\;\bar{\eta}^i_{\mu\nu}\;\mathrm{Tr}\;\delta_\beta A_\mu\;\delta_\gamma A_\nu \tag{1.21}$$

So far we have shown only that J^i define almost complex structures. To prove hyperKählerity, one must also show integrability which, after some gymnastics, is possible using the formulae above. A more detailed discussion of the geometry of the moduli space in this language can be found in [19, 20] and more generally in [21, 22]. For physicists the simplest proof of hyperKählerity follows from supersymmetry as we shall review in section 1.3.

It will prove useful to return briefly to discuss the isometries. In Kähler and hyperKähler manifolds, it's often important to state whether isometries are compatible with the complex structure J. If the complex structure doesn't change as we move along the isometry, so that the Lie derivative $\mathcal{L}_k J = 0$, with k the Killing vector, then the isometry is said to be *holomorphic*. In the instanton moduli space $\mathcal{I}_{k,N}$, the $SU(N)$ gauge group action is tri-holomorphic, meaning it preserves all three complex structures. Of the $SO(4) \cong SU(2)_L \times SU(2)_R$ rotational symmetry, one half, $SU(2)_L$, is tri-holomorphic, while the three complex structures are rotated under the remaining $SU(2)_R$ symmetry.

1.2.2 An Example: A Single Instanton in $SU(2)$

In the following subsection we shall show how to derive metrics on $\mathcal{I}_{k,N}$ using the powerful ADHM technique. But first, to get a flavor for the ideas, let's take a more pedestrian route for the simplest case of a $k=1$ instanton in $SU(2)$. As we saw above, there are three types of collective coordinates.

i) The four translational modes are $\delta_{(\nu)}A_\mu = \partial A_\mu/\partial X^\nu + \mathcal{D}_\mu\Omega_\nu$ where Ω_ν must be chosen to satisfy (1.17). Using the fact that $\partial/\partial X^\nu = -\partial/\partial x^\nu$, it is simple to see that the correct choice of gauge is $\Omega_\nu = A_\nu$, so that the zero mode is simply given by $\delta_\nu A_\mu = F_{\mu\nu}$, which satisfies the gauge fixing condition by virtue of the original equations of motion (1.2). Computing the overlap of these translational zero modes then gives

$$\int d^4x \;\mathrm{Tr}\;(\delta_{(\nu)}A_\mu\,\delta_{(\rho)}A_\mu) = S_{\rm inst}\;\delta_{\nu\rho} \tag{1.22}$$

ii) One can check that the scale zero mode $\delta A_\mu = \partial A_\mu/\partial\rho$ already satisfies the gauge fixing condition (1.17) when the solution is taken in singular gauge (1.9). The overlap integral in this case is simple to perform, yielding

$$\int d^4x \;\mathrm{Tr}\;(\delta A_\mu\,\delta A_\mu) = 2S_{\rm inst} \tag{1.23}$$

iii) Finally, we have the gauge orientations. These are simply of the form $\delta A_\mu = \mathcal{D}_\mu\Lambda$, but where Λ does not vanish at infinity, so that it corresponds to a global gauge transformation. In singular gauge it can be checked that the three $SU(2)$ rotations $\Lambda^i = [(x-X)^2/((x-X)^2+\rho^2)]\sigma^i$ satisfy the gauge fixing constraint. These give rise to an $SU(2)\cong\mathbf{S}^3$ component of the moduli space with radius given by the norm of any one mode, say, Λ^3

$$\int d^4x \;\mathrm{Tr}\;(\delta A_\mu\,\delta A_\mu) = 2S_{\rm inst}\,\rho^2 \tag{1.24}$$

Note that, unlike the others, this component of the metric depends on the collective coordinate ρ, growing as ρ^2. This dependence means that the $\mathbf{S}^3$ arising from $SU(2)$ gauge rotations combines with the $\mathbf{R}^+$ from scale transformations to form the space $\mathbf{R}^4$. However, there is a discrete subtlety. Fields in the adjoint representation are left invariant under the center $Z_2\subset SU(2)$, meaning that the gauge rotations give rise to $\mathbf{S}^3/\mathbf{Z}_2$ rather than $\mathbf{S}^3$. Putting all this together, we learn that the moduli space of a single instanton is

$$\mathcal{I}_{1,2}\cong\mathbf{R}^4\times\mathbf{R}^4/\mathbf{Z}_2 \tag{1.25}$$

where the first factor corresponds to the position of the instanton, and the second factor determines its scale size and $SU(2)$ orientation. The normalization of the flat metrics on the two $\mathbf{R}^4$ factors is given by (1.22) and (1.23). In this case, the hyperKähler structure on $\mathcal{I}_{1,2}$ comes simply by viewing each $\mathbf{R}^4\cong\mathbb{H}$, the quaternions. As is clear from our derivation, the singularity at the origin of the orbifold $\mathbf{R}^4/\mathbf{Z}_2$ corresponds to the small instanton $\rho\to 0$.

1.3 Fermi Zero Modes

So far we've only concentrated on the pure Yang-Mills theory (1.1). It is natural to wonder about the possibility of other fields in the theory: could they also have non-trivial solutions in the background of an instanton, leading to further collective coordinates? It turns out that this doesn't happen for bosonic fields (although they do have an important impact if they gain a vacuum expectation value as we shall review in later sections). Importantly, the fermions do contribute zero modes.

Consider a single Weyl fermion λ transforming in the adjoint representation of $SU(N)$, with kinetic term $i\mathrm{Tr}\,\bar{\lambda}\,\bar{\not{\mathcal{D}}}\lambda$. In Euclidean space, we treat λ and $\bar{\lambda}$ as independent variables, a fact which leads to difficulties in defining a real action. (For the purposes of this lecture, we simply ignore the issue - a summary of the problem and its resolutions can be found in [13]). The equations of motion are

$$\bar{\not{\mathcal{D}}}\lambda \equiv \bar{\sigma}^\mu \mathcal{D}_\mu \lambda = 0 \quad , \qquad \not{\mathcal{D}}\bar{\lambda} \equiv \sigma^\mu \mathcal{D}_\mu \bar{\lambda} = 0 \tag{1.26}$$

where $\not{\mathcal{D}} = \sigma^\mu \mathcal{D}_\mu$ and the 2×2 matrices are $\sigma^\mu = (\sigma^i, -i1_2)$. In the background of an instanton $F = {}^\star F$, only λ picks up zero modes. $\bar{\lambda}$ has none. This situation is reversed in the background of an anti-instanton $F = -{}^\star F$. To see that $\bar{\lambda}$ has no zero modes in the background of an instanton, we look at

$$\bar{\not{\mathcal{D}}}\,\not{\mathcal{D}} = \bar{\sigma}^\mu \sigma^\nu D_\mu \mathcal{D}_\nu = \mathcal{D}^2\, 1_2 + F^{\mu\nu}\bar{\eta}^i_{\mu\nu}\sigma^i \tag{1.27}$$

where $\bar{\eta}^i$ are the anti-self-dual 't Hooft matrices defined in (1.10). But a self-dual matrix $F_{\mu\nu}$ contracted with an anti-self-dual matrix $\bar{\eta}_{\mu\nu}$ vanishes, leaving us with $\bar{\not{\mathcal{D}}}\,\not{\mathcal{D}} = \mathcal{D}^2$. And the positive definite operator $\mathcal{D}^2$ has no zero modes. In contrast, if we try to repeat the calculation for λ, we find

$$\not{\mathcal{D}}\,\bar{\not{\mathcal{D}}} = \mathcal{D}^2\, 1_2 + F^{\mu\nu}\eta^i_{\mu\nu}\sigma^i \tag{1.28}$$

where η^i are the self-dual 't Hooft matrices (1.11). Since we cannot express the operator $\not{\mathcal{D}}\,\bar{\not{\mathcal{D}}}$ as a total square, there's a chance that it has zero modes. The index theorem tells us that each Weyl fermion λ picks up $4kN$ zero modes in the background of a charge k instanton. There are corresponding Grassmann collective coordinates, which we shall denote as χ, associated to the most general solution for the gauge field and fermions. But these Grassmann collective coordinates occasionally have subtle properties. The quick way to understand this is in terms of supersymmetry. And often the quick way to understand the full power of supersymmetry is to think in higher dimensions.

1.3.1 Dimension Hopping

It will prove useful to take a quick break in order to make a few simple remarks about instantons in higher dimensions. So far we've concentrated on solutions to the self-duality equations in four-dimensional theories, which are objects localized in Euclidean spacetime. However, it is a simple matter to embed the solutions in higher dimensions simply by insisting that all fields are independent of the new coordinates. For example, in $d = 4 + 1$ dimensional theories one can set $\partial_0 = A_0 = 0$, with the spatial part of the gauge field satisfying $F = {}^\star F$. Such configurations have finite energy and the interpretation of particle like solitons. We shall describe some of their properties when we come to applications. Similarly, in $d = 5 + 1$, the instantons are string like objects, while in $d = 9 + 1$, instantons are five-branes. While this isn't a particularly deep insight, it's a useful trick to keep in mind when considering the fermionic zero modes of the soliton in supersymmetric theories as we shall discuss shortly.

When solitons have a finite dimensional worldvolume, we can promote the collective coordinates to fields which depend on the worldvolume directions. These correspond to massless excitations living on the solitons. For example, allowing the translational modes to vary along the instanton string simply corresponds to waves propagating along the string. Again, this simple observation will become rather powerful when viewed in the context of supersymmetric theories.

A note on terminology: Originally the term "instanton" referred to solutions to the self-dual Yang-Mills equations $F = {}^\star F$. (At least this was true once Physical Review lifted its censorship of the term!). However, when working with theories in spacetime dimensions other than four, people often refer to the relevant finite action configuration as an instanton. For example, kinks in quantum mechanics are called instantons. Usually this doesn't lead to any ambiguity but in this review we'll consider a variety of solitons in a variety of dimensions. I'll try to keep the phrase "instanton" to refer to (anti)-self-dual Yang-Mills instantons.

1.3.2 Instantons and Supersymmetry

Instantons share an intimate relationship with supersymmetry. Let's consider an instanton in a $d = 3 + 1$ supersymmetric theory which could be either $\mathcal{N} = 1$, $\mathcal{N} = 2$ or $\mathcal{N} = 4$ super Yang-Mills. The supersymmetry transformation for any adjoint Weyl fermion takes the form

$$\delta\lambda = F^{\mu\nu}\sigma_\mu\bar{\sigma}_\nu\epsilon \quad , \quad \delta\bar{\lambda} = F^{\mu\nu}\bar{\sigma}_\mu\sigma_\nu\bar{\epsilon} \tag{1.29}$$

where, again, we treat the infinitesimal supersymmetry parameters ϵ and $\bar{\epsilon}$ as independent. But we've seen above that in the background of a self-dual solution $F = {}^\star F$

the combination $F^{\mu\nu}\bar{\sigma}_\mu\sigma_\nu = 0$. This means that the instanton is annihilated by half of the supersymmetry transformations $\bar{\epsilon}$, while the other half, ϵ, turn on the fermions λ. We say that the supersymmetries arising from ϵ are broken by the soliton, while those arising from $\bar{\epsilon}$ are preserved. Configurations in supersymmetric theories which are annihilated by some fraction of the supersymmetries are known as BPS states (although the term Witten-Olive state would be more appropriate [23]).

Both the broken and preserved supersymmetries play an important role for solitons. The broken ones are the simplest to describe, for they generate fermion zero modes $\lambda = F^{\mu\nu}\sigma_\mu\bar{\sigma}_\nu\epsilon$. These "Goldstino" modes are a subset of the $4kN$ fermion zero modes that exist for each Weyl fermion λ. Further modes can also be generated by acting on the instanton with superconformal transformations.

The unbroken supersymmetries $\bar{\epsilon}$ play a more important role: they descend to a supersymmetry on the soliton worldvolume, pairing up bosonic collective coordinates X with Grassmannian collective coordinates χ. There's nothing surprising here. It's simply the statement that if a symmetry is preserved in a vacuum (where, in this case, the "vacuum" is the soliton itself) then all excitations above the vacuum fall into representations of this symmetry. However, since supersymmetry in $d = 0 + 0$ dimensions is a little subtle, and the concept of "excitations above the vacuum" in $d = 0 + 0$ dimensions even more so, this is one of the places where it will pay to lift the instantons to higher dimensional objects. For example, instantons in theories with 8 supercharges (equivalent to $\mathcal{N} = 2$ in four dimensions) can be lifted to instanton strings in six dimensions, which is the maximum dimension in which Yang-Mills theory with eight supercharges exists. Similarly, instantons in theories with 16 supercharges (equivalent to $\mathcal{N} = 4$ in four dimensions) can be lifted to instanton five-branes in ten dimensions. Instantons in $\mathcal{N} = 1$ theories are stuck in their four-dimensional world.

Considering Yang-Mills instantons as solitons in higher dimensions allows us to see this relationship between bosonic and fermionic collective coordinates. Consider exciting a long-wavelength mode of the soliton in which a bosonic collective coordinate X depends on the worldvolume coordinate of the instanton s, so $X = X(s)$. Then if we hit this configuration with the unbroken supersymmetry $\bar{\epsilon}$, it will no longer annihilate the configuration, but will turn on a fermionic mode proportional to $\partial_s X$. Similarly, any fermionic excitation will be related to a bosonic excitation.

The observation that the unbroken supersymmetries descend to supersymmetries on the worldvolume of the soliton saves us a lot of work in analyzing fermionic zero modes: if we understand the bosonic collective coordinates and the preserved supersymmetry, then the fermionic modes pretty much come for free. This includes some rather subtle interaction terms.

For example, consider instanton five-branes in ten-dimensional super Yang-Mills. The worldvolume theory must preserve 8 of the 16 supercharges. The only such theory in $5+1$ dimensions is a sigma-model on a hyperKähler target space [24] which, for instantons, is the manifold $\mathcal{I}_{k,N}$. The Lagrangian is

$$\mathcal{L} = g_{\alpha\beta}\partial X^\alpha \partial X^\beta + i\bar{\chi}^\alpha D_{\alpha\beta}\chi^\beta + \tfrac{1}{4}R_{\alpha\beta\gamma\delta}\bar{\chi}^\alpha\chi^\beta\bar{\chi}^\gamma\chi^\delta \tag{1.30}$$

where ∂ denotes derivatives along the soliton worldvolume and the covariant derivative is $D_{\alpha\beta} = g_{\alpha\beta}\partial + \Gamma^\gamma_{\alpha\beta}(\partial X_\gamma)$. This is the slick proof that the instanton moduli space metric must be hyperKähler: it is dictated by the 8 preserved supercharges.

The final four-fermi term couples the fermionic collective coordinates to the Riemann tensor. Suppose we now want to go back down to instantons in four dimensional $\mathcal{N}=4$ super Yang-Mills. We can simply dimensionally reduce the above action. Since there are no longer worldvolume directions for the instantons, the first two terms vanish, but we're left with the term

$$S_{\rm inst} = \tfrac{1}{4}R_{\alpha\beta\gamma\delta}\bar{\chi}^\alpha\chi^\beta\bar{\chi}^\gamma\chi^\delta \tag{1.31}$$

This term reflects the point we made earlier: zero modes cannot necessarily be lifted to collective coordinates. Here we see this phenomenon for fermionic zero modes. Although each such mode doesn't change the action of the instanton, if we turn on four Grassmannian collective coordinates at the same time then the action does increase! One can derive this term without recourse to supersymmetry but it's a bit of a pain [25]. The term is very important in applications of instantons.

Instantons in four-dimensional $\mathcal{N}=2$ theories can be lifted to instanton strings in six dimensions. The worldvolume theory must preserve half of the 8 supercharges. There are two such super-algebras in two dimensions, a non-chiral $(2,2)$ theory and a chiral $(0,4)$ theory, where the two entries correspond to left and right moving fermions respectively. By analyzing the fermionic zero modes one can show that the instanton string preserves $(0,4)$ supersymmetry. The corresponding sigma-model doesn't contain the term (1.31). (Basically because the $\bar{\chi}$ zero modes are missing). However, similar terms can be generated if we also consider fermions in the fundamental representation.

Finally, instantons in $\mathcal{N}=1$ super Yang-Mills preserve $(0,2)$ supersymmetry on their worldvolume.

In the following sections, we shall pay scant attention to the fermionic zero modes, simply stating the fraction of supersymmetry that is preserved in different theories. In many cases this is sufficient to fix the fermions completely: the beauty of supersymmetry is that we rarely have to talk about fermions!

1.4 The ADHM Construction

In this section we describe a powerful method to solve the self-dual Yang-Mills equations $F = {}^\star F$ due to Atiyah, Drinfeld, Hitchin and Manin and known as the ADHM construction [26]. This will also give us a new way to understand the moduli space $\mathcal{I}_{k,N}$ and its metric. The natural place to view the ADHM construction is twistor space. But, for a physicist, the simplest place to view the ADHM construction is type II string theory [27, 28, 29]. We'll do things the simple way.

The brane construction is another place where it's useful to consider Yang-Mills instantons embedded as solitons in a $p+1$ dimensional theory with $p \geq 3$. With this in mind, let's consider a configuration of N Dp-branes, with k D$(p-4)$-branes in type II string theory (Type IIB for p odd; type IIA for p even). A typical configuration is drawn in figure 1. We place all N Dp-branes on top of each other so that, at low energies, their worldvolume dynamics is described by

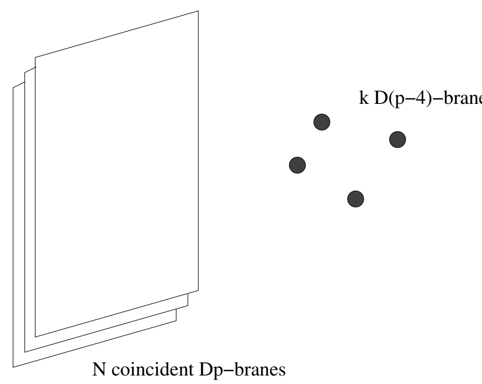

Figure 1: Dp-branes as instantons.

$$d = p + 1 \ U(N) \text{ Super Yang-Mills with 16 Supercharges}$$

For example, if $p = 3$ we have the familiar $\mathcal{N} = 4$ theory in $d = 3 + 1$ dimensions. The worldvolume theory of the Dp-branes also includes couplings to the various RR-fields in the bulk. This includes the term

$$\mathrm{Tr} \int_{Dp} d^{p+1}x \ \ C_{p-3} \wedge F \wedge F \tag{1.32}$$

where F is the $U(N)$ gauge field, and C_{p-3} is the RR-form that couples to D$(p-4)$-branes. The importance of this term lies in the fact that it relates instantons on the Dp-branes to D$(p-4)$ branes. To see this, note that an instanton with non-zero $F \wedge F$ gives rise to a source $(8\pi^2/e^2) \int d^{p-3}x \ C_{p-3}$ for the RR-form. This is the same source induced by a D$(p-4)$-brane. If you're careful in comparing the factors of 2 and π and such like, it's not hard to show that the instanton has precisely the mass and charge of the D$(p-4)$-brane [3, 5]. They are the same object! We have the important result that

$$\text{Instanton in D}p\text{-Brane} \;\equiv\; \text{D}(p-4)\text{-Brane} \tag{1.33}$$

The strategy to derive the ADHM construction from branes is to view this whole story from the perspective of the D$(p-4)$-branes [27, 28, 29]. For definiteness, let's revert back to $p=3$, so that we're considering D-instantons interacting with $D3$-branes. This means that we have to write down the $d=0+0$ dimensional theory on the D-instantons. Since supersymmetric theories in no dimensions may not be very familiar, it will help to keep in mind that the whole thing can be lifted to higher p.

Suppose firstly that we don't have the D3-branes. The theory on the D-instantons in flat space is simply the dimensional reduction of $d=3+1$ $\mathcal{N}=4$ $U(k)$ super Yang-Mills to zero dimensions. We will focus on the bosonic sector, with the fermions dictated by supersymmetry as explained in the previous section. We have 10 scalar fields, each of which is a $k\times k$ Hermitian matrix. For later convenience, we split them into two batches:

$$(X^\mu, \hat{X}^m) \qquad \mu=1,2,3,4; \quad m=5,\ldots,10 \tag{1.34}$$

where we've put hats on directions transverse to the D3-brane. We'll use the index notation $(X^\mu)^\alpha{}_\beta$ to denote the fact that each of these is a $k\times k$ matrix. Note that this is a slight abuse of notation since, in the previous section, $\alpha=1,\ldots,4k$ rather than $1,\ldots,k$ here. We'll also introduce the complex notation

$$Z = X_1 + iX_2 \quad , \quad W = X_3 - iX_4 \tag{1.35}$$

When X_μ and $\hat{X}_m$ are all mutually commuting, their $10k$ eigenvalues have the interpretation of the positions of the k D-instantons in flat ten-dimensional space.

What effect does the presence of the $D3$-branes have? The answer is well known. Firstly, they reduce the supersymmetry on the lower dimensional brane by half, to eight supercharges (equivalent to $\mathcal{N}=2$ in $d=3+1$). The decomposition (1.34) reflects this, with the $\hat{X}_m$ lying in a vector multiplet and the X_μ forming an adjoint hypermultiplet. The new fields which reduce the supersymmetry are N hypermultiplets, arising from quantizing strings stretched between the Dp-branes and D$(p-4)$-branes. Each hypermultiplet carries an $\alpha=1,\ldots k$ index, corresponding to the D$(p-4)$-brane on which the string ends, and an $a=1,\ldots,N$ index corresponding to the Dp-brane on which the other end of the string sits.. Again we ignore fermions. The two complex scalars in each hypermultiplet are denoted

$$\psi^\alpha{}_a \quad , \quad \tilde{\psi}^a{}_\alpha \tag{1.36}$$

where the index structure reflects the fact that ψ transforms in the $\mathbf{k}$ of the $U(k)$ gauge symmetry, and the $\bar{\mathbf{N}}$ of a $SU(N)$ flavor symmetry. In contrast $\tilde{\psi}$ transforms in the $(\bar{\mathbf{k}}, \mathbf{N})$ of $U(k) \times SU(N)$. (One may wonder about the difference between a gauge and flavor symmetry in zero dimensions; again the reader is invited to lift the configuration to higher dimensions where such nasty questions evaporate. But the basic point will be that we treat configurations related by $U(k)$ transformations as physically equivalent). These hypermultiplets can be thought of as the dimensional reduction of $\mathcal{N} = 2$ hypermultiplets in $d = 3 + 1$ dimensions which, in turn, are composed of two chiral multiplets ψ and $\tilde{\psi}$.

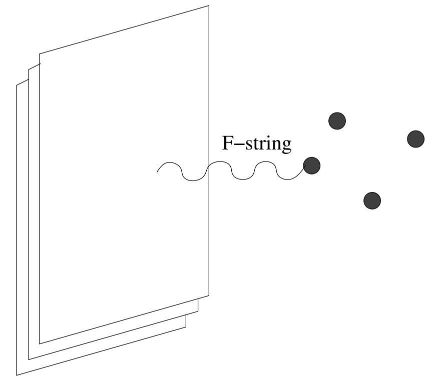

Figure 2: F-strings give rise to hypermultiplets.

The scalar potential for these fields is fixed by supersymmetry (Actually, supersymmetry in $d = 0 + 0$ dimensions is rather weak; at this stage we should lift up to, say $p = 7$, where so we can figure out the familiar $\mathcal{N} = 2$ theory on the D$(p-3)$=D3-branes, and then dimensionally reduce back down to zero dimensions). We have

$$V = \frac{1}{g^2} \sum_{m,n=5}^{10} [\hat{X}_m, \hat{X}_n]^2 + \sum_{m=5}^{10} \sum_{\mu=1}^{4} [\hat{X}_m, X_\mu]^2 + \sum_{a=1}^{N} (\psi^{a\dagger} \hat{X}_m^2 \psi_a + \tilde{\psi}^a \hat{X}_m^2 \tilde{\psi}_a^\dagger) \tag{1.37}$$

$$+ g^2 \,\mathrm{Tr}\, (\sum_{a=1}^{N} \psi_a \psi^{a\dagger} - \tilde{\psi}_a^\dagger \tilde{\psi}^a + [Z, Z^\dagger] + [W, W^\dagger])^2 + g^2 \,\mathrm{Tr}\, |\sum_{a=1}^{N} \psi_a \tilde{\psi}^a + [Z, W]|^2$$

The terms in the second line are usually referred to as D-terms and F-terms respectively (although, as we shall shall review shortly, they are actually on the same footing in theories with eight supercharges). Each is a $k \times k$ matrix. The third term in the first line ensures that the hypermultiplets get a mass if the $\hat{X}_m$ get a vacuum expectation value. This reflects the fact that, as is clear from the picture, the Dp-D$(p-4)$ strings become stretched if the branes are separated in the $\hat{X}^m$, $m = 5, \ldots, 10$ directions. In contrast, there is no mass for the hypermultiplets if the D$(p-4)$ branes are separated in the X_μ, $\mu = 1, 2, 3, 4$ directions. Finally, note that we've included an auxiliary coupling constant g^2 in (1.37). Strictly speaking we should take the limit $g^2 \to \infty$.

We are interested in the ground states of the D-instantons, determined by the solutions to $V = 0$. There are two possibilities

1. The second line vanishes if $\psi = \tilde{\psi} = 0$ and X_μ are diagonal. The first two terms vanish if $\hat{X}_m$ are also diagonal. The eigenvalues of X_μ and $\hat{X}_m$ tell us

where the k D-instantons are placed in flat space. They are unaffected by the existence of the D3-branes whose presence is only felt at the one-loop level when the hypermultiplets are integrated out. This is known as the "*Coulomb branch*", a name inherited from the structure of gauge symmetry breaking: $U(k) \rightarrow U(1)^k$. (The name is, of course, more appropriate in dimensions higher than zero where particles charged under $U(1)^k$ experience a Coulomb interaction).

2. The first line vanishes if $\hat{X}_m = 0$, $m = 5, \ldots, 10$. This corresponds to the D($p-4$) branes lying on top of the Dp-branes. The remaining fields ψ, $\tilde{\psi}$, Z and W are constrained by the second line in (1.37). Since these solutions allow $\psi, \tilde{\psi} \neq 0$ we will generically have the $U(k)$ gauge group broken completely, giving the name "*Higgs branch*" to this class of solutions. More precisely, the Higgs branch is defined to be the space of solutions

$$\mathcal{M}_{\rm Higgs} \cong \{\hat{X}_m = 0, V = 0\}/U(k) \tag{1.38}$$

where we divide out by $U(k)$ gauge transformations. The Higgs branch describes the D($p-4$) branes nestling inside the larger Dp-branes. But this is exactly where they appear as instantons. So we might expect that the Higgs branch knows something about this. Let's start by computing its dimension. We have $4kN$ real degrees of freedom in ψ and $\tilde{\psi}$ and a further $4k^2$ in Z and W. The D-term imposes k^2 real constraints, while the F-term imposes k^2 complex constraints. Finally we lose a further k^2 degrees of freedom when dividing by $U(k)$ gauge transformations. Adding, subtracting, we have

$$\dim(\mathcal{M}_{\rm Higgs}) = 4kN \tag{1.39}$$

which should look familiar (1.13). The first claim of the ADHM construction is that we have an isomorphism between manifolds,

$$\mathcal{M}_{\rm Higgs} \cong \mathcal{I}_{k,N} \tag{1.40}$$

1.4.1 The Metric on the Higgs Branch

To summarize, the D-brane construction has lead us to identify the instanton moduli space $\mathcal{I}_{k,N}$ with the Higgs branch of a gauge theory with 8 supercharges (equivalent to $\mathcal{N} = 2$ in $d = 3+1$). The field content of this gauge theory is

$$\begin{aligned} U(k) \text{ Gauge Theory} &+ \text{Adjoint Hypermultiplet } Z, W \\ &+ N \text{ Fundamental Hypermultiplets } \psi_a, \tilde{\psi}^a \end{aligned} \tag{1.41}$$

This auxiliary $U(k)$ gauge theory defines its own metric on the Higgs branch. This metric arises in the following manner: we start with the flat metric on $\mathbf{R}^{4k(N+k)}$, parameterized by ψ, $\tilde{\psi}$, Z and W. Schematically,

$$ds^2 = |d\psi|^2 + |d\tilde{\psi}|^2 + |dZ|^2 + |dW|^2 \tag{1.42}$$

This metric looks somewhat more natural if we consider higher dimensional D-branes where it arises from the canonical kinetic terms for the hypermultiplets. We now pull back this metric to the hypersurface $V = 0$, and subsequently quotient by the $U(k)$ gauge symmetry, meaning that we only consider tangent vectors to $V = 0$ that are orthogonal to the $U(k)$ action. This procedure defines a metric on $\mathcal{M}_{\rm Higgs}$. The second important result of the ADHM construction is that this metric coincides with the one defined in terms of solitons in (1.18).

I haven't included a proof of the equivalence between the metrics here, although it's not too hard to show (for example, using Maciocia's hyperKähler potential [22] as reviewed in [13]). However, we will take time to show that the isometries of the metrics defined in these two different ways coincide. From the perspective of the auxiliary $U(k)$ gauge theory, all isometries appear as flavor symmetries. We have the $SU(N)$ flavor symmetry rotating the hypermultiplets; this is identified with the $SU(N)$ gauge symmetry in four dimensions. The theory also contains an $SU(2)_R$ R-symmetry, in which $(\psi, \tilde{\psi}^\dagger)$ and $(Z, W^\dagger)$ both transform as doublets (this will become more apparent in the following section in equation (1.44)). This coincides with the $SU(2)_R \subset SO(4)$ rotational symmetry in four dimensions. Finally, there exists an independent $SU(2)_L$ symmetry rotating just the X_μ.

The method described above for constructing hyperKähler metrics is an example of a technique known as the hyperKähler quotient [30]. As we have seen, it arises naturally in gauge theories with 8 supercharges. The D- and F-terms of the potential (1.37) give what are called the triplet of "moment-maps" for the $U(k)$ action.

1.4.2 Constructing the Solutions

As presented so far, the ADHM construction relates the moduli space of instantons $\mathcal{I}_{k,N}$ to the Higgs branch of an auxiliary gauge theory. In fact, we've omitted the most impressive part of the story: the construction can also be used to give solutions to the self-duality equations. What's more, it's really very easy! Just a question of multiplying a few matrices together. Let's see how it works.

Firstly, we need to rewrite the vacuum conditions in a more symmetric fashion. Define

$$\omega_a = \begin{pmatrix} \psi^{\alpha}{}_{a} \\ \tilde{\psi}^{\dagger\alpha}{}_{a} \end{pmatrix} \tag{1.43}$$

Then the real D-term and complex F-term which lie in the second line of (1.37) and define the Higgs branch can be combined in to the triplet of constraints,

$$\sum_{a=1}^{N} \omega_a^\dagger \sigma^i \, \omega_a - i[X_\mu, X_\nu]\bar{\eta}^i_{\mu\nu} = 0 \tag{1.44}$$

where σ^i are, as usual, the Pauli matrices and $\bar{\eta}^i$ the 't Hooft matrices (1.10). These give three $k \times k$ matrix equations. The magic of the ADHM construction is that for each solution to the algebraic equations (1.44), we can build a solution to the set of non-linear partial differential equations $F = {}^\star F$. Moreover, solutions to (1.44) related by $U(k)$ gauge transformations give rise to the same field configuration in four dimensions. Let's see how this remarkable result is achieved.

The first step is to build the $(N + 2k) \times 2k$ matrix Δ,

$$\Delta = \begin{pmatrix} \omega^T \\ X_\mu \sigma^\mu \end{pmatrix} + \begin{pmatrix} 0 \\ x_\mu \sigma^\mu \end{pmatrix} \tag{1.45}$$

where $\sigma_\mu = (\sigma^i, -i\mathbf{1}_2)$. These have the important property that $\sigma_{[\mu}\bar{\sigma}_{\nu]}$ is self-dual, while $\bar{\sigma}_{[\mu}\sigma_{\nu]}$ is anti-self-dual, facts that we also used in Section 1.3 when discussing fermions. In the second matrix we've re-introduced the spacetime coordinate x_μ which, here, is to be thought of as multiplying the $k \times k$ unit matrix. Before proceeding, we need a quick lemma:

Lemma: $\Delta^\dagger \Delta = f^{-1} \otimes 1_2$

where f is a $k \times k$ matrix, and 1_2 is the unit 2×2 matrix. In other words, $\Delta^\dagger\Delta$ factorizes and is invertible.

Proof: Expanding out, we have (suppressing the various indices)

$$\Delta^\dagger \Delta = \omega^\dagger \omega + X^\dagger X + (X^\dagger x + x^\dagger X) + x^\dagger x 1_k \tag{1.46}$$

Since the factorization happens for all $x \equiv x_\mu \sigma^\mu$, we can look at three terms separately. The last is $x^\dagger x = x_\mu \bar{\sigma}^\mu x_\nu \sigma^\nu = x^2 \, 1_2$. So that works. For the term linear in x, we simply need the fact that $X_\mu = X_\mu^\dagger$ to see that it works. What's more tricky is the term that doesn't depend on x. This is where the triplet of D-terms (1.44) comes in. Let's write the relevant term from (1.46) with all the indices, including an $m, n = 1, 2$ index to denote the two components we introduced in (1.43). We require

$$\begin{aligned} & \omega^{\dagger\alpha}_{ma}\,\omega_{a\beta n} + (X_\mu)^\alpha{}_\gamma (X_\nu)^\gamma{}_\beta\, \bar{\sigma}^{\mu\, mp} \sigma^\nu{}_{pn} \sim \delta^m_n \\ \Leftrightarrow \quad & \mathrm{tr}_2\, \sigma^i \left[\omega\omega^\dagger + X^\dagger X\right] = 0 \quad\quad i = 1, 2, 3 \\ \Leftrightarrow \quad & \omega^\dagger \sigma^i \omega + X_\mu X_\nu \bar{\sigma}^\mu \sigma^i \sigma^\nu = 0 \end{aligned} \tag{1.47}$$

But, using the identity $\bar{\sigma}^\mu \sigma^i \sigma^\nu = 2i\bar{\eta}^i_{\mu\nu}$, we see that this last condition is implied by the vanishing of the D-terms (1.44). This concludes our proof of the lemma. □

The rest is now plain sailing. Consider the matrix Δ as defining $2k$ linearly independent vectors in $\mathbf{C}^{N+2k}$. We define U to be the $(N + 2k) \times N$ matrix containing the N normalized, orthogonal vectors. i.e

$$\Delta^\dagger U = 0 \quad\quad , \quad\quad U^\dagger U = 1_N \tag{1.48}$$

Then the potential for a charge k instanton in $SU(N)$ gauge theory is given by

$$A_\mu = iU^\dagger\, \partial_\mu U \tag{1.49}$$

Note firstly that if U were an $N \times N$ matrix, this would be pure gauge. But it's not, and it's not. Note also that A_μ is left unchanged by auxiliary $U(k)$ gauge transformations.

We need to show that A_μ so defined gives rise to a self-dual field strength with winding number k. We'll do the former, but the latter isn't hard either: it just requires more matrix multiplication. To help us in this, it will be useful to construct the projection operator $P = UU^\dagger$ and notice that this can also be written as $P = 1 - \Delta f \Delta^\dagger$. To see that these expressions indeed coincide, we can check that $PU = U$ and $P\Delta = 0$ for both. Now we're almost there:

$$\begin{aligned} F_{\mu\nu} &= \partial_{[\mu} A_{\nu]} - iA_{[\mu} A_{\nu]} \\ &= \partial_{[\mu}\, iU^\dagger \partial_{\nu]} U + iU^\dagger (\partial_{[\mu} U) U^\dagger (\partial_{\nu]} U) \\ &= i(\partial_{[\mu} U^\dagger)(\partial_{\nu]} U) - i(\partial_{[\mu} U^\dagger) U U^\dagger (\partial_{\nu]} U) \\ &= i(\partial_{[\mu} U^\dagger)(1 - UU^\dagger)(\partial_{\nu]} U) \\ &= i(\partial_{[\mu} U^\dagger) \Delta\, f\, \Delta^\dagger (\partial_{\nu]} U) \\ &= iU^\dagger (\partial_{[\mu} \Delta)\, f\, (\partial_{\nu]} U) \\ &= iU^\dagger \sigma_{[\mu} f \bar{\sigma}_{\nu]} U \end{aligned}$$

At this point we use our lemma. Because $\Delta^\dagger\Delta$ factorizes, we may commute f past σ_μ. And that's it! We can then write

$$F_{\mu\nu} = iU^\dagger f\sigma_{[\mu}\bar{\sigma}_{\nu]}U = {}^\star F_{\mu\nu} \tag{1.50}$$

since, as we mentioned above, $\sigma_{\mu\nu} = \sigma_{[\mu}\bar{\sigma}_{\nu]}$ is self-dual. Nice huh! What's harder to show is that the ADHM construction gives all solutions to the self-dualily equations. Counting parameters, we see that we have the right number and it turns out that we can indeed get all solutions in this manner.

The construction described above was first described in ADHM's original paper, which weighs in at a whopping 2 pages. Elaborations and extensions to include, among other things, $SO(N)$ and $Sp(N)$ gauge groups, fermionic zero modes, supersymmetry and constrained instantons, can be found in [31, 32, 33, 34].

An Example: The Single $SU(2)$ Instanton Revisited

Let's see how to re-derive the $k = 1$ $SU(2)$ solution (1.9) from the ADHM method. We'll set $X_\mu = 0$ to get a solution centered around the origin. We then have the 4×2 matrix

$$\Delta = \begin{pmatrix} \omega^T \\ x_\mu\sigma^\mu \end{pmatrix} \tag{1.51}$$

where the D-term constraints (1.44) tell us that $\omega^\dagger_{am}(\sigma^i)^m{}_n\omega^n_a = 0$. We can use our $SU(2)$ flavor rotation, acting on the indices $a, b = 1, 2$, to choose the solution

$$\omega^{\dagger a}{}_m\omega^m{}_b = \rho^2\delta^a{}_b \tag{1.52}$$

in which case the matrix Δ becomes $\Delta^T = (\rho 1_2, x_\mu\sigma^\mu)$. Then solving for the normalized zero eigenvectors $\Delta^\dagger U = 0$, and $U^\dagger U = 1$, we have

$$U = \begin{pmatrix} \sqrt{x^2/(x^2+\rho^2)}\, 1_2 \\ -\sqrt{\rho^2/x^2(x^2+\rho^2)}\, x_\mu\bar{\sigma}^\mu \end{pmatrix} \tag{1.53}$$

From which we calculate

$$A_\mu = iU^\dagger\partial_\mu U = \frac{\rho^2 x_\nu}{x^2(x^2+\rho^2)}\,\bar{\eta}^i_{\mu\nu}\,\sigma^i \tag{1.54}$$

which is indeed the solution (1.9) as promised.

1.4.3 Non-Commutative Instantons

There's an interesting deformation of the ADHM construction arising from studying instantons on a non-commutative space, defined by

$$[x_\mu, x_\nu] = i\theta_{\mu\nu} \tag{1.55}$$

The most simple realization of this deformation arises by considering functions on the space $\mathbf{R}^4_\theta$, with multiplication given by the $\star$-product

$$f(x) \star g(x) = \exp\left(\frac{i}{2}\theta_{\mu\nu}\frac{\partial}{\partial y^\mu}\frac{\partial}{\partial x^\nu}\right) f(y)g(x)\bigg|_{x=y} \tag{1.56}$$

so that we indeed recover the commutator $x_\mu \star x_\nu - x_\nu \star x_\mu = i\theta_{\mu\nu}$. To define gauge theories on such a non-commutative space, one must extend the gauge symmetry from $SU(N)$ to $U(N)$. When studying instantons, it is also useful to decompose the non-commutivity parameter into self-dual and anti-self-dual pieces:

$$\theta_{\mu\nu} = \xi^i\,\eta^i_{\mu\nu} + \zeta^i\,\bar{\eta}^i_{\mu\nu} \tag{1.57}$$

where η^i and $\bar{\eta}^i$ are defined in (1.11) and (1.10) respectively. At the level of solutions, both ξ and ζ affect the configuration. However, at the level of the moduli space, we shall see that the self-dual instantons $F = {}^\star F$ are only affected by the anti-self-dual part of the non-commutivity, namely ζ^i. (A similar statement holds for $F = -{}^\star F$ solutions and ξ). This change to the moduli space appears in a beautifully simple fashion in the ADHM construction: we need only add a constant term to the right hand-side of the constraints (1.44), which now read

$$\sum_{a=1}^{N} \omega_a^\dagger \sigma^i\,\omega_a - i[X_\mu, X_\nu]\bar{\eta}^i_{\mu\nu} = \zeta^i\,1_k \tag{1.58}$$

From the perspective of the auxiliary $U(k)$ gauge theory, the ζ^i are Fayet-Iliopoulous (FI) parameters.

The observation that the FI parameters ζ^i appearing in the D-term give the correct deformation for non-commutative instantons is due to Nekrasov and Schwarz [35]. To see how this works, we can repeat the calculation above, now in non-commutative space. The key point in constructing the solutions is once again the requirement that we have the factorization

$$\Delta^\dagger \star \Delta = f^{-1}\,1_2 \tag{1.59}$$

The one small difference from the previous derivation is that in the expansion (1.46), the $\star$-product means we have

$$x^\dagger \star x = x^2\, 1_2 - \zeta^i \sigma^i \tag{1.60}$$

Notice that only the anti-self-dual part contributes. This extra term combines with the constant terms (1.47) to give the necessary factorization if the D-term with FI parameters (1.58) is satisfied. It is simple to check that the rest of the derivation proceeds as before, with $\star$-products in the place of the usual commutative multiplication.

The addition of the FI parameters in (1.58) have an important effect on the moduli space $\mathcal{I}_{k,N}$: they resolve the small instanton singularities. From the ADHM perspective, these arise when $\psi = \tilde{\psi} = 0$, where the $U(k)$ gauge symmetry does not act freely. The FI parameters remove these points from the moduli space, $U(k)$ acts freely everywhere on the Higgs branch, and the deformed instanton moduli space $\mathcal{I}_{k,N}$ is smooth. This resolution of the instanton moduli space was considered by Nakajima some years before the relationship to non-commutivity was known [36]. A related fact is that non-commutative instantons occur even for $U(1)$ gauge theories. Previously such solutions were always singular, but the addition of the FI parameter stabilizes them at a fixed size of order $\sqrt{\theta}$. Reviews of instantons and other solitons on non-commutative spaces can be found in [37, 38].

1.4.4 Examples of Instanton Moduli Spaces

A Single Instanton

Consider a single $k = 1$ instanton in a $U(N)$ gauge theory, with non-commutivity turned on. Let us choose $\theta_{\mu\nu} = \zeta \bar{\eta}^3_{\mu\nu}$. Then the ADHM gauge theory consists of a $U(1)$ gauge theory with N charged hypermultiplets, and a decoupled neutral hypermultiplet parameterizing the center of the instanton. The D-term constraints read

$$\sum_{a=1}^N |\psi_a|^2 - |\tilde{\psi}_a|^2 = \zeta \quad , \quad \sum_{a=1}^N \tilde{\psi}_a \psi_a = 0 \tag{1.61}$$

To get the moduli space we must also divide out by the $U(1)$ action $\psi_a \to e^{i\alpha}\psi_a$ and $\tilde{\psi}_a \to e^{-i\alpha}\tilde{\psi}_a$. To see what the resulting space is, first consider setting $\tilde{\psi}_a = 0$. Then we have the space

$$\sum_{a=1}^N |\psi_a|^2 = \zeta \tag{1.62}$$

which is simply $\mathbf{S}^{2N-1}$. Dividing out by the $U(1)$ action then gives us the complex projective space $\mathbb{CP}^{N-1}$ with size (or Kähler class) ζ. Now let's add the $\tilde{\psi}$ back. We can turn them on but the F-term insists that they lie orthogonal to ψ, thus defining the co-tangent bundle of $\mathbb{CP}^{N-1}$, denoted $T^\star\mathbb{CP}^{N-1}$. Including the decoupled $\mathbf{R}^4$, we have [39]

$$\mathcal{I}_{1,N} \cong \mathbf{R}^4 \times T^\star\mathbb{CP}^{N-1} \tag{1.63}$$

where the size of the zero section $\mathbb{CP}^{N-1}$ is ζ. As $\zeta \to 0$, this cycle lying in the center of the space shrinks and $\mathcal{I}_{1,N}$ becomes singular at this point.

For a single instanton in $U(2)$, the relative moduli space is $T^\star\mathbf{S}^2$. This is the smooth resolution of the A_1 singularity $\mathbf{C}^2/\mathbf{Z}_2$ which we found to be the moduli space in the absence of non-commutivity. It inherits a well-known hyperKähler metric known as the Eguchi-Hanson metric [40],

$$ds^2_{EH} = \left(1 - 4\zeta^2/\rho^4\right)^{-1} d\rho^2 + \frac{\rho^2}{4}\left(\sigma_1^2 + \sigma_2^2 + \left(1 - 4\zeta^2/\rho^4\right)\sigma_3^2\right) \tag{1.64}$$

Here the σ_i are the three left-invariant $SU(2)$ one-forms which, in terms of polar angles $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$ and $0 \leq \psi \leq 2\pi$, take the form

$$\begin{aligned} \sigma_1 &= -\sin\psi\ d\theta + \cos\psi\sin\theta\ d\phi \\ \sigma_2 &= \cos\psi\ d\theta + \sin\psi\sin\theta\ d\phi \\ \sigma_3 &= d\psi + \cos\theta\ d\phi \end{aligned} \tag{1.65}$$

As $\rho \to \infty$, this metric tends towards the cone over $\mathbf{S}^3/\mathbf{Z}_2$. However, as we approach the origin, the scale size is truncated at $\rho^2 = 2\zeta$, where the apparent singularity is merely due to the choice of coordinates and hides the zero section $\mathbf{S}^2$.

Two $U(1)$ Instantons

Before resolving by a non-commutative deformation, there is no topology to support a $U(1)$ instanton. However, it is perhaps better to think of the $U(1)$ theory as admitting small, singular, instantons with moduli space given by the symmetric product $\text{Sym}^k(\mathbf{C}^2)$, describing the positions of k points. Upon the addition of a non-commutivity parameter, smooth $U(1)$ instantons exist with moduli space given by a resolution of $\text{Sym}^k(\mathbf{C}^2)$. To my knowledge, no explicit metric is known for $k \geq 3$ $U(1)$ instantons, but in the case of two $U(1)$ instantons, the metric is something rather familiar, since $\text{Sym}^2\mathbf{C}^2 \cong \mathbf{C}^2 \times \mathbf{C}^2/Z_2$ and we have already met the resolution of this space above. It is

$$\mathcal{I}_{k=2,N=1} \cong \mathbf{R}^4 \times T^\star\mathbf{S}^2 \tag{1.66}$$

endowed with the Eguchi-Hanson metric (1.64) where ρ now has the interpretation of the separation of two instantons rather than the scale size of one. This can be checked explicitly by computing the metric on the ADHM Higgs branch using the hyperKähler quotient technique [41]. Scattering of these instantons was studied in [42]. So, in this particular case we have $\mathcal{I}_{1,2} \cong \mathcal{I}_{2,1}$. We shouldn't get carried away though as this equivalence doesn't hold for higher k and N (for example, the isometries of the two spaces are different).

1.5 Applications

Until now we've focussed exclusively on classical aspects of the instanton configurations. But, what we're really interested in is the role they play in various quantum field theories. Here we sketch two examples which reveal the importance of instantons in different dimensions.

1.5.1 Instantons and the AdS/CFT Correspondence

We start by considering instantons where they were meant to be: in four dimensional gauge theories. In a semi-classical regime, instantons give rise to non-perturbative contributions to correlation functions and there exists a host of results in the literature, including exact results in both $\mathcal{N} = 1$ [43, 44] and $\mathcal{N} = 2$ [45, 34, 37] supersymmetric gauge theories. Here we describe the role instantons play in $\mathcal{N} = 4$ super Yang-Mills and, in particular, their relationship to the AdS/CFT correspondence [47]. Instantons were first considered in this context in [48, 49]. Below we provide only a sketchy description of the material covered in the paper of Dorey et al [50]. Full details can be found in that paper or in the review [13].

In any instanton computation, there's a number of things we need to calculate [7]. The first is to count the zero modes of the instanton to determine both the bosonic collective coordinates X and their fermionic counterparts χ. We've described this in detail above. The next step is to perform the leading order Gaussian integral over all modes in the path integral. The massive (i.e. non-zero) modes around the background of the instanton leads to the usual determinant operators which we'll denote as $\det \Delta_B$ for the bosons, and $\det \Delta_F$ for the fermions. These are to be evaluated on the background of the instanton solution. However, zero modes must be treated separately. The integration over the associated collective coordinates is left unperformed, at the price of introducing a Jacobian arising from the transformation between field variables and collective coordinates. For the bosonic fields, the Jacobian is simply $J_B = \sqrt{\det g_{\alpha\beta}}$, where $g_{\alpha\beta}$ is the metric on the instanton moduli space defined in (1.18). This is the role played by the instanton moduli space metric in four dimensions: it appears in the measure when performing the path integral. A related factor J_F occurs for fermionic zero modes. The final ingredient in an instanton calculation is the action $S_{\rm inst}$ which includes both the constant piece $8\pi k/g^2$, together with terms quartic in the fermions (1.31). The end result is summarized in the instanton measure

$$d\mu_{\rm inst} = d^{n_B}X\, d^{n_F}\chi\; J_B J_F\; \frac{\det\Delta_F}{\det^{1/2}\Delta_B}\, e^{-S_{\rm inst}} \tag{1.67}$$

where there are $n_B = 4kN$ bosonic and n_F fermionic collective coordinates. In supersymmetric theories in four dimensions, the determinants famously cancel [7] and we're left only with the challenge of evaluating the Jacobians and the action. In this section, we'll sketch how to calculate these objects for $\mathcal{N}=4$ super Yang-Mills.

As is well known, in the limit of strong 't Hooft coupling, $\mathcal{N}=4$ super Yang-Mills is dual to type IIB supergravity on $AdS_5\times\mathbf{S}^5$. An astonishing fact, which we shall now show, is that we can see this geometry even at weak 't Hooft coupling by studying the $d=0+0$ ADHM gauge theory describing instantons. Essentially, in the large N limit, the instantons live in $AdS_5\times\mathbf{S}^5$. At first glance this looks rather unlikely! We've seen that if the instantons live anywhere it is in $\mathcal{I}_{k,N}$, a $4kN$ dimensional space that doesn't look anything like $AdS_5\times\mathbf{S}^5$. So how does it work?

While the calculation can be performed for an arbitrary number of k instantons, here we'll just stick with a single instanton as a probe of the geometry. To see the AdS_5 part is pretty easy and, in fact, we can do it even for an instanton in $SU(2)$ gauge theory. The trick is to integrate over the orientation modes of the instanton, leaving us with a five-dimensional space parameterized by X_μ and ρ. The rationale for doing this is that if we want to compute gauge invariant correlation functions, the $SU(N)$ orientation modes will only give an overall normalization. We calculated the metric for a single instanton in equations (1.22)-(1.24), giving us $J_B\sim\rho^3$ (where we've dropped some numerical factors and factors of e^2). So integrating over the $SU(2)$ orientation to pick up an overall volume factor, we get the bosonic measure for the instanton to be

$$d\mu_{\rm inst}\sim\rho^3\,d^4X d\rho \tag{1.68}$$

We want to interpret this measure as a five-dimensional space in which the instanton moves, which means thinking of it in the form $d\mu=\sqrt{G}\,d^4X\,d\rho$ where G is the metric on the five-dimensional space. It would be nice if it was the metric on AdS_5. But it's not! In the appropriate coordinates, the AdS_5 metric is,

$$ds^2_{AdS} = \frac{R^2}{\rho^2}(d^4X + d\rho^2) \tag{1.69}$$

giving rise to a measure $d\mu_{AdS} = (R/\rho)^5 d^4X d\rho$. However, we haven't finished with the instanton yet since we still have to consider the fermionic zero modes. The fermions are crucial for quantum conformal invariance so we may suspect that their zero modes are equally crucial in revealing the AdS structure, and this is indeed the case. A single $k = 1$ instanton in the $\mathcal{N} = 4$ $SU(2)$ gauge theory has 16 fermionic zero modes. 8 of these, which we'll denote as ξ are from broken supersymmetry while the remaining 8, which we'll call ζ arise from broken superconformal transformations. Explicitly each of the four Weyl fermions λ of the theory has a profile,

$$\lambda = \sigma^{\mu\nu} F_{\mu\nu}(\xi - \sigma^\rho \zeta \, (x_\rho - X_\rho)) \tag{1.70}$$

One can compute the overlap of these fermionic zero modes in the same way as we did for bosons. Suppressing indices, we have

$$\int d^4x \, \frac{\partial\lambda}{\partial\xi}\frac{\partial\lambda}{\partial\xi} = \frac{32\pi^2}{e^2} \quad , \quad \int d^4x \, \frac{\partial\lambda}{\partial\zeta}\frac{\partial\lambda}{\partial\zeta} = \frac{64\pi^2\rho^2}{e^2} \tag{1.71}$$

So, recalling that Grassmannian integration is more like differentiation, the fermionic Jacobian is $J_F \sim 1/\rho^8$. Combining this with the bosonic contribution above, the final instanton measure is

$$d\mu_{\rm inst} = \left(\frac{1}{\rho^5}\, d^4X d\rho\right) d^8\xi d^8\zeta = d\mu_{Ads}\, d^8\xi d^8\zeta \tag{1.72}$$

So the bosonic part does now look like AdS_5. The presence of the 16 Grassmannian variables reflects the fact that the instanton only contributes to a 16 fermion correlation function. The counterpart in the AdS/CFT correspondence is that D-instantons contribute to R^4 terms and their 16 fermion superpartners and one can match the supergravity and gauge theory correlators exactly.

So we see how to get AdS_5 for $SU(2)$ gauge theory. For $SU(N)$, one has $4N - 8$ further orientation modes and $8N - 16$ further fermi zero modes. The factors of ρ cancel in their Jacobians, leaving the AdS_5 interpretation intact. But there's a problem with these extra fermionic zero modes since we must saturate them in the path integral in some way even though we still want to compute a 16 fermionic correlator. This is achieved by the four-fermi term in the instanton action (1.31). However, when looked at in the right way, in the large N limit these extra fermionic zero modes will generate the $\mathbf{S}^5$ for us. I'll now sketch how this occurs.

The important step in reforming these fermionic zero modes is to introduce auxiliary variables $\hat{X}$ which allows us to split up the four-fermi term (1.31) into terms quadratic in the fermions. To get the index structure right, it turns out that we need six such auxiliary fields, let's call them $\hat{X}^m$, with $m = 1, \ldots, 6$. In fact we've met these guys before: they're the scalar fields in the vector multiplet of the ADHM gauge theory. To see that they give rise to the promised four fermi term, let's look at how they appear in the ADHM Lagrangian. There's already a term quadratic in $\hat{X}$ in (1.37), and another couples this to the surplus fermionic collective coordinates χ so that, schematically,

$$\mathcal{L}_{\hat{X}} \sim \hat{X}^2 \omega^\dagger \omega + \bar{\chi} \hat{X} \chi \tag{1.73}$$

where, as we saw in Section 1.4, the field ω contains the scale and orientation collective coordinates, with $\omega^\dagger \omega \sim \rho^2$. Integrating out $\hat{X}$ in the ADHM Lagrangian does indeed result in a four-fermi term which is identified with (1.31). However, now we perform a famous trick: we integrate out the variables we thought we were interested in, namely the χ fields, and focus on the ones we thought were unimportant, the $\hat{X}$'s. After dealing correctly with all the indices we've been dropping, we find that this results in the contribution to the measure

$$d\mu_{\text{auxiliary}} = d^6\hat{X}\ (\hat{X}^m \hat{X}^m)^{2N-4} \exp\left(-2\rho^2 \hat{X}^m \hat{X}^m\right) \tag{1.74}$$

In the large N limit, the integration over the radial variable $|\hat{X}|$ may be performed using the saddle-point approximation evaluated at $|\hat{X}| = \rho$. The resulting powers of ρ are precisely those mentioned above that are needed to cancel the powers of ρ appearing in the bosonic Jacobian. Meanwhile, the integration over the angular coordinates in $\hat{X}^m$ has been left untouched. The final result for the instanton measure becomes

$$d\mu_{\text{inst}} = \left(\frac{1}{\rho^5}\, d^4X\, d\rho\, d^5\hat{\Omega}\right)\ d^8\xi d^8\zeta \tag{1.75}$$

And the instanton indeed appears as if it's moving in $AdS_5 \times \mathbf{S}^5$ as promised.

The above discussion is a little glib. The invariant meaning of the measure alone is not clear: the real meaning is that when integrated against correlators, it gives results in agreement with gravity calculations in $\text{AdS}_5 \times \mathbf{S}^5$. This, and several further results, were shown in [50]. Calculations of this type were later performed for instantons in other four-dimensional gauge theories, both conformal and otherwise [51, 52, 53, 54, 55]. Curiously, there appears to be an unresolved problem with performing the calculation for instantons in non-commutative gauge theories.

1.5.2 Instanton Particles and the $(2,0)$ Theory

There exists a rather special superconformal quantum field theory in six dimensions known as the $(2,0)$ theory. It is the theory with 16 supercharges which lives on N M5-branes in M-theory and it has some intriguing and poorly understood properties. Not least of these is the fact that it appears to have N^3 degrees of freedom. While it's not clear what these degrees of freedom are, or even if it makes sense to talk about "degrees of freedom" in a strongly coupled theory, the N^3 behavior is seen when computing the free energy $F \sim N^3 T^6$ [56], or anomalies whose leading coefficient also scales as N^3 [57].

If the $(2,0)$ theory is compactified on a circle of radius R, it descends to $U(N)$ $d = 4+1$ super Yang-Mills with 16 supercharges, which can be thought of as living on D4-branes in Type IIA string theory. The gauge coupling e^2, which has dimension of length in five dimensions, is given by

$$e^2 = 8\pi^2 R \tag{1.76}$$

As in any theory compactified on a spatial circle, we expect to find Kaluza-Klein modes, corresponding to momentum modes around the circle with mass $M_{\rm KK} = 1/R$. Comparison with the gauge coupling constant (1.76) gives a strong hint what these particles should be, since

$$M_{\rm kk} = M_{\rm inst} \tag{1.77}$$

and, as we discussed in section 1.3.1, instantons are particle-like objects in $d = 4+1$ dimensions. The observation that instantons are Kaluza-Klein modes is clear from the IIA perspective: the instantons in the D4-brane theory are D0-branes which are known to be the Kaluza-Klein modes for the lift to M-theory.

The upshot of this analysis is a remarkable conjecture: the maximally supersymmetric $U(N)$ Yang-Mills theory in five dimensions is really a six-dimensional theory in disguise, with the size of the hidden dimension given by $R \sim e^2$ [58, 59, 60]. As $e^2 \to \infty$, the instantons become light. Usually, as solitons become light, they also become large floppy objects, losing their interpretation as particle excitations of the theory. But this isn't necessarily true for instantons because, as we've seen, their scale size is arbitrary and, in particular, independent of the gauge coupling.

Of course, the five-dimensional theory is non-renormalizable and we can only study questions that do not require the introduction of new UV degrees of freedom. With this caveat, let's see how we can test the conjecture using instantons. If they're really Kaluza-Klein modes, they should exhibit Kaluza-Klein-eqsue behavior which includes a characteristic spectrum of threshold bound state of particles with k units of momentum going around the circle. This means that if the five-dimensional theory contains the information about its six dimensional origin, it should exhibit a threshold bound state of k instantons for each k. But this is something we can test in the semi-classical regime by solving the low-energy dynamics of k interacting instantons. As we have seen, this is given by supersymmetric quantum mechanics on $\mathcal{I}_{k,N}$, with the Lagrangian given by (1.30) where $\partial = \partial_t$ in this equation.

Let's review how to solve the ground states of $d = 0+1$ dimensional supersymmetric sigma models of the form (1.30). As explained by Witten, a beautiful connection to de Rahm cohomology emerges after quantization [61]. Canonical quantization of the fermions leads to operators satisfying the algebra

$$\{\chi_\alpha, \chi_\beta\} = \{\bar{\chi}_\alpha, \bar{\chi}_\beta\} = 0 \quad \text{and} \quad \{\chi_\alpha, \bar{\chi}_\beta\} = g_{\alpha\beta} \tag{1.78}$$

which tells us that we may regard $\bar{\chi}_\alpha$ and χ_β as creation and annihilation operators respectively. The states of the theory are described by wavefunctions $\varphi(X)$ over the moduli space $\mathcal{I}_{k,N}$, acted upon by some number p of fermion creation operators. We write $\varphi_{\alpha_1,\ldots,\alpha_p}(X) \equiv \bar{\chi}_{\alpha_1} \ldots \bar{\chi}_{\alpha_p} \varphi(X)$. By the Grassmann nature of the fermions, these states are anti-symmetric in their p indices, ensuring that the tower stops when $p = \dim(\mathcal{I}_{k,N})$. In this manner, the states can be identified with the space of all p-forms on $\mathcal{I}_{k,N}$.

The Hamiltonian of the theory has a similarly natural geometric interpretation. One can check that the Hamiltonian arising from (1.30) can be written as

$$H = QQ^\dagger + Q^\dagger Q \tag{1.79}$$

where Q is the supercharge which takes the form $Q = -i\bar{\chi}_\alpha p_\alpha$ and $Q^\dagger = -i\chi_\alpha p_\alpha$, and p_α is the momentum conjugate to X^α. Studying the action of Q on the states above, we find that $Q = d$, the exterior derivative on forms, while $Q^\dagger = d^\dagger$, the adjoint operator. We can therefore write the Hamiltonian as the Laplacian acting on all p-forms,

$$H = dd^\dagger + d^\dagger d \tag{1.80}$$

We learn that the space of ground states $H = 0$ coincide with the harmonic forms on the target space.

There are two subtleties in applying this analysis to instantons. The first is that the instanton moduli space $\mathcal{I}_{k,N}$ is singular. At these points, corresponding to small instantons, new UV degrees of freedom are needed. Presumably this reflects the non-renormalizability of the five-dimensional gauge theory. However, as we have seen, one can resolve the singularity by turning on non-commutivity. The interpretation of the instantons as KK modes only survives if there is a similar non-commutative deformation of the $(2, 0)$ theory which appears to be the case.

The second subtlety is an infra-red effect: the instanton moduli space is non-compact. For compact target spaces, the ground states of the sigma-model coincide with the space of harmonic forms or, in other words, the cohomology. For non-compact target spaces such as $\mathcal{I}_{k,N}$, we have the further requirement that any putative ground state wavefunction must be normalizable and we need to study cohomology with compact support. With this in mind, the relationship between the five-dimensional theory and the six-dimensional $(2,0)$ theory therefore translates into the conjecture

There is a unique normalizable harmonic form on $\mathcal{I}_{k,N}$ for each k and N

Note that even for a single instanton, this is non-trivial. As we have seen above, after resolving the small instanton singularity, the moduli space for a $k=1$ instanton in $U(N)$ theory is $T^\star(\mathbf{CP}^{N-1})$, which has Euler character $\chi = N$. Yet, there should be only a single groundstate. Indeed, it can be shown explicitly that of these N putative ground states, only a single one has sufficiently compact support to provide an L^2 normalizable wavefunction [62]. For an arbitrary number of k instantons in $U(N)$ gauge theory, there is an index theorem argument that this unique bound state exists [63].

So much for the ground states. What about the N^3 degrees of freedom. Is it possible to see this from the five-dimensional gauge theory? Unfortunately, so far, no one has managed this. Five dimensional gauge theories become strongly coupled in the ultra-violet where their non-renormalizability becomes an issue and we have to introduce new degrees of freedom. This occurs at an energy scale $E \sim 1/e^2N$, where the 't Hooft coupling becomes strong. This is parametrically lower than the KK scale $E \sim 1/R \sim 1/e^2$. Supergravity calculations reveal that the N^3 degrees of freedom should also become apparent at the lower scale $E \sim 1/e^2N$ [64]. This suggests that perhaps the true degrees of freedom of the theory are "fractional instantons", each with mass $M_{\rm inst}/N$. Let me end this section with some rampant speculation along these lines. It seems possible that the $4kN$ moduli of the instanton may rearrange themselves into the positions of kN objects, each living in $\mathbf{R}^4$ and each, presumably, carrying the requisite mass $1/e^2N$. We shall see a similar phenomenon occurring for vortices in Section 3.8.2. If this speculation is true, it would also explain why a naive quantization of the instanton leads to a continuous spectrum, rather strange behavior for a single particle: it's because the instanton is really a multi-particle state. However, to make sense of this idea we would need to understand why the fractional instantons are confined to lie within the instanton yet, at the same time, are also able to wander freely as evinced by the $4kN$ moduli. Which, let's face it, is odd! A possible explanation for this strange behavior may lie in the issues of non-normalizability of non-abelian modes discussed above, and related issues described in [65].

While it's not entirely clear what a fractional instanton means on $\mathbf{R}^4$, one can make rigorous sense of the idea when the theory is compactified on a further $\mathbf{S}^1$ with a Wilson line [66, 67]. Moreover, there's evidence from string dualities [68, 39] that the moduli space of instantons on compact spaces $\mathbf{M} = \mathbf{T}^4$ or $K3$ has the same topology as the symmetric product $\mathrm{Sym}^{kN}(\mathbf{M})$, suggesting an interpretation in terms of kN entities (strictly speaking, one needs to resolve these spaces into an object known as the Hilbert scheme of points over $\mathbf{M}$).