1. Instantons

1.1. Notation. Throughout this talk, we will use the following notation:

- \( G \): a (semi-simple) Lie group, typically \( SU(n) \)
- \( \mathfrak{g} \): its Lie algebra
- \( X \): a (simply connected) 4-fold (typically, \( S^4 \))
- \( P \): a principal \( G \)-bundle \( \pi : P \to X \)
- \( A \): a connection on \( P \)
- \( F_A \): its curvature
- \( \Omega^p_X(\mathfrak{g}) \): differential \( p \)-forms on \( X \) with values in \( \mathfrak{g} \), i.e., \( \Omega^1_X(\mathfrak{g}) := \Gamma(T^*X \otimes \mathfrak{g}) \).

1.2. Primer on Connections. Recall from Peng’s talk that a connection on \( P \to X \) can be thought of in three ways:

1. As a field of horizontal subspaces: \( T_pP = H_pP \oplus V_pP \), where \( V_pP = \ker(\pi_\ast) \).
2. As a \( \mathfrak{g} \)-valued 1-form \( A \in \Omega^1_X(\mathfrak{g}) \) which is invariant w.r.t. the induced \( G \)-action on \( \Omega^1_X(\mathfrak{g}) \).
3. Given a representation of \( G \) on \( W \), as a vector bundle connection on the associated bundle \( E := P \times_G W \to X \)

\( \nabla_A : \Omega^k_X(E) \to \Omega^{k+1}_X(E) \),

where \( \nabla_A \) is a linear map satisfying the Leibniz rule: \( \nabla(f \cdot s) = f \cdot \nabla s + df \cdot s \) for \( f \in C^\infty(X) \), \( s \in \Gamma(X,E) \).

The first two of these definitions are seen to be equivalent by setting \( H_pP = \ker A_p \) for (2) \( \to \) (1), or \( A_p = T_pP \to V_pP \) (projection) for (1) \( \to \) (2). For (3), we think of \( P \) as being the frame bundle for \( E \), and then describe a horizontal frame of sections.

For concreteness, we will mostly use (3) in this talk, i.e., fix a vector bundle \( E \to X \), and then think of \( P \) as the frame bundle of \( E \). However, everything can still be done for principal bundles, too.

From \( \nabla_A \), we can build a new operator

\( d_A : \Omega^k_X(E) \to \Omega^{k+1}_X(E) \)

by requiring that \( d_A = \nabla_A \) for sections of \( E \) and \( d_A(\omega \wedge \theta) = (d_A\omega) \wedge \theta + (-1)^{|\omega|} \omega \wedge d_A\theta \). In general, \( d_A^2 \neq 0 \), and we give this a special name: the curvature

\( F_A := d_A^2 : E \to \Omega^2_X(E) \).
Locally, 
\[ d_A = d + A \wedge, \]
where \( A \in \Omega^1_X(E) \) is an \( E \)-valued 1-form, and
\[ F_A = dA + A \wedge A \in \Omega^2_X(E). \]

The covariant derivative along \( v \in T_X \) of a section \( s \) is given by \( \iota_v dA s \).

If \( X \) has the additional structure of a Riemannian metric, the formal adjoint \( d_A^* \) to \( d_A \) can be defined:
\[ \int_X \langle d_A \phi, \psi \rangle \, d\mu = \int_X \langle \phi, d_A^* \psi \rangle \, d\mu. \]

If in addition \( \dim X = 4 \), then from Hodge theory, 2-forms on \( X \) decompose into self-dual and anti-self-dual parts. This extends to \( E \)-valued forms
\[ \Omega^2_X(E) = \Omega^+_X(E) \oplus \Omega^-_X(E), \]
so \( F_A = F^+_A + F^-_A \).

A connection \( A \) is called anti-self-dual (ASD) if \( F^+_A = 0 \) (i.e., \( *F_A = -F_A \)).

1.3. **Yang-Mills Theory.** Yang-Mills theory is a field theory defined for principal \( G \) bundles \( \pi: P \to X \). The fields of the theory are connections, and the action is (up to some constants)
\[ (1.1) \quad S(A) := \int_X |F_A|^2 \, d\mu. \]

\( S \) is conformally invariant in dimension 4: if \( g \mapsto cg \) is a conformal transformation, then \( d\mu \mapsto c^d d\mu \) and \( F_A \mapsto c^{-2} F_A \), so for \( \dim X = d = 4 \),
\[ \int_X c^{d-4} |F_A|^2 \, d\mu = \int_X |F_A|^2 \, d\mu. \]

For a \( G \)-invariant metric (which can be readily constructed if \( G \) is compact), this action is gauge-invariant. \( |F_A|^2 \) is sometimes called the **Yang-Mills density**.

**Proposition 1.1.** The Euler-Lagrange equations of this action are
\[ (1.2) \quad d_A^* F_A = 0. \]

**Proof.** This is an exercise in variational calculus. I’ll skip most of the algebra. Observe that
\[ F_{A+t\tau} = d(A + t\tau) + (A + t\tau) \wedge (A + t\tau) = F_A + td_A \tau + t^2 \tau \wedge \tau. \]

Then,
\[ |F_{A+t\tau}|^2 = |F_A|^2 + 2t \langle d_A \tau, F_A \rangle + t^2 (\cdots), \]
so
\[ 0 = \frac{d}{dt} S(A + t\tau) = 2 \int_X \langle d_A \tau, F_A \rangle \, d\mu = 2 \int_X \langle \tau, d_A^* F_A \rangle \, d\mu. \]

Hence the equations of motion are
\[ d_A^* F_A = 0. \]
An instanton is a topologically nontrivial solution to the classical equations of motion with finite action.

**Proposition 1.2.** Anti-self-dual connections are instantons, i.e., topologically nontrivial solutions to (1.2).

**Proof.** First we show that an ASD connection solves the equations of motion. The main fact is
\[ \ast d_A F_A = d_A \ast F_A, \]
so if A is ASD, \( d_A \ast F_A = -d_A F_A = 0 \) by the Bianchi identity.

When \( \dim X = 4 \) and \( G = SU(n) \), ASD connections are topologically nontrivial: for skew-adjoint matrices \((A^* = -A)\), i.e., \( u(n)\),
\[ \text{tr} (\xi \wedge \xi) = -|\xi|^2, \]
so
\[ \text{tr} (F_A^2) = -\left(|F_A^+|^2 - |F_A^-|^2\right) d\mu. \]

\[ |F_A|^2 = F_A \wedge * F_A = |F_A^+|^2 + |F_A^-|^2, \]
so a connection is ASD if and only if \( \text{tr} (F_A^2) = |F_A|^2 d\mu \) at every \( x \in X \). Recall that for \( SU(n) \) bundles, \( c_1(E) \) vanishes because \( \text{tr} (F_A) = 0 \), so
\[ c_2(E) = \frac{1}{8\pi^2} \int_X \text{tr} (F_A^2) \, d\mu. \]

Hence,
\[ S(A) = \int_X |F_A|^2 \, d\mu = \int_X |F_A^-|^2 \, d\mu + \int_X |F_A^+|^2 \, d\mu \geq 8\pi^2 c_2(E), \]
with the bound achieved precisely when \( A \) is ASD. For this reason, physicists often refer to \( c_2(E) \) as the “instanton number.” \( \square \)

2. ADHM Construction

Let
\[ F_{ij} := [\nabla_i, \nabla_j] = \frac{\partial}{\partial x_i} A_j - \frac{\partial}{\partial x_j} A_i + [A_i, A_j]. \]

Then, instanton equation \( F_A^+ = 0 \) becomes
\[
\begin{align*}
F_{12} + F_{34} &= 0, \\
F_{14} + F_{23} &= 0, \\
F_{13} + F_{42} &= 0.
\end{align*}
\]
(2.1)

The ADHM (Atiyah, Drinfeld, Hitchin, and Manin) construction gives a way of producing ASD connections from linear algebraic data. The idea is to take a “Fourier transform” of the ASD equations to produce a set of matrix equations which can be more readily solved.

Substituting \( D_1 := \nabla_1 + i \nabla_2, \ D_2 := \nabla_3 + i \nabla_4, \) the equations (2.1) become
\[
\begin{align*}
[D_1, D_2] &= (F_{13} + F_{42}) + i (F_{23} + F_{14}) = 0, \\
[D_1, D_1^+] + [D_2, D_2^+] &= -2i (F_{12} + F_{34}) = 0,
\end{align*}
\]
so we can reduce the ASD equations to these “complex” covariant derivatives.

Because the ASD equation and \( |F_A|^2 \) are conformally invariant, an ASD connection on \( \mathbb{R}^4 \) with \( S(A) < \infty \) can be regarded as an ASD connection on \( S^4 \).
2.1. ADHM Data. Let $U \cong \mathbb{C}^2$ as a complex manifold, with coordinates $(z_1, z_2)$. The inputs for the ADHM construction consist of:

(1) A $k$-dimensional complex vector space $H$ with a Hermitian metric.
(2) An $n$-dimensional complex vector space $E_\infty$, with Hermitian metric and symmetry group $SU(n)$.
(3) A linear map $T \in V^* \otimes \text{hom} (H, H)$ defining maps $\tau_1, \tau_2 : H \to H$.
(4) Linear maps $\pi : H \to E_\infty$ and $\sigma : E_\infty \to H$.

The ADHM equations are

\[
[\tau_1, \tau_2] + \sigma \pi = 0, \\
[\tau_1, \tau_1^*] + [\tau_2, \tau_2^*] + \sigma \sigma^* - \pi^* \pi = 0.
\]

If $\tau_1, \tau_2, \sigma, \pi$ satisfy these equations, then the maps

\[
\alpha := \begin{bmatrix} \tau_1 \\ \tau_2 \\ \pi \end{bmatrix}, \quad \beta := \begin{bmatrix} -\tau_2 & \tau_1 & \sigma \end{bmatrix}
\]

define a complex

\[
\begin{array}{ccc}
H & \overset{\alpha}{\longrightarrow} & H \oplus U \oplus E_\infty \\
\end{array}
\]

because

\[
\beta \alpha = [\tau_1, \tau_2] + \sigma \pi = 0.
\]

In fact, it defines a whole $\mathbb{C}^2$-family of complexes because we can replace $(\tau_1, \tau_2)$ by $(\tau_1 - z_1 \cdot 1, \tau_2 - z_2 \cdot 1)$ for any point $(z_1, z_2) =: x \in U$. We can then define a family of maps

\[
R_x : H \oplus U \oplus E_\infty \to H \oplus H
\]

and, if $\alpha_x$ is injective and $\beta_x$ is surjective, then $R_x$ is surjective and $\ker R_x = (\text{im } \alpha_x)^{\perp} \cap \ker \beta_x$.

**Definition.** A collection $(\tau_1, \tau_2, \sigma, \pi, E_\infty, H)$ of ADHM data is an ADHM system if

1. it satisfies the ADHM equations (2.2), and
2. the map $R_x$ is surjective for each $x \in U$.

2.2. ADHM Construction. How can we use this information to construct a connection? Suppose that we have an ADHM system. Now, construct the vector bundle $E \to U$ with fibers

\[
E_x = \ker R_x = \ker \beta_x / \text{im } \alpha_x.
\]

($E_x$ is the cohomology bundle of $\alpha, \beta$).

**Proposition 2.1.** There is a holomorphic structure $\mathcal{E}$ on $E$.

(Proof omitted in the interest of time).

Let $i : E_x \hookrightarrow H \oplus U \oplus E_\infty$ be inclusion, $P_\alpha^\beta : H \oplus U \oplus E_\infty \to (\text{im } \alpha_x)^{\perp}$ and $P_\beta^\alpha : H \oplus U \oplus E_\infty \to \ker \beta_x$ denote orthogonal projections, and $P_x := P_\alpha \circ P_\beta = P_\beta \circ P_\alpha$. 


be projection onto $E_x$. $H \otimes U \oplus E_\infty$ comes equipped with a flat product connection $d$, and we can define an induced connection $A$ on $E$ by, for a section $s : U \to E$, 
\[ d_A s := Pd_i(s). \]
By virtue of this construction, $A$ is unitary and compatible with the holomorphic structure on $E$. It is also ASD.

**Theorem 2.2 (ADHM).** The assignment $(\tau_1, \tau_2, \sigma, \pi) \mapsto d_A$ sets up a one-to-one correspondence between

1. equivalence classes of ADHM data for group $SU(n)$ and index $k$, and
2. gauge equivalence classes of finite energy ASD $SU(n)$-connections $A$ over $\mathbb{R}^4$ with $c_2(A) = k$.

Note that $(g, h) \in U(k) \times SU(n)$ acts on ADHM data by
\[
(\tau_1, \tau_2, \sigma, \pi) \mapsto (g\tau_1 g^{-1}, g\tau_2 g^{-1}, g\sigma h^{-1}, h\pi g^{-1}),
\]
so we mean classes of ADHM data up to this equivalence.

**2.3. Example: BPST Instanton.** The simplest example is to take $k = 1$ and $n = 2$. This corresponds to solutions on $SU(2)$ bundles with $c_2 = 1$. Then, $\tau_1, \tau_2$ are just complex numbers, $\sigma$ and $\pi$ are complex vectors, and the ADHM equations become
\[
\sigma \cdot \pi = 0, \quad |\sigma|^2 = |\pi|^2.
\]
Pick $\pi = (1, 0)$ and $\sigma = (0, 1)$, then for $(\tau_1, \tau_2)$, have
\[
\alpha^*_x = \begin{bmatrix} \tau_1 \\ \tau_2 \\ 1 \\ 0 \end{bmatrix}, \quad \beta_x = \begin{bmatrix} -\tau_2 \\ \tau_1 \\ 0 \\ 1 \end{bmatrix},
\]
and in general: replace $(\tau_1, \tau_2)$ by $(\tau_1 - z_1 \cdot 1, \tau_2 - z_2 \cdot 1)$ for any point $(z_1, z_2) =: x \in U$, so
\[
R_x = \begin{bmatrix} \tau_1 - z_1 \\ \tau_2 - z_2 \\ -\tau_2 - z_2 \\ \tau_1 - z_1 \end{bmatrix}.
\]
In particular, for $(\tau_1, \tau_2) = (0, 0)$, have
\[
R_x = \begin{bmatrix} -\tau_1 \\ -\tau_2 \\ z_2 \\ -z_1 \end{bmatrix}.
\]
A unitary basis for $R_x$ is
\[
\{\sigma_1, \sigma_2\} = \left\{ \frac{1}{1 + |x|^2} \begin{bmatrix} 1 \\ 0 \\ \frac{\overline{z}_1}{\overline{z}_2} \\ -\frac{z_1}{z_2} \end{bmatrix}, \frac{1}{1 + |x|^2} \begin{bmatrix} 1 \\ 0 \\ \frac{z_1}{\overline{z}_2} \\ -\frac{\overline{z}_1}{z_2} \end{bmatrix} \right\}.
\]
Suppose that in this trivialization we let $z_1 = x_1 + ix_2$, $z_2 = x_3 + ix_4$, and the connection matrix is
\[
A = \sum A_i dx_i,
\]
so $A_i$ is the matrix with $(p, q)$th entry
\[
\langle \nabla_i \sigma_p, \sigma_q \rangle = \left\langle \frac{\partial \sigma_p}{\partial x_i}, \sigma_q \right\rangle.
\]
Then, written out in full, the connection form is

\[ A = \frac{1}{1 + |x|^2} (\theta_1 i + \theta_2 j + \theta_3 k), \]

where \(i, j, k\) are a standard basis for \(su(2)\) and

\[
\begin{align*}
\theta_1 &= x_1 dx_2 - x_2 dx_1 - x_3 dx_4 + x_4 dx_3 \\
\theta_2 &= x_1 dx_3 - x_3 dx_1 - x_4 dx_2 + x_2 dx_4, \\
\theta_3 &= x_1 dx_4 - x_4 dx_1 - x_3 dx_2 + x_2 dx_3
\end{align*}
\]

is such that \(d\theta_1, d\theta_2, d\theta_3\) is a basis for the ASD two-forms on \(\mathbb{R}^4\). The curvature \(F_A = dA + A \wedge A\) is then

\[ F_A = \left( \frac{1}{1 + |x|^2} \right)^2 (d\theta_1 i + d\theta_2 j + d\theta_3 k), \]

and we can recover the other degrees of freedom lost in our choices of \(\pi, \sigma, \tau_1, \tau_2\) by translations \(x \mapsto x - y\) and dilations \(x \mapsto x/\lambda\) to obtain other connections with

\[ |F_A(y, \lambda)| = \frac{\lambda^2}{\left( \lambda^2 + |x - y|^2 \right)^2}. \]

3. Moduli Space of ASD Connections

**Definition 3.1.** Let \(E \to X\) be a bundle over a compact, oriented Riemannian 4-manifold \(X\). The **moduli space of ASD connections** \(M_E\) is the set of gauge equivalence classes of ASD connections on \(E\).

Recall that a **gauge transformation** is an automorphism \(u : E \to E\) respecting the structure on the fibers and reducing to the identity map on \(X\). It acts on a connection by the rule

\[ \nabla_{u(A)} s = u \nabla_A (u^{-1} s) = \nabla_A s - (\nabla_A u) u^{-1} s, \]

where the covariant derivative \(\nabla_A u\) is formed by regarding it as a section of the vector bundle \(\text{End}(E)\). In local coordinates, this looks like

\[ u(A) = uAu^{-1} - (du) u^{-1}. \]

The curvature transforms as a tensor under gauge transformations:

\[ F_{u(A)} = u F_A u^{-1}. \]

For connections on principal bundles \(P \to X\), this has a somewhat nicer expression: If \(u : P \to P\) satisfies

\[
\begin{align*}
(1) & \quad u(p \cdot g) = u(p) \cdot g \quad \text{and} \\
(2) & \quad \pi(u(p)) = \pi(p)
\end{align*}
\]

for all \(g \in G\), and \(A \in \Omega^1_P(g)\) is a connection,

\[ u(A) := (u^{-1})^* A. \]

Now we turn to some results about the structure of this moduli space.
3.1. Uhlenbeck’s Theorems. First, there are a few technical results due to Uhlenbeck that allow us to leverage tools from the study of elliptic differential equations to make statements about ASD connections.

**Theorem 3.2 (Uhlenbeck).** There are constants $\epsilon_1, M > 0$ such that any connection $A$ on the trivial bundle over $B^4$ with $||F_A||_{L^2} < \epsilon_1$ is gauge equivalent to a connection $\tilde{A}$ over $B^4$ with

1. $d^*\tilde{A} = 0$,
2. $\lim_{|x| \to 1} \tilde{A}_r = 0$, and
3. $||\tilde{A}||_{L^2} \leq M||F_A||_{L^2}$.

Moreover for suitable constants $\epsilon_1, M$, $\tilde{A}$ is uniquely determined by these properties, up to $\tilde{A} \mapsto u_0\tilde{A}u_0^{-1}$ for a constant $u_0$ in $U(n)$.

First, some notes about the theorem:

$$||\tilde{A}||_{L^2}^2 = \int_{B^4} |\nabla \tilde{A}|^2 + |\tilde{A}|^2 d\mu$$

is the Sobolev norm. $d^*\tilde{A}$ is the “Coulomb” gauge condition (the importance of which will be explained in the following section). Finally, $\lim_{|x| \to 1} \tilde{A}_r = 0$ means that, for $\tilde{A}_r(\rho, \sigma)$ a function on $S^3$, this function tends to 0 as $r \to 1$.

The main power of Uhlenbeck’s Theorem is that it turns a system of nonlinear, nonelliptic differential equations into an elliptic one. This section provides a sketch of why that might be a desirable thing to do. Recall the $d^+$ operator, defined by

$$d^+ = \left(\frac{1}{2} (1 + *)\right) \circ d,$$

which maps

$$d^+ : \Omega^1_X \to \Omega^1_X.$$

The ASD equation $F_A^+ = 0$ then becomes, in local coordinates,

$$(3.1) \quad d^+ A + (A \wedge A)^+ = 0.$$

This is a nonlinear, non-elliptic equation.

When $d^* A = 0$,

$$d^* + d : \oplus_i \Omega^{i+1}_X \to \oplus_i \Omega^i_X,$$

is elliptic, so if $H^1(X) = 0$, then all 1-forms are orthogonal to $\ker (d + d^*)$.

Elliptic differential operator theory implies that

$$(3.2) \quad ||A||_{L^2_k} \leq \text{const.} \left(||d^* A||_{L^2_{k-1}} + ||dA||_{L^2_{k-1}}\right)$$

for all $k$. When $d^* A = 0$, this becomes

$$||A||_{L^2_k} \leq \text{const.} \cdot ||F_A||_{L^2_{k-1}},$$

and the ASD equation can be replaced by the elliptic differential equation

$$\delta A = 0,$$

where $\delta = d^* + d^*$ is an elliptic operator.

The main consequence of Uhlenbeck’s Theorem relevant to the discussion of ASD connections comes from combining it with the following theorem:
Theorem 3.3 (Uhlenbeck). There exists a constant $\epsilon_2 > 0$ such that if $\tilde{A}$ is any ASD connection on the trivial bundle over $B^4$ which satisfies $d^* \tilde{A} = 0$ and $\|\tilde{A}\|_{L^4} \leq \epsilon_2$, then for all interior domains $D \subset B^4$ and $l \geq 1$,

$$\|\tilde{A}\|_{L^2(D)} \leq M_{l,D}\|F_{\tilde{A}}\|_{L^2(B^4)}$$

for a constant $M_{l,D}$ depending only on $l$ and $D$.

Combining this with Theorem 3.2 gives

Corollary 3.4. For any sequence of ASD connections $A_\alpha$ over $B^4$ with $\|F(A_\alpha)\|_{L^2} \leq \epsilon$, there is a subsequence $A_\alpha'$ and gauge equivalent connections $\tilde{A}_\alpha'$ which converge in $C^\infty$ on the open ball.

3.2. Results about the Moduli Space. Putting our previous results together, we get the following statements:

Theorem 3.5 (Uhlenbeck’s Removable Singularities). Let $A$ be a unitary connection over the punctured ball $B^4 \setminus \{0\}$ which is ASD with respect to a smooth metric on $B^4$. If

$$\int_{B^4 \setminus \{0\}} |F_A|^2 < \infty,$$

then there is a smooth ASD connection over $B^4$ gauge equivalent to $A$ over the punctured ball.

Note that this theorem implies that, for example, the ADHM construction gives all of the ASD connections on $S^4$ (not just $\mathbb{R}^4$).

Let $M_k(G)$ denote the moduli space of ASD connections up to gauge transformation with $c_2 = k$, and $\overline{M}_k(G)$ denote the closure of $M_k(G)$ in the space of “ideal connections.” An ideal connection is a connection with curvature densities possibly having $\delta$-measure concentrations at up to $k$ points of $X$, i.e., of the form

$$|F_A|^2 + 8\pi^2 \sum_{i=1}^{n} \delta_{x_i}$$

Then,

Theorem 3.6. Any infinite sequence in $M_k$ has a weakly convergent subsequence in $\overline{M}_k$, with limit point in $\overline{M}_k$.

Corollary 3.7. The space $\overline{M}_k$ is compact.

What do these spaces look like locally? Let $\mathcal{G}$ denote the group of gauge transformations of $E \to X$, and

$$\Gamma_A = \{ u \in \mathcal{G} : u(A) = A \},$$

the isotropy group of $A$. Then,

Proposition 3.8. If $A$ is an ASD connection over $X$, a neighborhood of $[A]$ in $M$ is modeled on a quotient $f^{-1}(0)/\Gamma_A$, where

$$f : \ker \delta_A \to \coker d_A^+$$

is a $\Gamma_A$-equivariant map and $\delta_A = d_A^* + d_A^+$.
REFERENCES

