Yang-Mills Instantons

Seongmin Kim, Taeyoon Kim

June 20, 2024

1 Introduction : What is an Instanton?

What is instanton? Instanton is a solution to the classical field equations of motion in Euclidean space. Instanton solutions of Euclidean EL equation are localized in (Euclidean) space and time, and have finite (Euclidean) action. This is why it is called instanton. Instantons only appear in Euclidean space, and they are not solutions to the equations of motion in Minkowski space. So, they are not physical particles or fields, in Real spacetime. So, why do we care about instantons? The answer is that we essentially need wick rotation and Euclidean space to perform path integral calculations in (Minkowski) quantum field theory. So even we are interested in Minkowski space, we should consider instantons, which are classical solutions in Euclidean space. In detail, how do instantons affect the QFT in Minkowski space? We will discuss how the instanton effect appears in the path integral formulation, and how it is related to various physical phenomena.

2 Instanton effect in Path Integral : a Toy Model

2.1 $0 + 1$ Dimensional Toy Model : QM with a double well potential

Let's consider a simple toy model, a quantum mechanical system with a double well potential. The action of the system is given by

$$S = \int dt \left(\frac{1}{2}\dot{x}^2 - V(x) \right) \tag{1}$$

where the potential is given by $V(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2$. The classical Vaccume structure of the system is shown in Fig. 1.

Figure 1: The vaccume structure of the double well potential, there are two minima at $x = \pm 1$.

Classically, the system has two minima at $x = \pm 1$, but quantum mechanically, the system can tunnel between the two minima. So the true quantum ground state of the system is a superposition of the two classical ground states, and this splitting of the ground state is totally determined by the tunneling amplitude between the two minima. Now, we will evaluate the tunneling amplitude using the path integral formulation of quantum mechanics, including the instanton solution of the system.

2.2 Path Integral Formulation and Wick Rotation

When we consider the path integral of the system, the path integral is given by $Z = \int \mathcal{D}x e^{iS[x]}$. But, this integral doesn't converge well, so we need to consider the path integral in Euclidean space, $Z = \int \mathcal{D}x e^{-S_E[x]}$, where $S_E[x] = \int dt \left(\frac{1}{2}\dot{x}^2 + V(x) \right)$. And to evaluate the original path integral, we just need to wick rotate the time variable, $t \to -i\tau$.

This looks like just a mathematical trick, and not an essential part of the theory, but in fact, this wick rotation is essential. Why? There is no way to consider the instanton effect in the path integral formulation in Minkowski space, because the instanton solutions only appear in Euclidean space. Then, is there any posibility to evaluate the instanton effect in Minkowski space? Like using the perturbative expansion of the path integral? The answer is no. The instanton effect is a non-perturbative effect, and it cannot be evaluated by the perturbative expansion of the path integral. So, the wick rotation is essential to evaluate the instanton effect in the path integral formulation of quantum field theory.

2.3 Instanton Solutions in Euclidean Space

The instanton solution of the system is given by the bounce solution, which is a solution to the Euclidean equation of motion, $\delta S_E = 0$. The bounce solution is a solution that interpolates between the two minima of the potential, and it is localized in Euclidean time. The action of the bounce solution is finite, and it can be evaluated by the Euclidean action of the bounce solution, $S_{\text{bounce}} = \int d\tau \left(\frac{1}{2}\dot{x}^2 + V(x)\right)$.

2.4 Instanton Contribution to the Path Integral

The path integral of the system is given by the sum of all possible paths. In the semiclassical limit, the path integral is dominated by the classical solutions, (i.e., the instanton solutions) and the small fluctuations around them. So the full path integral is given by the sum of all instanton contributions, with small fluctuations around them.

$$Z = \int \mathcal{D}x e^{-S[x]} = \sum_{\text{instantons}} e^{-S_{\text{inst}}} \int \mathcal{D}\delta x e^{-S_{\text{fluct}}[\delta x]} \tag{2}$$

where S_{inst} is the action of the instanton solution, and S_{fluct} is the action of the small fluctuations around the instanton solution. This is how the instanton effect appears in the path integral formulation of quantum field theory. Normally, the instanton effect is not easy to calculate, but in this toy model, we can calculate the instanton effect by hand, with the bounce solution. To evaluate the path integral, we need to consider all the possible classical paths. Any classical path between the two minima of the potential can be considered as a composition of n single bounce solutions, at time $t_1, ...t_n$, and the instanton action of this classical path is given by $S = \sum_{i=1}^{n} S_{\text{inst}} = nS_{\text{inst}}$.

So, the Euclidean path integral, from $x = -1$ to $x = 1$, is given by the sum as :

$$\langle x = 1|e^{-HT}|x = -1\rangle = \sum_{n \text{ odd}} e^{-nS_{\text{inst}}} \int_0^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{n-1}}^T \mathrm{d}t_n \int \mathcal{D}\delta x e^{-S_{\text{fluct}}} \tag{3}$$

The path integral of the fluctuation can be factorized into $n + 1$ parts, which are the fluctuation path integrals around each instanton solution, and one for the fluctuation around the classical ground state. Using this factorization, we can approximated, $\int \mathcal{D}\delta x e^{-S_{\text{fluct}}} \approx \left(\int \mathcal{D}\delta x_0 e^{-S_{\text{fluct}0}}\right)^n e^{-T/2} = K^n e^{-T/2}$ Finaly, we get

$$\langle x = 1|e^{-HT}|x = -1\rangle = e^{-T/2} \sum_{n \text{ odd}} \frac{T^n}{n!} K^n e^{-nS_{\text{inst}}} = e^{-T/2} \sinh(KTe^{-S_{\text{inst}}}) \tag{4}$$

So this is the matrix element of e^{-HT} from $x = -1$ to $x = 1$, using the path integral including instanton solution. To evaluate the propagator in Minkowski space, one cas just change the time variable back to Minkowski time, $T \to -iT$. Then, the propagator is given by $\langle x = 1|e^{-itH}|x = -1\rangle = e^{-it/2} \sin(Kte^{-S_{\text{inst}}})$.

2.5 Tunneling Amplitude, Instanton Action and WKB Approximation

In Hamiltonian formalism, the oscillatory behavior of the propagator

$$\langle x = 1|e^{-itH}|x = -1\rangle = e^{-it/2} \sin(Kte^{-S_{\text{inst}}}) \tag{5}$$

can be understood as an oscilation due to the tunneling amplitude between the two minima of the potential.

The tunneling amplitude, $\langle x = 1|H|x = -1\rangle$, can be evaluated with the WKB approximation, which can be evaluated by approximating the shrodinger equation with the WKB wavefunction.

The result is well known. The tunneling amplitude from two classical vaccume with enegry E, from $x = a$ to $x = b$ is generally given by

$$\langle x = a|H|x = b\rangle \propto \exp\left(-\int_a^b dx\ \sqrt{2m(V(x) - E)}\right) \tag{6}$$

We can easily show that this WKB approximation is equivalent to the path integral result :

$$\dot{x}^2 = 2(V(x) - E)/m \to \dot{x} = \sqrt{2(V(x) - E)/m} \tag{7}$$

$$S_{\text{inst}} = \int d\tau \left(\frac{1}{2}m\dot{x}^2 + V(x)\right) = \int d\tau \left(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{x}^2 + E\right) \tag{8}$$

$$= \int dx\ m\dot{x} + E\tau = \int dx\ \sqrt{2m(V(x) - E)} + E\tau \tag{9}$$

So the instanton action is the same as the WKB action, and the tunneling amplitude is given by the instanton action, $e^{-S_{\text{inst}}}$. Instanton is a essential ingredient to evaluate the tunneling amplitude in the path integral formulation, and the same result can be obtained by the WKB approximation. Here one might wonder, can't we just obtatain the tunneling amplitude by path integral in Minkowski space? Well, the answer is no. We will discuss this in the next section.

2.6 The non-perturbative behavior of the propagator

The propagator in the double well potential was calculated as

$$\langle x = 1|e^{-itH}|x = -1\rangle = e^{-it/2}\sin(Kte^{-S_{\text{inst}}}) \tag{10}$$

In this section, we will discuss the non-perturbative behavior of the propagator. Here, the term "non-perturbative" means that phisical quantities are not analytic in $\hbar$, and cannot be evaluated by the perturbative expansion of the path integral. So we have to re-introduce $\hbar$ to our calculation, and show that this is not an analytic function of $\hbar$, at $\hbar = 0$.

With $\hbar$, the propagator is given by

$$\langle x = 1|e^{-itH/\hbar}|x = -1\rangle = e^{-it/2}\sin(Ke^{-S_{\text{inst}}/\hbar}t/\hbar) \tag{11}$$

The $e^{-S_{\text{inst}}/\hbar}$ term is non-analytic in $\hbar$, and it is not possible to evaluate this term by the perturbative expansion of the path integral. This is because the path integral expansion is in fact, the expansion in $\hbar$. This looks unfamiliar because we usually think of the path integral as the expansion in the coupling constant, and we use $\hbar = 1$ conventionally. But in fact, in the feynman diagrams, the order of $\hbar$ is the order of the loop expansion, so the perturbative path integral expansion gives an asymptotic series in $\hbar$.

So if you don't have a brand new idea to evaluate path integral non-perturbatively in Minkowski space, you have to go to the Euclidean space, and consider the instanton solutions to evaluate all the non-perturbative effects. Going to Euclidean space is not just a mathematical trick, but an essential part of the path integral, in order to overcome the limit of the perturbative expansion of the path integral.

3 Yang-Mills Instanton

Now, you have a basic understanding of the instanton effect in the path integral formulation of quantum theory. Why is it essential, and how does it appear in the path integral, and how does it affect the physical phenomena. The phisics of Yangs-Mills instanton is not so much different from the toy model we discussed, but the mathematical structure is highly more complicated. In this section, we will focus on the mathematical structure of the Yang-Mills instanton, and discuss the "moduli space" of the instanton solutions.

Yang Mills Theory

Remark 1 (Exterior Covariant derivative)**.** *One can define the exterior covariant derivative.*

$$d_\nabla : \Omega^k(M, E) \to \Omega^{k+1}(M, E)$$

With the following properties. For $s \in \Gamma(M, E)$,

$$d_\nabla s = \nabla s$$

For $\omega \in \Omega^k(M, E)$, $\eta \in \Omega^l(M, E)$

$$d_\nabla(\omega \wedge \eta) = d_\nabla\omega \wedge \eta + (-1)^k\omega \wedge d_\nabla\eta$$

Note that this exterior covariant derivative gets along well with our former calculation.

Note 1. *Select a local frame of E as $\{e_1 \cdots e_r\}$. $\omega = \sum \omega^i \otimes e_i$, $\omega_i \in \Omega^k(M)$. One can calculate the exterior covariant derivative.*

$$d_\nabla\omega = d_\nabla e_i \wedge \omega_i + e_i \otimes d\omega_i = A^j_i e_j \wedge \omega^i + e_i \otimes d\omega^i = e_j \otimes (d\omega^j + A^j_i \wedge \omega^i)$$

For some covariant derivative matrix $A \in \Omega^1(M, \text{End}E)$ for local frames.

$$d_\nabla\omega = d\omega + A \wedge \omega$$

Since it is crucial to calculate covariant derivative for endomorphism valued forms, one needs to define the following covariant derivative.

Remark 2 (Covariant Derivative on End$E$)**.** *For $\eta \in \Gamma(M, \mathrm{End}E)$, $s \in \Gamma(M, E)$, the covariant derivative of η is defined as the following.*

$$\nabla_X^{\mathrm{End}E}(\eta)(s) = \nabla_X^E(\eta s) - \eta \nabla_X^E(s)$$

Which inherits the property of the Leibniz rule.

It is valid to define new exterior covariant derivative for endomorphism valued forms.

Proposition 1 (Exterior Covariant Derivative for Endomorphism)**.** *Exterior Covariant derivative for endomorphism also must have the following property. For $\eta \in \Omega^k(M, \mathrm{End}E)$, $\omega \in \Omega^l(M, E)$,*

$$d_\nabla^E(\eta \wedge \omega) = d_\nabla^{\mathrm{End}E}(\eta) \wedge \omega + (-1)^k \eta \wedge d_\nabla^E(\omega)$$

Proof. Rewrite that $\eta = A \otimes \alpha$, $\omega = s \otimes \beta$ For $\alpha, \beta \in \Omega^*(M)$.

$$\begin{aligned}
d_\nabla^E(A(s) \otimes \alpha \wedge \beta) &= \nabla^E(A(s)) \otimes \alpha \wedge \beta + A(s) d_\nabla^E(\alpha \wedge \beta) \\
&= (\nabla^{\mathrm{End}E} A(s) + A(\nabla^E s)) \otimes \alpha \wedge \beta + A(s)(d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta) \\
&= (\nabla^{\mathrm{End}E} A \wedge \alpha + A \otimes \alpha) \wedge (s \otimes \beta) + (-1)^k (A \otimes \alpha)(\nabla^E s \otimes \beta + s \otimes d\beta) \\
&= d_\nabla^{\mathrm{End}E}(A \otimes \alpha) \wedge s \otimes \beta + (-1)^k A \otimes \alpha \wedge d_\nabla^E(s \otimes \beta)
\end{aligned}$$

□

By this calculation, we can assure that endomorphism valued forms admit exterior covariant derivative. And now we omit the End$E$ for such symbols.
For a vector bundle E over M, one can define the Yang-Mills action defined on the space of connections of E.

Definition 1 (Yang-Mills Action)**.**

$$S_{YM}(\nabla) = \int_M \frac{1}{2g^2} \mathrm{tr}(F_\nabla \wedge \star F_\nabla) = \int_M \frac{1}{2g^2} \langle F_\nabla, F_\nabla \rangle \mathrm{Vol}_g$$

In coordinate basis,

$$S_{YM}(\nabla) = \int_M d^4x \frac{1}{2g^2} \mathrm{tr}(F^{\mu\nu} F_{\mu\nu})$$

While F is endomorphism valued 2 forms on M.

More physically, one can make a same description for principal G bundle over M. The associated vector bundle(i.e. its lie algebra bundle) would give the same construction on Yang-Mills Action. It is valid to consider its saddle point, checking classic solutions. This is the natural reason why one has the tr for the action, since the natural metric for (semi-) simple lie algebra is the Killing form.

Proposition 2 (Yang-Mills Equation of motion)**.** *F that lies on the critical point of the action functional satisfies the equation, $d_\nabla^\star F = 0$, while $d_\nabla^\star$ represents the adjoint of d_∇.*

Proof. Consider the following $\nabla_t = \nabla + tA$ for arbitrary A. Which yields the t-dependent $F_t = F_\nabla + t d_\nabla A + \frac{t^2}{2}[A, A]$

$$\left.\frac{d}{dt}\right|_{t=0} S_{YM}(\nabla + tA) = \int_M \frac{1}{g^2} \mathrm{Vol}_g \langle F_\nabla, d_\nabla A \rangle = \int_M \frac{1}{g^2} \mathrm{Vol}_g \langle d_\nabla^\star F_\nabla, A \rangle = 0$$

Giving that $d_\nabla^\star F = 0$. □

Note 2. *One can calculate $d_\nabla d_\nabla \omega$ for $\omega \in \Omega^k(M, E)$.*

$$\begin{aligned}
d_\nabla d_\nabla \omega = d_\nabla(d\omega + A \wedge \omega) &= d(A \wedge \omega) + A \wedge (d\omega + A \wedge \omega) \\
&= (dA + A \wedge A)\omega = F_\nabla \wedge \omega
\end{aligned}$$

Remark 3. *The Bianchi identity, $d_\nabla F = 0$ must be satisfied for any endomorphism valued 2 forms.*

Proof. For any $\omega \in \Omega^1(M, E)$,

$$\begin{aligned}
(d_\nabla)^3 \omega = d_\nabla(F_\nabla \wedge \omega) &= d_\nabla F_\nabla \wedge \omega + F_\nabla \wedge d_\nabla \omega \\
&= (d_\nabla)^2 d_\nabla \omega = F_\nabla \wedge d_\nabla \omega
\end{aligned}$$

Therefore we have $d_\nabla F_\nabla = 0$ □

Through this process, one can derive PDE that F_∇ must satisfy for classic Yang Mills gauge theory. To simplify,

$$d_\nabla^\star F_\nabla = 0, d_\nabla F_\nabla = 0$$

Yang Mills Instanton

Definition 2 (Hodge star operator). *One can define the following Hodge star operator for n-dimensional Riemannian manifold (M, g).*

$$\star : \Omega^k(M) \longrightarrow \Omega^{n-k}(M)$$

such that for any $\eta \in \Omega^k(M)$,

$$\langle \eta, \omega \rangle \mathrm{Volg} = \eta \wedge \star\omega$$

While the Volg *represents the volume form of the Riemannian manifold M.*

Note 3 (Hodge star operator for index notation). *Choosing the coordinate chart $(U, \phi = (x^1, \cdots, x^n))$ for M, one can explicitly write down the Hodge star operator for k forms.*

$$\star(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = \frac{\sqrt{\det g_{ij}}}{(n-k)!} g^{i_1 j_1} \cdots g^{i_k j_k} \epsilon_{j_1 \cdots j_n} dx^{j_{k+1}} \wedge \cdots \wedge dx^{j_n}$$

Note 4. *One can calculate the following $\star\star : \Omega^k(M) \to \Omega^k(M)$. And it is known that $\star\star\,\omega = (-1)^{k(n-k)}\omega$. And for $n = 4$, $k = 2$, $\star\star = \mathrm{id}$ on $\Omega^2(M)$.*

While it is natural to define such star operator on vector valued or endomorphism valued forms, to just apply it on the differential form part. Therefore, one can define the following.

Definition 3 (Yang Mills Instanton). *$F \in \Omega^2(M, \mathrm{End}E)$ that satisfies*

$$\star F = \pm F$$

are called self dual or anti self dual curvature form on M, and they are called the instanton solution of the Yang Mills equation.

Note that one can always decompose F into self dual and anti self dual parts.

$$F = \frac{F + \star F}{2} + \frac{F - \star F}{2} = F_+ + F_-$$

Lemma 1. *$d^\star_\nabla = (-1)^k \star^{-1} d\star$ on $\Omega^k(M)$*

Proof.

$$\langle d\omega, \eta \rangle = \int_M d\omega \wedge \star\eta = \int_M d(\omega \wedge \star\eta) - (-1)^{k-1}\omega \wedge d \star \eta = \int_M (-1)^k \omega \wedge d \star \eta$$

$$= \int_M \omega \wedge \star\star^{-1} (-1)^k d \star \eta = \langle \omega, (-1)^k \star^{-1} d \star \eta \rangle$$

□

Note 5. *(Anti-)Self dual F are the solutions for the Yang Mills equation. The Bianchi identity assures that $d_\nabla F_\nabla = 0$. While $d^\star_\nabla F_\nabla = d_\nabla \star d_\nabla \star F_\nabla = \pm d_\nabla \star d_\nabla F_\nabla = 0$ for Instanton solutions.*

In fact, the Instanton solutions are the absolute minima of the Yang Mills action. This physically means that Instanton solutions are the vacuum solution for the Yang Mills equation.

Proposition 3 (Instanton as the Absolute minima). *YM Instanton is the absolute minima of the YM action.*

Proof.

$$S_{YM}(\nabla) = \frac{1}{2g^2} \int_M \mathrm{tr}(F_\nabla \wedge \star F_\nabla) = \frac{1}{2g^2} \int_M \mathrm{tr}((F_+ + F_-) \wedge \star(F_+ + F_-))$$

$$\frac{1}{2g^2} \int_M \mathrm{tr}((F_+ + F_-) \wedge (F_+ - F_-)) = \frac{1}{2g^2} \left(||F_+||^2 + ||F_-||^2 \right)$$

While

$$c(\nabla) = \frac{1}{2g^2} \int_M \mathrm{tr}(F_\nabla \wedge F_\nabla) = \frac{1}{2g^2} \left(||F_+||^2 - ||F_-||^2 \right)$$

This assures that

$$|c(\nabla)| \leq S_{YM}(\nabla)$$

And the equality holds when $F = F_\pm, F_\mp = 0$ □

The additional information from the former proposition is the value of the absolute minimum. Where the $c(\nabla)$ is the 2nd Chern class of the connection ∇. One can actually give the calculation so that

$$c(\nabla) = \frac{8\pi^2}{g^2} n$$

for some integer $n \in \mathbb{Z}$.

Remark 4 (Chern class has integral coefficients). *Chern class is the Characteristic class, which must naturally related to the cohomology of integral coefficients.*

Characteristic class is the functor from $[-, \mathrm{BG}]$ (i.e. the homotopy classes of maps from space to the classifying space of the Lie group G) to $H^*(-, \mathbb{Z})$. While one knows the categorical equivalences between 2 groupoids, $\mathrm{Bun}_G(X)/\sim$ and $[-, \mathrm{BG}]$. While the $\mathrm{Bun}_G(X)/\sim$ represents the category of principal G bundles over X modulo gauge transformations. This gives us the following diagram.

$$\begin{tikzcd} \mathrm{Bun}_G(-)/\sim \arrow[r, leftrightarrow] & {[-, \mathrm{BG}]} \arrow[d, "Ch.Class"] \\ & H^*(-, \mathbb{Z}) \end{tikzcd}$$

If we are given a data of $\mathrm{Bun}_G(-)/\sim$, there exists a way to have some $H^*(-, \mathbb{R})$. Which is known as the Chern-Weil homomorphism. The only remaining step is to check that such object made by Chern Weil is the characteristic class. The mathematical definition of the Chern class has some axioms, and one can verify if the Chern-Weil satisfies the axioms. While the Chern class can give us the integrality of $U(N)$ bundle, it is possible to explain the integrality of the subgroups of $U(N)$.

Moduli space of YM Instanton

As a sense that Instanton is the classic solution to the YM equation, it is crucial to investigate such solutions. Atiyah-Hitchin-Singer (1977,1978) showed that the moduli space of Instanton, is a finite dimensional manifold for a certain condition. Such condition is for the base 4-manifold, and We will simply review how to calculate the dimension of the moduli space.

Theorem 1 (Atiyah-Hitchin-Singer). *Let M be a compact, self-dual Riemannian manifold with positive scalar curvature. Let P be a principal G-bundle over M, where G is a compact semisimple Lie group. Then the space of moduli of irreducible self-dual connections is either empty or a manifold of dimension*

$$p_1(\mathfrak{g}) - \frac{1}{2} G(\chi - \tau)$$

While χ is the Euler characteristic of M and τ is the signature of M.

First is to understand backgrounds for the theorem. For simplicity, we only think of the connected manifolds.

Definition 4 (Holonomy Group). *For a given connection ∇, with a loop $\gamma : [0, 1] \to M$ that $\gamma(0) = \gamma(1) = p$, the holonomy $P_\gamma(\nabla)$ is a parallel transport around the loop.*

Such parallel transport is defined to be End$E$, in a vector bundle setting. While in the principal bundle setting, such parallel transport can be understood as the transformation on fiber. (This can be proven by choosing the chart, and making it into a problem of ODE.) This means that the parallel transport is not the problem about vectors, but about the Lie group.

Proposition 4 (Gauge invariance of Prallel transport). *For any gauge transformation $\Phi : M \to G$, the parallel transformation of ∇ through a curve γ with Φ, can be written as the following.*

$$P_\gamma^\Phi = \Phi(\gamma(1)) P_\gamma \Phi(\gamma(0))^{-1}$$

Which is trivial from the fact that parallel transport on principal G bundle is the transform on fiber.

Definition 5 (Irreducible Connection).

The irreducibility of connection allows us to claim that gauge transformation acts freely on the space of connections. One knows that all Holonomy groups for each points of M, are isomorphic. Using the gauge invariance of the parallel transport,

$$P_\gamma^\Phi(x) = \Phi^{-1}(x) P_\gamma(x) \Phi(x)$$

If Φ leaves the connection, i.e. if it lies on the kernel of the action,

$$P_\gamma^\Phi(x) = P_\gamma(x) = \Phi^{-1}(x)P_\gamma(x)\Phi(x)$$

Which claims that $\Phi(x)$ is in the center of G. Since ant $g \in G$ can be realized by some γ. For groups like $\mathrm{U}(N)$ or $\mathrm{SU}(N)$, it is trivial that such center is 1, up to scalar. Meaning that it acts (almost) freely. This would play an important role in calculating the dimension of the Instanton (Anti-) self dual moduli space. Moreover, it is known that any semisimple Lie group admits a 0-dimensional center.

Sketch of proof for Thm 1. The main is to investigate the tangent space of the manifold $T_A\mathscr{A}_+/\mathscr{G}$, the self dual Instanton modulo gauge transformation.

$$0 \longrightarrow \Omega^0(M,\mathfrak{g}) \xrightarrow{d_\nabla} \Omega^1(M,\mathfrak{g}) \xrightarrow{d_\nabla^-} \Omega^2_-(M,\mathfrak{g}) \longrightarrow 0$$

The first claim is that if $\tau \in T_A\mathscr{A}/\mathscr{G}$ then $\tau \in \mathrm{Im}d$. Consider the following gauge transformation. Taking the local chart and for any lie algebra $X \in \mathfrak{g}$,

$$A(t) = e^{-tX}Ae^{tX} + e^{-tX}de^{tX}$$

And the differential of the map, i.e. the tangent connection at A, is as the following.

$$\left.\frac{d}{dt}\right|_{t=0} A(t) = [X, A] - dX = -d_\nabla X$$

This proves that the element of $T_A\mathscr{A}/\mathscr{G}$ lies in $\mathrm{Im}d_\nabla$. The second is that the element of $T_A\mathscr{A}_+$ lies at the kernel of 2nd map. Which is trivial by definition. Therefore our main interest reduces to calculating the first cohomology of the complex.
From here, it highly depends on the original paper of Atiyah-Hitchin-Singer. Witzenböck positivity argument shows that the 2nd cohomology must vanish. For a covariant exterior derivative d_∇,

$$\Delta = d_\nabla^-(d_\nabla^-)^\star = \frac{1}{2}d_\nabla d_\nabla^\star + \frac{R}{6} - W_-$$

R represents the Ricci scalar, and W_- represents the anti self dual part of Weyl tensor. In our setting, this shows that d_∇^- is positive-definite and hence any $\omega \in H^2_-(M,\mathfrak{g})$ is $d_\nabla^\star\omega$ up to positive constant. Which proves that $h_2 = 0$.
The 1st cohomology is the kernel of d_∇. For any lie algebra X,

$$-\left.\frac{d}{dt}\right|_{t=0} A(t) = d_\nabla X$$

And this would give that A is invariant under e^{tX}, which we know that the irreducibility of the bundle shows $X = 0$.
For compact M, it is known that the complex is elliptic, and hence it index equals the topological index by Atiyah-Singer index theorem. To calculate its analytic index, one needs to define the following Dirac operator.

$$D = d_\nabla^\star + d_\nabla^- : \Omega^1(M,\mathfrak{g}) \to \Omega^0(M,\mathfrak{g}) \oplus \Omega^2_-(M,\mathfrak{g})$$

Using the Atiyah-Singer Index theorem,

$$\mathrm{ind}D = \int_M \mathrm{ch}\left(\bigoplus_i (-1)^i E_i\right) \frac{\mathrm{Td}(TM \otimes_\mathbb{R} \mathbb{C})}{\mathrm{e}(TM)}$$

The E_is are given as the following.

$$\begin{aligned} E_0 &= \mathfrak{g} \otimes \Omega^0(M)_\mathbb{C} \\ E_1 &= \mathfrak{g} \otimes \Omega^1(M)_\mathbb{C} \\ E_2 &= \mathfrak{g} \otimes \Omega^2_-(M)_\mathbb{C} \end{aligned}$$

Since the Chern character is splitted by tensoring,

$$\mathrm{ch}\left(\bigoplus_i (-1)^i E_i\right) = \mathrm{ch}(\mathfrak{g})\mathrm{ch}(\Omega^0(M) - \Omega^1(M) + \Omega^2_-(M))_\mathbb{C}$$

The Chern character of $\mathfrak{g}$ would be $\mathrm{dim}G + \frac{1}{2}p_1(\mathfrak{g})$, while the remaining terms would give $b_0 - b_1 + b_2^-$ for the 4 form part. While the zero form part is highly nontrivial, which is known as 2. Through such discussion, the index would be $\mathrm{dim}G(b_0 - b_1 + b_2) + p_1(\mathfrak{g})$. Through some calculation, one can know that it is $p_1(\mathfrak{g}) + \frac{1}{2}\mathrm{dim}G(\chi - \tau)$ $\square$

Note that Atiyah-Hitchin-Singer's original paper uses an equivalent elliptic operator on Spin bundle, through an isomorphism between spin bundle and bundle of differential forms.

Proposition 5. SU(2) *instanton moduli over* S^4 *is* $8k-3$ *dimensional.*

Proof. One knows that $\chi(S^4) = 2$, $\tau(S^4) = 0$. The pontryagin class for $\mathfrak{su}(2)$ can be constructed from the product bundle, $E \otimes E$. Where E is the 2 dimesional representation of SU(2), vector bundle of $c_2(E) = -k$. Therefore one can consider that $E \otimes E = S^2 E \oplus \wedge^2 E$. The former represents the symmetric part, while the latter represents the antisymmetric part. The antisymmetric part is of rank 2 choose 2, therefore a line bundle. While the symmetric part is of a rank 3. Taking the chern character would give that,

$$(2 + c_2(E))^2 = 3 + \frac{1}{2} p_1(\mathfrak{su}(2)) + 1$$

Considering only 4 dimensional part, one would have that $p_1(\mathfrak{su}(2)) = 8k$ and hence the total dimension is $8k - 3$. □

BPST Instanton

There exists a method to construct instanton in $\mathbb{R}^4$. One can have the ansatz

$$A_\mu(x) = \alpha \sigma_{\mu\nu} \partial_\nu \log \phi(x^2).$$

While $\sigma_{\mu\nu}$ are generators for rotation in $\mathbb{R}^4$,satisfying

$$[\sigma_{\mu\nu}, \sigma_{\rho\sigma}] = -2 \left(\delta_{\mu\rho}\sigma_{\nu\sigma} + \delta_{\nu\sigma}\sigma_{\mu\rho} - \delta_{\mu\sigma}\sigma_{\nu\rho} - \delta_{\nu\rho}\sigma_{\mu\sigma}\right)$$

Then the field strength and dual field strength are:

$$F_{\mu\nu} = \alpha\sigma_{\nu\rho}\partial_\mu\partial_\rho \log\phi - \alpha\sigma_{\mu\rho}\partial_\nu\partial_\rho \log\phi^2 + \alpha^2[\sigma_{\mu\rho}, \sigma_{\nu\sigma}](\partial_\rho \log\phi)(\partial_\sigma \log\phi)$$

$$= (\alpha\sigma_{\nu\rho}\partial_\mu\partial_\rho \log\phi - (\mu \leftrightarrow \nu)) + 2\alpha^2 \left(\sigma_{\mu\sigma}(\partial_\nu \log\phi)(\partial_\sigma \log\phi) - (\mu \leftrightarrow \nu)\right) - 2\alpha^2\sigma_{\mu\nu}(\partial \log\phi)^2;$$

$$*F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}$$

$$= \alpha\epsilon_{\mu\nu\rho\sigma}\sigma_{\nu\lambda}\partial_\rho\partial_\lambda \log\phi + 2\alpha^2\epsilon_{\mu\nu\rho\sigma}\sigma_{\rho\lambda}(\partial_\sigma \log\phi)(\partial_\lambda \log\phi) - \alpha^2\epsilon_{\mu\nu\rho\sigma}\sigma_{\rho\sigma}(\partial \log\phi)^2$$

$$= \sigma_{\nu\rho}\left(\alpha\partial_\rho\partial_\mu \log\phi - 2\alpha^2(\partial_\rho \log\phi)(\partial_\mu \log\phi) - (\mu \leftrightarrow \nu)\right) + \sigma_{\mu\nu}(\alpha\partial^2 \log\phi).$$

To equip self duality, $-2\alpha^2(\partial \text{log}\phi)^2 = \alpha\partial^2\phi$ needs to be satisfied. Substituting $\phi \to \phi^{1/2\alpha}$, one can get that $\phi^{-1}\partial^2\phi = 0$, which is just laplace equation.
For $\phi = \frac{\rho^2}{(x-a)^2} + C$, C needs to be 1 for a condition that sufficiently large x gives 0 on A.

$$A_\mu(x) = -\sigma_{\mu\nu}\frac{\rho^2(x-a)_\nu}{(x-a)^2((x-a)^2+\rho^2)}$$

Under $U = \frac{i\sigma_i x_i}{|x|}$,

$$U(\partial_\mu + A_\mu)U^{-1} = -\bar{\sigma}_{\mu\nu}\frac{(x-a)_\nu}{((x-a)^2+\rho^2)}$$

The field strength would be followed.

$$F_{\mu\nu} = \partial_\mu\left[-\bar{\sigma}_{\nu\rho}\frac{(x-a)_\nu}{(x-a)^2+\rho^2}\right] - (\mu \leftrightarrow \nu) + \frac{(x-a)_\rho(x-a)_\sigma}{[(x-a)^2+\rho^2]^2}[\bar{\sigma}_{\mu\rho}, \bar{\sigma}_{\nu\sigma}]$$

$$= \left[\frac{\bar{\sigma}_{\mu\nu}}{(x-a)^2+\rho^2} + \frac{2(x-a)_\mu(x-a)_\rho}{[(x-a)^2+\rho^2]^2}\bar{\sigma}_{\nu\rho}\right] - (\mu \leftrightarrow \nu)$$

$$-\frac{2}{[(x-a)^2+\rho^2]^2}\left((x-a)^2\bar{\sigma}_{\mu\nu} - (x-a)_\rho(x-a)_\mu\bar{\sigma}_{\rho\nu} - (x-a)_\nu(x-a)_\rho\bar{\sigma}_{\mu\sigma}\right)$$

$$= \frac{2\rho^2\bar{\sigma}_{\mu\nu}}{[(x-a)^2+\rho^2]^2}$$

each a corresponds to the translational transformation of instanton, with ρ a sizing transformation. These are the conformal transformations, making the instanton valid. The winding number(i.e. the 2nd chern class integrated)

$$k = -\frac{1}{16\pi^2}\int \mathrm{d}^4x \,\mathrm{tr}\, F^*_{\mu\nu}F_{\mu\nu} = -\frac{1}{16\pi^2}\int \mathrm{d}^4x \frac{4\rho^4}{\left[(x-a)^2+\rho^2\right]^4}\,\mathrm{tr}\,\bar{\sigma}_{\mu\nu}\bar{\sigma}_{\mu\nu} = 1$$

These matches well with our observations. Since the dimension of instanton moduli, SU(2) bundle over S^4 is equivalent for instanton moduli over $\mathbb{R}^4$, vanishing at infinity. And such degree of freedom is realized as 5 conformal transformations. While $8k$ is physically intuitive, considering the symmetries of SU(2), but 3 degrees of freedom are dropped for the gauge transformation.
Therefore, if one doesn't consider the vanishing gauge field, it is possible to consider the gauge configuration on 2 half sphere of S^4, intersecting at S^3. This would give the matching condition as in the magnetic monopole construction.
As a persepective of pure gauge story, the only information about the instanton at infinities are the winding numbers. Which needs to be true, since the action functional $\mathrm{tr}(F\wedge F)$ is locally exact, as $d\mathrm{tr}(AdA+\frac{2}{3}A^3)$. Meaning that for S^4 configuration, the stokes theorem gives the information on boundary, as the Chern Simons form.
With our former consideration on characteristic class, it is valid to consider the $[S^4, \mathrm{BSU}(2)] = \pi_4(BSU(2)) \cong \mathbb{Z}$. While this $\mathbb{Z}$ represents the winding number. In a monoidal perspective, gauge configuration can be multiplied. And such information is in the 2nd chern class. Physically,

$$J^\mu = \frac{1}{24\pi^2}\mathrm{tr}(\partial_\nu g g^{-1}\partial_\rho g g^{-1}\partial_\sigma g^{-1})$$

Is the topological current, counting such instanton numbers. Which can be interpreted in a pure gauge persepective,

$$J \approx \mathrm{tr}(A^3)$$

Which is asymptotically a Chern Simons term on boundary.

4 Effects of Instantons in real world physics

Now, let's return to the physics. In the toy model case, the instanton effect was essential to calculate the tunneling effect. Though we did not mention it, the instanton possibly affects any other term, which is not perturbative in the coupling constant, or $\hbar$. In the real world, the instanton effect is crucial to understand the vacuum structure of QCD. First, we study the classical vaccume structure of YM theory.

4.1 Vaccume structure of YM theory

4.2 Easy analogy: QM on a circle

4.3 The θ-term

4.4 The θ-term in QCD

4.5 Chiral Anomaly and θ-vacuum

5 Instanton effects in Axion phtsics

6 References