Notes on Instantons

Leonardo Almeida Lessa

November 2019

Contents

We introduce the concept of instantons first in Quantum Mechanics (Sec. 1) and then in Quantum Field Theory (Sec. 2). Similarly to the WKB approximation, the technique behind instantons involves a small-$\hbar$ nonperturbative approximation. It also connects the topology of spacetime to the the vacuum structure of quantum theories (Sec. 2.3).

Of the many applications of instantons, we discuss the confinement effect in the Abelian Higgs Model in 1+1 dimensions (Sec. 3.1) and the 't Hooft solution of the U(1) problem (Sec. 3.2).

These notes are based mostly on Chapter 3 of Sidney Coleman's "Aspects of Symmetry" [1]. Other references include Chapter 3 of "Condensed Matter Field Theory", by A. Altland and B. Simon [2],

Warning: Euclidean Spaces

As we will work almost solely within Euclidean Space, we shall remind ourselves of the necessary changes:

$$\begin{aligned} x^4 &= ix^0, \\ t &= i\tau, \\ S_E &= -iS. \end{aligned}$$

If, for example, the relation between the action S of a massive particle in Minkowski space

$$S = \int_{\mathbb{R}} \mathrm{d}\tau \left[\frac{m}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}\tau} \right)^2 - V \right],$$

and the corresponding Euclidean action S_E is the following

$$\begin{aligned} S &= -i \int_{i\mathbb{R}} \mathrm{d}t \left[-\frac{m}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)^2 - V \right], \\ &= i \int_{i\mathbb{R}} \mathrm{d}t \left[\frac{m}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)^2 + V \right], \\ &= iS_E. \end{aligned}$$

Note that the kinetic and potential terms of S_E have the same plus sign.

> ! Since the following discussion uses only Euclidean spaces, **we will drop the subscript** E from Euclidean quantities.

1 Instantons in Quantum Mechanics

We will first introduce the instantons in the context of Nonrelativistic Quantum Mechanics (NRQM), which will form the basis for the calculations in QFT. In all cases, the definition of instantons is

Instantons are (nontrivial) solutions $\overline{\phi}$ of the classical equations of motion of an action S, i.e.

$$\frac{\delta S}{\delta \phi}[\overline{\phi}] = 0.$$

More that that, the main idea behind instantons is that we can have a nonperturbative results arising from a small $\hbar$ (semiclassical) approximation. As we shall see, this amounts to nontrivial solutions of the equation of motion, around which we will calculate quantum perturbations.

Nonperturbative results in a small $\hbar$ approximation may seem strange, but there are known examples of this behavior in ordinary NRQM, such as the WKB approximation. It is a small $\hbar$ approximation that predicts effects that don't appear in any order in perturbation theory, such as tunneling probabilities.

1.1 Saddle Point Approximation

Suppose a field action S has a stationary point $\overline{\phi}$, which, by definition, satisfies the equations of motion of S. In this case, we can do a Taylor expansion around $\overline{\phi}$ to calculate the partition function:

$$\begin{aligned} \mathcal{Z} &= N \int \mathcal{D}\phi \ \exp\{-S[\phi]/\hbar\}, \\ &\approx N \int \mathcal{D}\phi \ \exp\left\{ -S[\overline{\phi}]/\hbar - \frac{1}{2} \int \mathrm{d}^4x \, \mathrm{d}^4y \ \Delta\phi(x) \frac{1}{\hbar} \frac{\delta^2 S[\overline{\phi}]}{\delta\phi(x)\delta\phi(y)} \Delta\phi(y) \right\}, \\ &= N e^{-S[\overline{\phi}]/\hbar} \det\left(\frac{1}{2\pi\hbar} \frac{\delta^2 S[\overline{\phi}]}{\delta\phi(x)\delta\phi(y)} \right)^{-1/2} [1 + \mathcal{O}(\hbar)], \end{aligned}$$

where $\Delta\phi = \phi - \overline{\phi}$ and N is a normalization factor. This formula also works for NRQM action with substitutions $x \to t, y \to t', \phi \to x^1$. In the case of $S = \int \mathrm{d}t \left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + V(x(t))\right]$, the path integral becomes

$$\mathcal{Z} = \tilde{N} e^{-S[\overline{x}]/\hbar} \det\left[m(-\partial_t^2 + \omega^2)\right]^{-1/2}[1 + \mathcal{O}(\hbar)],$$

where $m\omega^2 = V''(\overline{x}(t))$ and the operator $m(-\partial_t^2 + \omega^2)$ is to be viewed as a diagonal infinite matrix in t and t'. If S has many stationary points, we sum over them.

> ! The stationary point condition for the NRQM action is $m\ddot{x} = V'(x)$, Newton's Law with **inverted potential** $-V$.

The partition function will be useful for analyzing the states of lowest energies, since $\langle x_f | e^{-HT/\hbar} | x_i \rangle = \mathcal{Z}$ with boundary conditions $x(0) = x_i$ and $x(T) = x_f$. If $H|n\rangle = E_n |n\rangle$, with E_0 being the energy of the ground state,

$$\begin{aligned} G(x_f, x_i; T) := \langle x_f | e^{-HT/\hbar} | x_i \rangle &= \sum_n e^{-E_n T/\hbar} \langle x_f | n \rangle \langle n | x_i \rangle, \\ &\overset{T\to\infty}{=} e^{-E_0 T/\hbar} \langle x_f | 0 \rangle \langle 0 | x_i \rangle, \end{aligned} \tag{1.1}$$

so the ground state mode dominates the behavior of $\mathcal{Z}$ in the limit $T \to +\infty$.

1.2 Quadratic Potential

To test our saddle point approximation, we will calculate the propagator

$$G(0, 0; T) = \langle x_f = 0 | e^{-HT/\hbar} | x_i = 0 \rangle$$

for the locally quadratic potential $V(x) \sim \frac{1}{2}\omega^2 x^2$ (see Figure 1). The stationary points of the action are classical solutions for the inverted potential $-V$ (dashed curve) with $x(0) = x(T) = 0$. For a quadratic potential, only $\overline{x} = 0$ satisfy these requirements, for which $S[\overline{x}] = 0$.

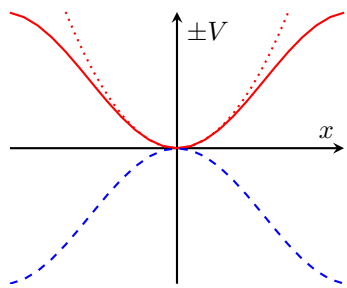


Figure 1: Quadratic potential

Now, it remains to calculate the operator determinant. The eigenvectors of the $-\partial_t^2 + \omega^2$ with boundary conditions $x(0) = x(T) = 0$ are $r_n(t) = \sin(\pi n t/T)$, $n = 1, 2, \ldots$, with eigenvalues $\epsilon_n = (\pi n/T)^2 + \omega^2$. Thus,

$$\det\left(-\partial_t^2 + \omega^2\right)^{-1/2} = \prod_{n=1}^{\infty} \left[\left(\frac{\pi n}{T}\right)^2 + \omega^2\right]^{-1/2},$$

which appears in the propagator as $G(0, 0; T) = J \det\left(-\partial_t^2 + \omega^2\right)^{-1/2}$. To calculate the prefactor J, we use that $G(0, 0; T)$ should reduce to the free propagator

$$G_{\text{free}}(x_f, x_i; T) = (2\pi\hbar T)^{-1/2} \exp\left\{-\frac{(x_f - x_i)^2}{2\hbar T}\right\},$$

in the limit $\omega \to 0$. Thus,

$$\begin{aligned} G(0, 0; T) &= \frac{G(0, 0; T)}{G(0, 0; T)|_{\omega=0}} G_{\text{free}}(0, 0; T), \\ &= \prod_{n=1}^{\infty} \left[1 + \left(\frac{T\omega}{\pi n}\right)\right]^{-1/2} (2\pi\hbar T)^{-1/2}, \\ &= \sqrt{\frac{\omega}{2\pi\hbar \sinh \omega T}} \overset{T\to\infty}{\longrightarrow} \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2}, \end{aligned} \tag{1.2}$$

[1] NRQM is just a 0 space + 1 time = 1 dimensional field theory.

where we used that $\prod_{n\geq 1}[1+(x/n\pi)]^{-1} = x/\sinh x$. In the limit $T\to\infty$, we can use (1.1) and (1.2) to conclude that $E_0 = \frac{1}{2}\omega\hbar[1+\mathcal{O}(\hbar)]$ and $|\langle x=0|n=0\rangle|^2 = (\omega/\pi\hbar)^{1/2}[1+\mathcal{O}(\hbar)]$, which reproduces the semiclassical results.

> ! From now on, we **will drop the** $O(\hbar)$ when it is clear from the context that we are approximating at first order in $\hbar$.

1.3 Double Well

Now, let's consider a double well potential (see Figure 2) with minima in $\pm a$. We can investigate the propagators $G(\pm a, \mp a; T)$, between $-a$ and a, and between themselves $G(\pm a, \pm a; T)$. Looking at the inverted potential (dashed curve), we can see that there are classical solutions that start at $x=-a$ and go to $x=a$. Their energy is zero, so they satisfy $d\bar{x}/dx = \sqrt{2V}$ and the action for them is $S_0 = \int \mathrm{d}t\left[\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + V\right] = \int_{-a}^{a}\mathrm{d}x\sqrt{2V}$.

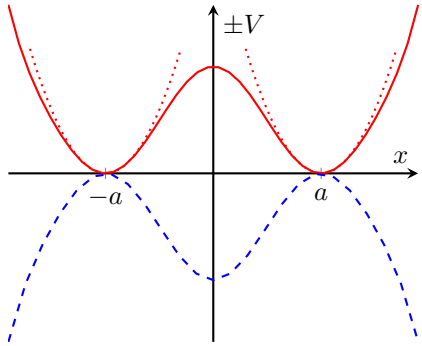


Figure 2: Double well potential

Moreover, these solutions are non trivial stationary points of the action and are called **instantons** (see Figure 3). The inverse solution is called an anti-instanton and we can combine any number of them centered at different times to construct a multi-instanton solution.

The dynamical part of an instanton is localized within a time $\sim\omega^{-1}$. Thus, taking $T\to\infty$, we can assume the transitions between minima are instantaneous and infinitely far apart for a *multi-instanton solution* (see Figure 4). Therefore, we can treat them as independent instantons.

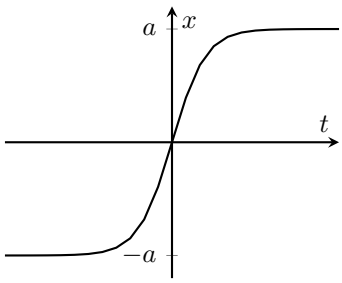


Figure 3: Instanton solution

According to the instanton gas approximation, we need to sum over all of configurations that take a certain minimum point to another one, and then integrate over all their centers. The result of this for the propagator is

$$G(-a,-a;T) = \sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}\sum_{\text{even } n}\frac{T^n}{n!}(Ke^{-S_0/\hbar})^n,$$

where the external factor $\sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}$ comes from the quadratic potential propagator; K is a correction term[2]; $T^n/n!$ comes from the integration over all instantons' centers; and the sum over even n guarantees that the system comes back to the initial state. Analogously,

$$G(\pm a,-a;T) = \frac{1}{2}\sqrt{\frac{\omega}{\pi\hbar}}e^{-\omega T/2}\left[\exp\left(Ke^{-S_0/\hbar}T\right)\mp\exp\left(-Ke^{-S_0/\hbar}T\right)\right]. \tag{1.3}$$

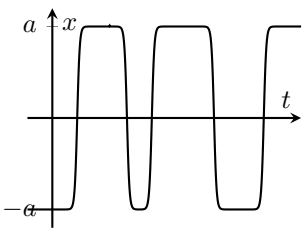


Figure 4: Multi-instanton

Comparing equations (1.1) and (1.3), we see not only the ground state contributing to the propagator, but *two* low energy states, with energies $E_\pm = \frac{1}{2}\hbar\omega \pm \hbar K e^{-S_0/\hbar}$, separated by an exponentially small factor in $\hbar$, the barrier penetration factor, characteristic of an instanton solution. These states are exactly the symmetric and antisymmetric combinations of the ground states of the harmonic oscillators.

Observation: The $\lambda\phi^4$ model with negative mass $\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4$ exhibits spontaneous symmetry breaking (SSB) of the $\phi\to-\phi$ symmetry. If we discretize spacetime and ignore the spatial derivative term, there would be a double well potential for each spacetime point. According to our analysis, the energy difference between the lowest energy states would be $\sim e^{-VS_0/\hbar}$, where V is the volume of the space. Taking $V\to\infty$, the ground state becomes two-fold degenerate, as expected from SSB.

[2] K corrects the fact that we need to calculate the operator determinant around an instanton solution, not just around the minima. For an explicit calculation, see [1].

1.4 Periodic Potential

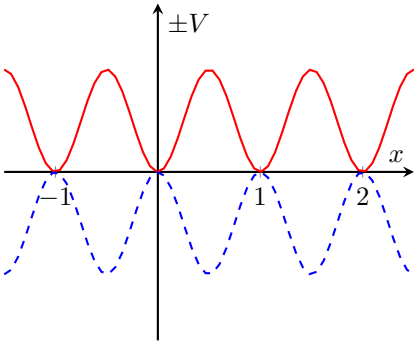


Figure 5: Periodic potential

For our last NRQM example, we will consider a periodic potential (see Figure 5). Just like the double well potential, the (anti-)instantons can jump from a minimum at $x = j \in \mathbb{Z}$ to an adjacent one $x = j \pm 1$, but now there is no constraint on x (See Figure 6). Repeating the same reasoning as before, we can find the propagator between x_i and x_f with $x_f - x_i = k \in \mathbb{Z}$ by summing over configuration with n instantons and $\overline{n}$ anti-instantons, subject to $n - \overline{n} = k$:

$$G(x_f - x_i = k; T) = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \sum_{n,\overline{n}} \frac{T^{n+\overline{n}}}{n!\overline{n}!} \times (K e^{-S_0/\hbar})^{n+\overline{n}} \delta_{n-\overline{n},k}.$$

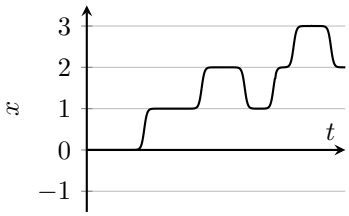


Figure 6: Multi-instanton

Using that $\delta_{ab} = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{i\theta(a-b)}$, we arrive at

$$G(k; T) = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{ik\theta} \exp\left[2KT\cos\theta e^{-S_0/\hbar}\right],$$

which corresponds to a continuum of low energy states $|\theta\rangle$ for $\theta \in [0, 2\pi)$ with energy $E(\theta) = \frac{1}{2}\hbar\omega + 2\hbar K \cos\theta e^{-S_0/\hbar}$ and wave function $\langle\theta|j\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} (2\pi)^{-1/2} e^{ij\theta}$, exactly what is predicted by Bloch waves.

2 Instantons in Gauge Field Theories

We will now see how instantons can affect the vacuum structure of Quantum Field Theories, in special Gauge Field Theories. But first we need to make Coleman's notation explicit.

Notation

For a gauge group G with Lie algebra $\mathfrak{g}$, we can choose a basis $\{T^a\} \subset \mathfrak{g}$ such that, in some representation R, $\mathrm{Tr}(T^a T^b) = C(R)\delta^{ab}$. It is also useful to define the Lie-algebra-valued forms $\mathcal{A} = A_\mu \,\mathrm{d}x^\mu$, *which is i times the usual vector potential*, and $\mathcal{F} = \frac{1}{2} F_{\mu\nu} \,\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$ with components

$$A_\mu = g A^a_\mu T^a$$
$$F_{\mu\nu} = g F^a_{\mu\nu} T^a = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

and the Cartan inner product $(x, y) = \frac{1}{C(R)} \mathrm{Tr}(xy)$, $x, y \in \mathfrak{g}$, so that the Euclidean gauge action becomes just $S = \frac{1}{4g^2} \int \mathrm{d}^4x (F, F)$, with $(F, F) = (F_{\mu\nu}, F_{\mu\nu})$. For example, for $G = SU(2)$, we have $T^a = \frac{1}{2i}\sigma^a$ and $C(\mathrm{fund}) = -\frac{1}{2}$, by convention. This is such that A^a_μ and $F^a_{\mu\nu}$ are real numbers.

2.1 A Minimum of Homotopy Groups

Two continuous maps $f, g : S \to T$ are said to be *homotopic* if there exists a continuous function $H : [0, 1] \times S \to T$ such that $H(0, s) = f(s)$ and $H(1, s) = g(s)$, for any $s \in S$. Intuitively, f can be continuously deformed into T. If $S = S^1$, or, equivalently, $S = [0, 1]$ with endpoints identified, then f and g are loop curves in the topological space T and the homotopic property defines an equivalence relation between them. Moreover, the space of all loops based in a point $p = f(0) = f(1) \in T$ has a group structure defined by adjoining them as (see Figure 7)

$$(f \vee g)(s) = \begin{cases} f(2s), & \text{if } 0 < s < 1/2 \\ g(2s-1), & \text{if } 1/2 < s < 1 \end{cases}.$$

Turns out[3] that the space of all equivalence classes under homotopy is a group under this operation as well, called the **fundamental group** $\pi_1(T)$. For example, $\pi_1(S^1) \cong \mathbb{Z}$, whose curves are indexed by the number of turns within S^1, called the **winding number**. This can be generalized to maps $f : S^n \to T$, and the corresponding group is $\pi_n(T)$, called the **n-th homotopy group**. For more information, we recommend Nakahara's book [4].

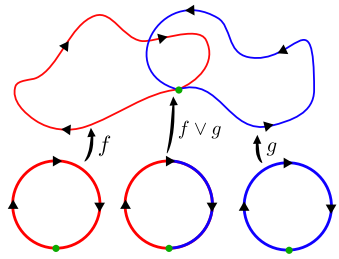


Figure 7: Adjoining Curves

2.2 Topological Solutions and Winding Number

A common feature of instantons that was displayed in our previous examples is that they are nontrivial solution that connect vacua in the border of the field domain. In the case of the double well, for example, the domain border is $\{-T, +T\}$ and an instanton connects the vacuum state $x = -a$ at $t = -T$ to the other vacuum $x = a$ at $t = +T$. In QFT, this is no different, but with the additional complication that the domain boundary is topologically a S^{d-1}, for a d-dimensional QFT.

Additionally, we will study *finite action* stationary field configurations, since an infinite action would not contribute to the integral because of the term $e^{-S[\overline{\phi}]}$ in the path integral. For a pure euclidean gauge theory, this limits the asymptotic behavior of the fields, since

$$\int \mathrm{d}^4x \; F^a_{\mu\nu}F^a_{\mu\nu} < \infty \;\Rightarrow\; F^a_{\mu\nu} = \mathcal{O}(r^{-3}) \;\overset{?}{\Rightarrow}\; A^a_\mu = \mathcal{O}(r^{-2})$$

where r is the radius of the Euclidean $M = \mathbb{R}^4$ space.

This analysis seems fine but we need to remember that we can also have vanishing $F^a_{\mu\nu}$ if $A_\mu = g\partial_\mu g^{-1} + \mathcal{O}(r^{-2})$, for a gauge transformation $g : S^3 \to G$ that only depends on angular coordinates. Moreover, we can change g to any other homotopic gauge map $\tilde{g} : S^3 \to G$ by doing a globally defined gauge transformation on A_μ, but not to non-homotopic ones. This has a profound consequence:

For each $g \in \pi_3(G)$ in the third homotopy group of the gauge group G,
we get a distinct asymptotic expression[4] for the potential $A_\mu = g\partial_\mu g^{-1} + \mathcal{O}(r^{-2})$.

Turns out that we need only to consider $G = SU(2)$ in four dimensions[5], for which $\pi_3(SU(2)) \cong \pi_3(S^3) \cong \mathbb{Z}$. As this relation suggests, $SU(2)$ is topologically related to S^3. Indeed, any element g of $SU(2)$ is of the form $g = a + i\mathbf{b}\cdot\sigma$, where $(a, \mathbf{b}) \in S^3$. It can also be shown that the representatives of $\pi_3(SU(2))$ in Table 1 exhaust the group. Consequently, the winding number $\nu \in \mathbb{Z}$ classifies the asymptotic behavior of A_μ and is gauge invariant.

Table 1: Elements of $\pi_3(SU(2))$

Winding Number ν	Group Map g_ν
0	1
1	$x_4 + i\mathbf{x}\cdot\sigma$
ν	$[g_1(x)]^\nu$

Defining the Hodge dual of $F_{\mu\nu}$ by $(*F)_{\mu\nu} := \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}$, we can calculate the winding number of A_μ via the integral

$$\int_M \mathrm{d}^4x \; (F, *F) = \int_{S^3} \mathrm{d}^3S_\mu G^\mu = 32\pi^2\nu, \tag{2.1}$$

where $(F, *F) = (F_{\mu\nu}, *F_{\mu\nu})$ and $G_\mu = \epsilon_{\mu\nu\rho\sigma}(A_\nu, F_{\rho\sigma} - \frac{2}{3}A_\rho A_\sigma)$ [6].

2.3 Many Vacua of Gauge Theories

As we have seen earlier, the asymptotic behavior of the fields is reflected on the vacua, so we expect the winding number to somehow classify low energy states of the gauge theory. To see this,

[3] See any book on fundamental groups or algebraic topology, for example [3].

[4] Equivalently, we may compactify $\mathbb{R}^4 \to S^4$ by adding a point ∞ at the infinity. In this context, "asymptotic" is replaced by locally around ∞.

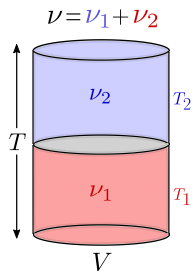
[5] See page 291 of Coleman's book [1].

[6] This result comes from an area of profound connection between geometry and physics called Characteristic Classes. The interested reader is referred to Chapter 11 of Nakahara [4] and, for a simpler proof, to Page 289 of Coleman [1].

we will work with QFT in a spacetime box of space volume V and time range T, so boundary conditions that persist in the limit $V, T \to \infty$ are made explicit.

For a variety of reasons, we will adopt the axial gauge A_3, one of them being the lack of ghosts. Moreover, we fix the tangential components of A_μ at the boundary, so the variation of the action $\delta S = \frac{1}{g^2}\int \mathrm{d}^3 S^\mu F_{\mu\nu}\delta A^\nu + \cdots$ is zero. In particular, the field at the boundary must be of fixed winding number.

Given this, we now proceed to calculate the path integral $F(V,T,\nu) := N\int \mathcal{D}A^{(\nu)}\, e^{-S}$, over fields $A_\mu^{(\nu)}$ of winding number ν [7]. At large T, we can separate our domain into two parts of duration T_1 and T_2, with $T_1 + T_2 = T$ and neglect effects that could come from the boundary between them. Because of (2.1), the winding numbers of these parts are added, so we sum over all combinations that give a total winding number ν



$$F(V, T_1 + T_2, \nu) = \sum_{\nu_1+\nu_2=\nu} F(V, T_1, \nu_1) F(V, T_2, \nu_2). \tag{2.2}$$

Unfortunately, the terms $F(V,T,\nu)$ do not describe an eigenstate of the Hamiltonian, since that would give a simple multiplication law, instead of (2.2). However, we can correct this via a Fourier transform:

$$F(V,T,\theta) := \sum_{\nu\in\mathbb{Z}} e^{i\nu\theta} F(V,T,\nu) = N\sum_{\nu\in\mathbb{Z}} \int \mathcal{D}A^{(\nu)}\, e^{-S} e^{i\nu\theta}, \quad \theta \in [0, 2\pi),$$

which satisfies $F(V, T_1 + T_2, \theta) = F(V, T_1, \theta) F(V, T_2, \theta)$. Assuming strong continuity of $F(V,T,\theta)$ with T and using (1.1), there exists an energy eigenstate $|\theta\rangle$ such that $F(V,T,\theta) \propto \langle\theta| e^{-HT} |\theta\rangle$. These $|\theta\rangle$ states, called **theta vacua**, are similar to the corresponding ones of the periodic potential example, as the winding number is analogous to a total change in x (see Section 1.4). The analogy goes further if we define the $n-$states by $|n\rangle = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} e^{in\theta} |\theta\rangle$, such that $F(V,T,\nu)$ is a probability amplitude to go from $|n\rangle$ to $|n+\nu\rangle$, since

$$\langle n+\nu| e^{-HT} |n\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\nu\theta} \langle\theta| e^{-HT} |\theta\rangle \propto F(V,T,\nu).$$

The path integral corresponding to $F(V,T,\theta)$ can be seen as the usual partition function of an effective action

$$S_{\text{eff}} = S - \frac{i\theta}{64\pi^2}\int \mathrm{d}^4x\, \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}. \tag{2.3}$$

Interestingly, even though we know from equation (2.1) that this extra term is a divergence, it has physical consequences, such as choosing the vacuum of the gauge theory. Furthermore, perturbation theory is connected only to the $\nu = 0$ gauge fields, so the θ vacua is within the nonperturbative regime [5].

2.4 General Features of Instantons

Keep in mind that we have only showed the *existence* of such instantons and vacua but not their exact expression. Surprisingly, many features of Instantons can be explored without constructing them explicitly.

Imitating what we did in the periodic potential case, we can calculate $F(V,T,\theta)$ considering a diluted instanton gas of n instantons of $\nu = 1$ and $\overline{n}$ anti-instantons of $\nu = -1$:

$$\langle\theta| e^{-HT} |\theta\rangle \propto \sum_{n,\overline{n}} (Ke^{-S_0})^{n+\overline{n}} \frac{(VT)^{n+\overline{n}}}{n!\overline{n}!} e^{i(n-\overline{n})\theta}, \tag{2.4}$$

$$= \exp\big(2Ke^{-S_0} VT\cos(\theta)\big), \tag{2.5}$$

[7] We have set $\hbar = 1$. We can still count the power of $\hbar$ via the powers of g^2, since they always come together in our notation.

where VT comes from an integration over all possible centers of the instantons. By (1.1), we conclude that a θ vacuum has an energy density of

$$\mathcal{E}(\theta) = -2K\cos(\theta)e^{-S_0}. \tag{2.6}$$

This result confirms that the θ vacua are all different from each other and that their energy density is small, because of the factor $e^{-S_0/\hbar}$.

Using (2.2), we can also see that the field strength has a non-zero expectation value:

$$\begin{aligned}\langle\theta|\,(F,*F)\,|\theta\rangle &= \frac{1}{VT}\langle\theta|\left[\int \mathrm{d}^4x\ (F,*F)\right]|\theta\rangle\,,\\ &= -\frac{32\pi^2 i}{VT}\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\left(\int \mathcal{D}A^{(\nu)}\ e^{-S}e^{i\nu\theta}\right),\\ &= -64\pi^2 iKe^{-S_0}\sin(\theta),\end{aligned} \tag{2.7}$$

where we used translational invariance of the θ vacuum in the first line. In the abelian case, we have $(F,*F)\propto \mathbf{E}\cdot\mathbf{B}$, so the EM fields are non-zero in the θ vacuum, for $\theta\neq 0$.

2.5 Constructing the Instanton

We know that instantons have a definite winding number, but we lack an explicit expression for them. For this, we could solve the classical equation of motion for the pure gauge action, which is $D_\mu F_{\mu\nu} = \partial_\mu F_{\mu\nu} + [A_\mu, F_{\mu\nu}] = 0$. However, this is a complicated second-order differential equation in A_μ. Fortunately, we can simplify our problem by realizing that, by Cauchy-Schwartz inequality in $L^2(\mathbb{R}^4)$,

$$\int \mathrm{d}^4x\ (F,F) = \sqrt{\left(\int \mathrm{d}^4x\ (F,F)\right)\left(\int \mathrm{d}^4x\ (*F,*F)\right)} \geq \left|\int \mathrm{d}^4x\ (F,*F)\right|,$$

thus, by (2.2), $S\geq \frac{8\pi^2}{g^2}|\nu|$. The equality is attained iff $*F=\lambda F$, $\lambda\in\mathbb{R}$, but using that $**F=F$, only $\lambda=1$ (self-dual, $\nu>0$) or $\lambda=-1$ (anti-self-dual, $\nu<0$) solutions are present. In either case, the equation of motion is automatically satisfied[8], we get a minimum action solution for a given winding number *and* $F=\pm *F$ is an easier first-order differential equation.

Using rotational invariance, the solution of $F=*F$ for $\nu=1$ is

$$A_\mu = f(r)g_1\partial_\mu g_1^{-1}, \quad \text{with } f(r)=\frac{r^2}{r^2+\rho^2}, \tag{2.8}$$

where $\rho>0$ is the "size of the instanton" that comes from scale invariance of the theory (See Figure 8). Since we take $V,T\to+\infty$, we can see from (2.8) that the instanton in localized in space and in time.

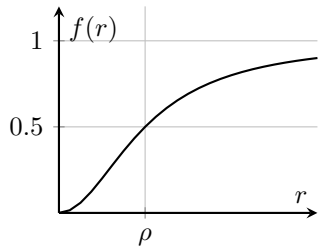


Figure 8: Radial part $f(r)$

All other solutions for $\nu=1$ are conformal and gauge transformations of this one. Furthermore, we don't need solutions with higher winding number. First because we can always imitate one with many $\nu=\pm1$ instantons far apart, and second because they would give a minute contribution to the energy density (2.6), proportional to $e^{-\nu S_0}$.

From now on, we will discuss applications of instantons. The only detail left over was the calculation of K, which involves renormalization group and infrared divergences. Hence, to not lose focus, the interested reader is directed to section 7.3.6 of Coleman [1].

[8] Since $D_\mu(*F)_{\mu\nu}=0$, a non-abelian version of the homogeneous Maxwell's equations, called the Bianchi identity (see Section 10.3.5 of Nakahara [4]).

3 Applications of Instantons

3.1 Abelian Higgs Model in 1+1 Dimensions

Usually, a spontaneous symmetry breaking model gives rise to a massive field and a Goldstone boson. If we introduce gauge symmetry, the massless gauge vector "absorbs" the Goldstone boson and becomes massive, in turn *reducing the interaction range* by an exponentially decaying factor in the propagators. I will now show that instantons reverse this effect in the abelian Higgs model in 1+1 dimensions by creating an effective Coulomb force between charges and hindering the Higgs phenomenon.

The Lagrangian for this model is of $U(1)$ gauge group with Higgs coupling,

$$\mathcal{L} = \frac{1}{4e^2}(F,F) + D_\mu\psi^* D_\mu\psi + \frac{\lambda}{4}(\psi^*\psi)^2 + \frac{\mu^2}{2}\psi^*\psi,$$

with $\mu^2 < 0$, such that the ψ potential has a Mexican hat shape, which allows for symmetry breaking.

If we repeat the same construction done in Section 2, we find that the instantons of this model are classified by the homotopy group $\pi_1(U(1)) \cong \mathbb{Z}$ via

$$\nu = \frac{i}{4\pi}\int_M \mathrm{d}^2x\ \varepsilon_{\mu\nu}F_{\mu\nu}. \tag{3.1}$$

The energy density is the same as in (2.6) and the analogous of (2.7) is $\langle\theta|\, E\, |\theta\rangle = -4\pi K e^{-S_0}\sin\theta$, with $E = F_{21}$ being the electric field.

To analyze the interaction between charges, we will calculate the energy difference Δ due to the presence of two static charges $\pm q$ a distance L' apart (See Figure 9). First, we define R to be the spacetime region enclosed by the charges' world lines in a time period T'.

Following Wilson [6], we claim that

$$\Delta = -\lim_{T'\to\infty}\frac{1}{T'}\ln\langle\theta|\, W\, |\theta\rangle,$$

where

$$W = \exp\left\{-\frac{q}{e}\oint_{\partial R} A_\mu \mathrm{d}x_\mu\right\} = \exp\left\{-\frac{q}{2e}\oint_R F_{\mu\nu}\varepsilon_{\mu\nu}\mathrm{d}^2x\right\},$$

and

$$\langle\theta|\, W\, |\theta\rangle = \frac{\sum_{\nu\in\mathbb{Z}}\int \mathcal{D}A^{(\nu)}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ We^{-S}e^{i\nu\theta}}{\sum_{\nu\in\mathbb{Z}}\int \mathcal{D}A^{(\nu)}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ e^{-S}e^{i\nu\theta}}$$

My semiclassical explanation for this is that Δ is a potential difference between the charges, since $-\frac{1}{T'}\ln W \approx -\frac{q}{e}\int_0^{L'} E(x)\mathrm{d}x$ [9].

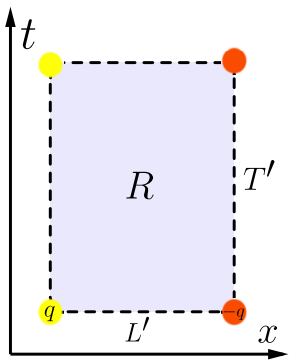


Figure 9: Wilson Line

If we take L' and T' large enough, but still much smaller than L and T, we can suppose there is an instanton solution of winding number ν_I inside R and another of winding number ν_O outside R. As such, we can use (3.1) with $\nu = \nu_I$ to do the W integral:

$$\begin{aligned}
\langle\theta|\, W\, |\theta\rangle &= \frac{\sum_{\nu\in\mathbb{Z}}\int \mathcal{D}A^{(\nu)}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ e^{2\pi q i\nu_I/e}e^{-S}e^{i\nu\theta}}{\sum_{\nu\in\mathbb{Z}}\int \mathcal{D}A^{\nu}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ e^{-S}e^{i\nu\theta}},\\
&= \frac{\sum_{\nu_I,\nu_O\in\mathbb{Z}}\left(\int \mathcal{D}A^{(\nu_I)}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ e^{-S_I}e^{i\nu_I(\theta+2\pi q/e)}\right)\left(\int \mathcal{D}A^{(\nu_O)}\ e^{-S_O}e^{i\nu_O\theta}\right)}{\sum_{\nu\in\mathbb{Z}}\int \mathcal{D}A^{\nu}\ \mathcal{D}\psi^*\ \mathcal{D}\psi\ e^{-S}e^{i\nu\theta}},\\
&\overset{(2.5)}{=} \exp\{2Ke^{-S_0}[L'T'\cos(\theta+2\pi q/e) + (LT - L'T')\cos(\theta) - LT\cos(\theta)]\},
\end{aligned}$$

so that $\Delta = 2L'Ke^{-S_0}[\cos(\theta) - \cos(\theta+2\pi q/e)]$. Since Δ depends linearly in L', then there exists a **constant force between charges**, as expected from a Coulomb force in 1+1 dimensions[10].

Unfortunately, this argument doesn't generalize to 4 dimensions, since the area bounded by the loop integral of W has measure zero and, as such, W can be made trivial by a gauge transformation.

[9] A more accurate justification can be found in Schwartz [7], p. 531.

[10] To see this, we can use Gauss' Law $\boldsymbol{\nabla}\cdot\mathbf{E} = \rho/\varepsilon_0$ in 1 spatial dimension or treat each charge as a infinite plane in 3 spatial dimensions.

3.2 't Hooft Solution of the U(1) Problem

In Minkowski space, the QCD Lagrangian with quarks $q = (u, d)$ is

$$\mathcal{L}_{QCD} = -\frac{1}{4g^2}(F_{\mu\nu}, F^{\mu\nu}) + \sum_{f=1,2} \bar{q}_f(iD_\mu\gamma^\mu - m_f)q_f.$$

In the massless limit $m_f = 0$, QCD has a $U(2)_V \times U(2)_A$ symmetry acting on the flavor indices (see Table 2) that holds approximately well because of the small quark masses $m_u, m_d < \Lambda_{QCD}$. More specifically, $U(2)_V$ symmetry is realized and $SU(2)_A$ is spontaneously broken, as we can see from the quark condensates $\langle\bar{u}u\rangle = \langle\bar{d}d\rangle \neq 0$, giving rise to approximate Goldstone bosons, the pions. However, $U(1)_A$ is neither realized (there is no associated multiplet) nor spontaneously broken [5]. If it were the latter, there would be another particle as light as the pions, but the next candidate, the η' meson, is too heavy[11].

Table 2: $U(2)_V \times U(2)_A$ approximate symmetry of QCD.

Symmetry		Action	Consequence
$U(2)_V$	$SU(2)_V$	$e^{i\vec{\alpha}\cdot\vec{\sigma}/2}q$	Isospin Multiplet
	$U(1)_V$	$e^{i\alpha}q$	Barion number conservation
$U(2)_A$	$SU(2)_A$	$e^{i\vec{\alpha}\cdot\vec{\sigma}\gamma_5/2}q$	SSB: pions $\pi^0, \pi^\pm$
	$U(1)_A$	$e^{i\alpha\gamma_5}q$	See main text

You may think this is not a problem, since the axial symmetry is anomalous, for its current $j_5^\mu = \sum_f \bar{q}_f\gamma^\mu\gamma^5 q_f$ satisfies

$$\partial_\mu j_5^\mu = \frac{N}{16\pi^2}(F, *F) = \frac{N}{16\pi^2}\partial_\mu G^\mu, \tag{3.2}$$

where N is the number of massless quarks and $G_\mu = \epsilon_{\mu\nu\rho\sigma}(A^\nu, F^{\rho\sigma} - \frac{2}{3}A^\rho A^\sigma)$. But if we define a conserved current $J_5^\mu = j_5^\mu - \frac{N}{16\pi^2}G^\mu$, we would get corresponding Ward identities and Goldstone poles in Green's functions with J_5^μ, pointing again at a nonexistent Goldstone boson. This is the **U(1) Problem**.

't Hooft gave an explanation [9] of how this problem can be solved using instantons, which I will briefly comment here. As we saw in section 2.3, we need to include a term proportional to $\theta(F, *F)$ in the effective action to account for instantons and to select a θ vacuum. In turn, this extra term *explicitly breaks* the chiral symmetry, since, from (3.2), a chiral rotation applied to the effective action changes the θ vacuum. Therefore, the $U(1)_A$ transformation is not a symmetry to begin with.

Note that $(F, *F) \propto \mathbf{E}_a \cdot \mathbf{B}_a$ breaks CP symmetry. If we include this CP symmetry violation with the one coming from the electroweak sector in the quark mass matrix, we arrive at an effective $\bar{\theta}$ angle that gives an electric dipole moment for the neutron $d_n \approx \bar{\theta}\frac{em_q}{M_n^2}$. In order to comply to experimental upper bounds on d_n, the effective angle has to satisfy $\bar{\theta} \leq 10^{-9}$ rad [5]. This is called the **Strong CP Problem**.

4 Conclusion

Despite their simple definition, we have seen here many interesting features of instantons, including:

- How they affect the vacuum structure of quantum theories, via the equation (1.1);
- They are topological in nature, in the sense that the topological properties of the spacetime and of the gauge group are solely responsible for many qualitative features (see Section 2.2)

[11] Weinberg proved an upper limit of $\sqrt{3}m_\pi$ for the supposed $U(1)_A$ Goldstone boson [8]. The eta prime meson exceeds this limit by $m_{\eta'}/m_{\pi^+} > 6$.

- Their far-reaching consequences for the theory. For example, the existence of many vacua in a gauge theory and their appearance in the effective action (2.3), and the effects we discussed in the application (see Section 3).

References

[1] Sidney Coleman. *Aspects of Symmetry: Selected Erice Lectures*. Cambridge University Press, Feb. 18, 1988. Chap. 7. ISBN: 978-0-521-31827-3.

[2] Alexander Altland and Ben D. Simons. *Condensed Matter Field Theory*. Cambridge University Press, Mar. 11, 2010. Chap. 3, pp. 95–134. ISBN: 978-0-521-76975-4.

[3] Elon Lages Lima. *Grupo Fundamental e Espaços de Recobrimento*. 4th ed. IMPA, 2012. ISBN: 978-85-244-0086-5.

[4] Mikio Nakahara. *Geometry, Topology and Physics*. 2nd ed. Institute of Physics Publishing, 2003. Chap. 1, 9, 10.

[5] R. D. Peccei. "QCD, Strong CP and Axions". In: *arXiv:hep-ph/9606475* (June 28, 1996). arXiv: `hep-ph/9606475`. URL: `http://arxiv.org/abs/hep-ph/9606475` (visited on 11/17/2019).

[6] Kenneth G. Wilson. "Confinement of quarks". In: *Physical Review D* 10.8 (Oct. 15, 1974), pp. 2445–2459. DOI: `10.1103/PhysRevD.10.2445`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.10.2445` (visited on 11/13/2019).

[7] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Edição: New. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03473-0.

[8] Steven Weinberg. "The U(1) problem". In: *Physical Review D* 11.12 (June 15, 1975), pp. 3583–3593. DOI: `10.1103/PhysRevD.11.3583`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.11.3583` (visited on 11/17/2019).

[9] G. 't Hooft. "How instantons solve the U(1) problem". In: *Physics Reports* 142.6 (Sept. 1, 1986), pp. 357–387. ISSN: 0370-1573. DOI: `10.1016/0370-1573(86)90117-1`. URL: `http://www.sciencedirect.com/science/article/pii/0370157386901171` (visited on 11/09/2019).