Lecture 6: Instantons and the θ vacua

Cem Eröncel

cem.eroncel@.itu.edu.tr

Lecture date: March 28, 2023

Last update: April 7, 2023

In the last lecture, we have investigated instanton solutions in Quantum Mechanics. In this lecture, we will do a similar exercise for the Yang-Mills theory. In the meantime, we also discover the non-trivial structure of the QCD vacuum, and dig out a hidden parameter from the theory.

1 Brief review and the notation

Let us briefly review the Yang-Mills theories and at the same time specify the notation. The Lagrangian for a generic Yang-Mills theory in Minkowski space reads

$$S_{\mathrm{YM}} = -\frac{1}{2}\int \mathrm{d}^4x \, \mathrm{Tr}\big(G_{\mu\nu}G^{\mu\nu}\big) = -\frac{1}{4}\int \mathrm{d}^4x \, G^a_{\mu\nu}G^{a\,\mu\nu}, \tag{1}$$

where the $G_{\mu\nu}$ is the Lie-Algebra valued field strength defined by

$$G_{\mu\nu} = \partial_\mu A_\mu - \partial_\nu A_\mu - ig\left[A_\mu, A_\nu\right] =: G^a_{\mu\nu}T^a, \tag{2}$$

with A_μ being the Lie-algebra valued gauge potential. The latter is written as a linear combination of the group generators $\{T^a\}$:

$$A_\mu = A^a_\mu T^a. \tag{3}$$

The components of $G_{\mu\nu}$ are

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu, \tag{4}$$

where f^{abc} are the structure constants such that

$$\left[T^a, T^b\right] = i f^{abc} T^c. \tag{5}$$

For QCD, the gauge group is $\mathcal{G} = \mathsf{SU}(3)_c$, the gauge potentials $\left\{A^a_\mu\right\}^8_{a=1}$ are the gluons, and $T^a = \lambda^a/2$ where $\{\lambda^a\}^8_{a=1}$ are the Gell-Mann matrices.

Under a gauge transformation $U \in \mathcal{G}$, the gauge potential transforms as

$$A_\mu \mapsto U A_\mu U^{-1} + \frac{i}{g} U \partial_\mu U^{-1}, \tag{6}$$

while the transformation law for the field strength is

$$G_{\mu\nu} \mapsto U G_{\mu\nu} U^{-1}. \tag{7}$$

Notice that the gauge potential has a non-trivial transformation law in contrast to the field strength. This will play a crucial role in the coming discussion.

Euclidean Yang-Mills action

We will need the Euclidean Yang-Mills action to study instantons. We perform the Wick rotation via $x^4 := ix^0$. This implies

$$\partial_4 = \frac{\partial}{\partial x^4} = -i\partial_0 \quad \text{and} \quad A_4 = -iA_0. \tag{8}$$

From Eq. (4) we also get

$$G^a_{i4} = -iG^a_{i0} \quad \text{and} \quad G^{a\,i4} = iG^{a\,i0}. \tag{9}$$

while the other components remain unchanged. Therefore, the Yang-Mills action in (1) in Euclidean coordinates reads

$$S_{\text{YM}} = i\int \mathrm{d}^4x_E\, G^a_{\mu\nu}G^{a\,\mu\nu}. \tag{10}$$

Recall that we defined the Euclidean action as $S_E := -iS$ where S is the Wick-rotated original action. Therefore we express the Euclidean Yang-Mills action as

$$S^E_{\text{YM}} = \int \mathrm{d}^4x_E\, G^a_{\mu\nu}G^{a\,\mu\nu}. \tag{11}$$

With our Minkowski metric convention $(+,-,-,-)$, the Wick rotation yields a flat metric $g_{\mu\nu} = -\delta_{\mu\nu}$ on four-dimensional Euclidean spacetime. Therefore, while working in Euclidean space we can forget about the difference between the upper and lower spacetime indices and write the Euclidean action simply as[1]

[1] The overall minus sign vanishes because we have lowered two indices in Eq. (12).

$$S^E_{\text{YM}} = \int \mathrm{d}^4x_E\, G^a_{\mu\nu}G^a_{\mu\nu}. \tag{12}$$

Until the end of the lecture, we will drop the E-subscript in order not to clutter the notation. This means that every expression should be understood as it is written in Euclidean space where $\mu = 1,2,3,4$ until we Wick-rotate back to the Minkowski space at the end of the lecture.

2 Topology of the Yang-Mills vacuum

We want to explore the vacuum structure of the Yang-Mills theory, and a semi-classical analysis is suitable for this task. As we have seen in the previous lecture, the first step of semi-classical approximation is identifying the field configurations that minimizes the Euclidean action. From Eq. (12) we can easily see that the Euclidean Yang-Mills action is positive definite, therefore it is minimized when the field strength vanishes:

$$G^a_{\mu\nu}\Big|_{\text{VAC}} = 0. \tag{13}$$

This however, does not mean that the vector potentials $\left\{A^a_\mu\right\}$ should also vanish. From the gauge transformation laws given in Eqs. (6) and (7) we can see that if one starts with a zero gauge potential $A_\mu = 0$ and apply a gauge transformation $U \in \mathcal{G}$, the resulting field strength still vanishes but the vector potential becomes

$$A_\mu\Big|_{\text{VAC}} = \frac{i}{g} U(x)\partial_\mu U^{-1}(x). \tag{14}$$

These gauge fields that can be written as a gauge transformation of zero are called pure gauges, and they correspond to the *zero-energy states*. Since these are directly related to the Yang-Mills vacuum, it would be beneficial to have a classification for these fields which is given by the classification of the gauge transformations U. This is what we shall study now.

Each gauge transformation $U(x)$ is in general a map from the Euclidean space $\mathbb{R}^4$ to the gauge group $\mathcal{G}$. We can employ the gauge freedom in the Yang-Mills theory and choose the temporal gauge where we set $A_0 = 0$. This way we can restrict ourselves to the time-independent gauge transformations so instead of Eq. (14) we can write

$$\mathbf{A}\Big|_{\text{VAC}} = \frac{i}{g} U(\mathbf{x})\boldsymbol{\nabla} U^{-1}(\mathbf{x}). \tag{15}$$

Furthermore, we want to restrict ourselves to the zero-energy states that can be connected to each other via the tunneling transformations for which the Euclidean action in Eq. (12) is *finite*. This restriction is needed for two reasons:

1. The non-perturbative analysis we are performing relies on the semi-classical approximation, and from the previous lecture we know that the only the field configurations with finite Euclidean action have non-zero contribution in the semi-classical approximation.

2. The Euclidean action is related to the Hamiltonian of the system, so a transition with infinite action will need an infinite amount of energy. Therefore, we shall not take these states into the account.

It turns out this condition requires us to restrict the gauge transformations to the ones that approach to the identity at spatial infinity.[2] Namely,

$$U(\mathbf{x}) \to 1 \quad \text{as} \quad |\mathbf{x}| \to \infty. \tag{16}$$

[2] Curtis G. Callan, Roger Dashen, and David J. Gross (May 1978). In: *Phys. Rev. D* 17 (10), pp. 2717–2763.

This effectively identifies the boundary of the domain $\mathbb{R}^3$ with a single point, thereby converts $\mathbb{R}^3$ to $\mathbb{S}^3$. Therefore, we can think U as a map

$$U : \mathbb{S}^3 \to \mathcal{G}. \tag{17}$$

Winding number

We are interested in QCD so we choose $\mathcal{G} = \mathsf{SU}(3)_c$. It turns out we can further restrict $\mathsf{SU}(3)_c$ to its subgroup $\mathsf{SU}(2)$ thanks to a powerful theorem by Raoul Bott[3] which states that any continuous mapping of $\mathbb{S}^3$ into $\mathcal{G}$ can be continuously deformed into a mapping into an $\mathsf{SU}(2)$ subgroup of $\mathcal{G}$. Finally, any element $V \in \mathsf{SU}(2)$ can be written as

[3] Raoul Bott (1956). eng. In: *Bulletin de la Société Mathématique de France* 84, pp. 251–281.

$$V = v_0 + i v_a \tau^a \quad \text{such that} \quad v_0^2 + v_a v_a = 1, \tag{18}$$

where v_0 and $\{v_a\}$ are real parameters, and $\{\tau^a\}$ are some basis vectors. This is known as the Queternion representation of $\mathsf{SU}(2)$ which maps it to $\mathbb{S}^3$ that is also its group manifold. All this discussion shows that we are interested in the maps

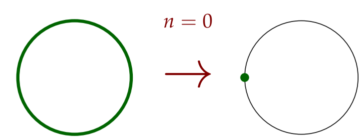

Figure 1: A trivial map which maps $\mathbb{S}^1$ into a single point.

$$U : \mathbb{S}^3 \to \mathbb{S}^3. \tag{19}$$

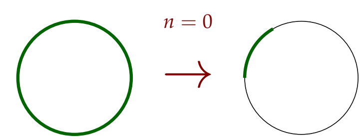

Figure 2: A mapping from $\mathbb{S}^1$ into a finite subset of $\mathbb{S}^1$.

The classification of these maps is accomplished by the homotopy theory. Two such maps are said to be in different homotopy classes if they cannot be *continuously* deformed into each other. This is not special to the maps from $\mathbb{S}^3$ to $\mathbb{S}^3$. It can be applied to any map from $\mathbb{S}^m$ to $\mathbb{S}^m$.

In order to gain some intuition, let us start with a much simpler case which is the classification of the maps from $\mathbb{S}^1$ to $\mathbb{S}^1$, i.e. from circle to circle. We have the following options:

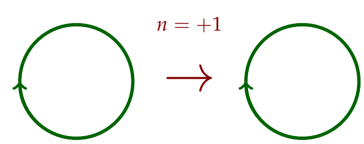

Figure 3: A one-to-one mapping from $\mathbb{S}^1$ into $\mathbb{S}^1$.

- We can map $\mathbb{S}^1$ into a single point, Figure 1. Let us give the number $n = 0$ to this map and denote it by the trivial map.
- We can map $\mathbb{S}^1$ into a finite subset of $\mathbb{S}^1$, Figure 2. However, we can transform this map smoothly to the trivial map so it is in the same homotopy class.
- We can map $\mathbb{S}^1$ into $\mathbb{S}^1$ as a one-to-one fashion, Figure 3. We cannot transform this map smoothly to the trivial map, so it is in a different homotopy class which we label as $n = +1$.

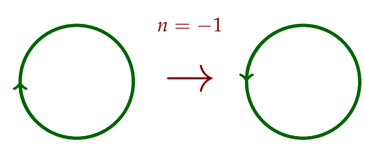

Figure 4: A one-to-one mapping from $\mathbb{S}^1$ into $\mathbb{S}^1$.

- We can again define a one-to-one map but reverse the direction of the increasing angle, Figure 4. This map cannot be deformed to neither of the previous maps so it should have a different label which we set as $n = -1$.
- We can define maps which wraps the $\mathbb{S}^1$ at the domain more than one times to the $\mathbb{S}^1$ at the target in both directions, Figure 5.

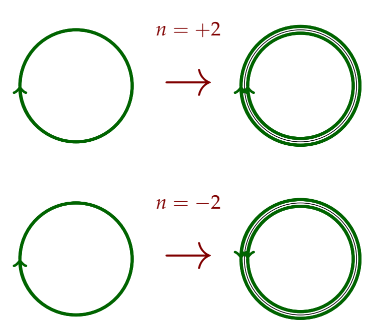

Figure 5: Mappings that warp $\mathbb{S}^1$ into $\mathbb{S}^1$ multiple times.

This example shows that we can classify the maps from $\mathbb{S}^1$ to $\mathbb{S}^1$ by a countable infinite number of equivalence classes labeled by an integer $n \in \mathbb{Z}$. This integer is called the winding number.

If one also defines a composition law for these equivalence classes where the winding numbers simply add, then these homotopy classes become elements of a group which is the homotopy group. It is labeled by

$$\pi_m(\mathbb{S}^n), \tag{20}$$

where m and n are the dimensions of the domain and target spheres respectively. In other words, the homotopy group of the maps from $\mathbb{S}^m$ to $\mathbb{S}^n$ is denoted by $\pi_m(\mathbb{S}^n)$.

For the maps from $\mathbb{S}^1$ to $\mathbb{S}^1$ we have shown that

$$\pi_1(\mathbb{S}^1) = \mathbb{Z}. \tag{21}$$

This result can be straightforwardly generalized to d dimensions as $\pi_d(\mathbb{S}^d)$. In particular,

$$\pi_3(\mathbb{S}^3) = \mathbb{Z}, \tag{22}$$

which is the result that we were seeking. With this results we can classify the zero-energy states, i.e. the pure gauges, by the winding number of U by defining

$$A_{(n)}\Big|_{\text{VAC}} = \frac{i}{g} U_{(n)}(\mathbf{x}) \boldsymbol{\nabla} U^{-1}_{(n)}, \tag{23}$$

where $U_{(n)}$ has the winding number n. This analysis shows that the QCD has a countably infinite number of degenerate "candidate" vacua $\{|n\rangle\}$ labeled by the winding number. We say "candidate" since the true QCD vacuum state is neither of these as we shall see soon.

Yang-Mills Instantons

It might be confusing to learn that they are gauge transformations that cannot be continuously converted into each other. We know that $\mathsf{SU}(3)_c$ is simply connected and therefore we should be able to go from an element to another element continuously. Indeed, we can do so. When obtaining the homotopy classification for the gauge transformations, we are restricted ourselves to the gauge transformations that are time-independent and obey the boundary condition (16). It is possible to go from a gauge transformation $U_{(n)}$ to another $U_{(m)}$ with different winding numbers $n \neq m$, but we should abandon the pure gauge condition. Each topological class correspond to an absolute action minimum, and these different minima are separated by a finite, non-zero action barrier called a sphaleron barrier.

Let us assume that we start at the pure gauge configuration $A_\mu^{(n)}\Big|_{\text{VAC}}$ at $x^4 = -\infty$ with the winding number n, and will end at the pure gauge $A_\mu^{(m)}\Big|_{\text{VAC}}$ at $x^4 = +\infty$ with the winding number $m = n + 1$. The Euclidean equation of motion is

$$\partial_\mu G_{\mu\nu} + g\left[A_\mu, G_{\mu\nu}\right] =: D_\mu G_{\mu\nu} = 0. \tag{24}$$

The solution to this equation $A_\mu^{(I)}(x)$ subject to the boundary conditions mentioned above is called a Yang-Mills instanton. Explicitly, the solution is[4]

$$A_\mu^{(I)}(x) = \frac{2}{g}\frac{\eta^a_{\mu\nu}(x_\nu - z_\nu)\tau^a}{(x-z)^2+\rho^2}, \tag{25}$$

where $\{\tau^a\}$ are the $\mathrm{SU}(2)$ generators, ρ is a parameter corresponding to the 4D size of the instanton, z denotes the instanton center, and finally $\eta^a_{\mu\nu}$ is the 't Hooft symbol:[5,6]

$$\eta^a_{\mu\nu} = \delta_{a\mu}\delta_{4\nu} - \delta_{a\nu}\delta_{4\mu} + \epsilon_{a\mu\nu 4}. \tag{26}$$

Some properties can be read directly from the solution:

- Its "spin", i.e. its Lorentz index is coupled with the color orientation via the 't Hooft symbol.
- Its non-perturbative character is imprinted by the diverging weak coupling limit.
- The solution states that the tunneling process is localized in a region of size ρ in space and time, around its center z.

By repeating the same exercise, but now choosing $m = n - 1$ will give us a Yang-Mills anti-instanton. The solution is identical to the instanton solution in Eq. (25) except one uses the anti-self-dual 't Hooft symbol:[7]

$$\overline{\eta}^a_{\mu\nu} = -\delta_{a\mu}\delta_{4\nu} + \delta_{a\nu}\delta_{4\mu} + \epsilon_{a\mu\nu 4}. \tag{29}$$

Instanton action

From the instanton solution given in Eq. (25), we can easily get the corresponding instanton field strength as

$$G_{\mu\nu}^{(I)}(x) = -\frac{4\rho^2}{g}\frac{\eta^a_{\mu\nu}\tau^a}{\left[(x-z)^2+\rho^2\right]}. \tag{30}$$

From this one can calculate the instanton action in Euclidean space as

$$S_I = \frac{1}{2}\int \mathrm{d}^4x\, \mathrm{Tr}\, G_{\mu\nu}^{(I)} G_{\mu\nu}^{(I)} = \frac{8\pi^2}{g^2}. \tag{31}$$

The anti-instanton action turns out to be the same. In fact, for multi-instanton solutions from n to m with $|m-n| > 1$, one gets $S_I = |m-n|S_I$.[8]

We see that the action is finite and greater than zero as expected. It also does not depend on

[4] A.A. Belavin et al. (1975). In: *Physics Letters B* 59.1, pp. 85–87. ISSN: 0370-2693.

[5] Since we are in Euclidean space, we can be sloppy about the location of the indices.

[6] With the convention $\epsilon_{1234} = +1$.

[7] The name comes from the fact that $\overline{\eta}^a_{\mu\nu}$ is anti-self-dual, namely

$$\overline{\eta}^a_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\overline{\eta}^a_{\rho\sigma}, \tag{27}$$

whereas $\eta^a_{\mu\nu}$ is self-dual:

$$\eta^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\eta^a_{\rho\sigma}. \tag{28}$$

[8] This result does not apply for the combination of both.

- the center z of the instanton due to the translation Invariance,
- on the gauge transformation U due to the gauge Invariance,
- and on the size of the instanton ρ due to the scale invariance of the Yang-Mills Lagrangian.

The scale invariance is a classical symmetry which gets broken in the quantum theory by the so-called trace anomaly.

3 Quark zero modes and index theorems

Previously in Lecture 4: Anomalies, we have derived the chiral anomaly that arises by performing a chiral transformation $q \mapsto e^{i\alpha\gamma_5} q$ to a single quark q as

$$\partial_\mu J_{\mu 5} = \frac{\alpha g^2}{8\pi^2} \operatorname{Tr} G_{\mu\nu} \widetilde{G}_{\mu\nu}. \tag{32}$$

We have also stated that $G^a_{\mu\nu} \widetilde{G}^a_{\mu\nu}$ can be written as a total derivative of the Chern-Simons current $\mathcal{K}^\mu$ as

$$G^a_{\mu\nu} \widetilde{G}^a_{\mu\nu} = \partial_\mu \epsilon_{\mu\alpha\beta\gamma} \left(A^a_\alpha G^a_{\beta\psi} - \frac{g_s}{3} f^{abc} A^a_\alpha A^b_\beta A^c_\gamma \right) =: \partial_\mu \mathcal{K}^\mu. \tag{33}$$

This implies that the integral of Eq. (32) over the Euclidean spacetime should yield a quantity that depends only on the boundary, i.e. it should be topological. Now, we will demonstrate this explicitly, and in the meanwhile discover other features.

We define the following quantity in Euclidean space:

$$Q := \frac{g^2}{16\pi^2} \int d^4x \operatorname{Tr} G_{\mu\nu} \widetilde{G}_{\mu\nu}, \tag{34}$$

which we call as the topological charge due to the reasons that will be clear soon. With a flashback to our anomaly lecture, we can see that this term can be expressed as[9]

$$Q = -\lim_{\Lambda\to\infty} \sum_m e^{-\lambda_m^2/\Lambda^2} \int d^4x\, \psi_m^\dagger \gamma^5 \psi_m, \tag{35}$$

[9] In contrast to the anomaly lecture, we are not using the E-subscript in order not to clutter the notation.

where $\{\lambda_m\}$ and $\{\psi_m\}$ are the eigenvalues and the eigenfunctions of the Euclidean Dirac operator:

$$\not{D}\psi_m(x) = \lambda_m \psi_m(x). \tag{36}$$

It can be shown that the Dirac operator $\not{D}$ anti-commutes with γ^5. Then

$$\not{D}\gamma^5 \psi_n(x) = -\gamma^5 \not{D}\psi_n(x) = -\lambda_n \gamma^5 \psi_n(x). \tag{37}$$

So for each eigenfunction $\psi_n(x)$ with a positive eigenvalue $\lambda_n > 0$, there exists another eigenfunction

$$\psi_{-n}(x) := \gamma^5 \psi_n(x) \tag{38}$$

with the eigenvalue

$$\lambda_{-n} = -\lambda_n. \tag{39}$$

Thus, non-vanishing eigenvalues appear in the spectrum in pairs with opposite signs.

The remaining eigenfunctions $\psi_{0,k}(x)$ with eigenvalues $\lambda_k = 0$ are called zero modes. They can be written in a basis, the so-called chiral basis [10]

$$\psi^{\pm}_{0,k} = \frac{1}{2}\left(1 \pm \gamma^5\right)\psi_{0,k} =: \mathcal{P}_{\pm}\psi_{0,k}, \tag{40}$$

so that γ^5 is diagonalized, i.e.

$$\gamma^5\psi^{\pm}_{0,k} = \pm\psi^{\pm}_{0,k}. \tag{41}$$

Now we return to the evaluation of Eq. (35). Eq. (38) implies

$$\int d^4x\, \psi^{\dagger}_m \gamma^5 \psi_m = \int d^4x\, \psi^{\dagger}_m \psi_{-m} = 0, \tag{42}$$

for eigenfunctions with non-zero eigenvalues since ψ_m and $\psi^{\dagger}_m$ should be orthogonal to each other. Thus, the sum in (35) is only over the zero modes. By trivially evaluating the limit we get

$$Q = -\sum_k \int d^4x\, \psi^{\dagger}_{0,k}\gamma^5\psi_{0,k}. \tag{43}$$

By using the fact that $\mathcal{P}^2_{\pm} = \mathcal{P}_{\pm}$ and $\mathcal{P}^{\dagger}_{\pm} = \mathcal{P}_{\pm}$ we can show that

$$\psi^{\dagger}_{0,k}\gamma^5\psi_{0,k} = \left(\psi^{+}_{0,k}\right)^{\dagger}\psi^{+}_{0,k} - \left(\psi^{-}_{0,k}\right)^{\dagger}\psi^{-}_{0,k}. \tag{44}$$

Therefore

$$\begin{aligned} Q &= \sum_{m=1}^{n_-} \int d^4x \left(\psi^{-}_{0,k}\right)^{\dagger}\psi^{-}_{0,k} - \sum_{m=1}^{n_+} \int d^4x \left(\psi^{+}_{0,k}\right)^{\dagger}\psi^{+}_{0,k} \\ &= n_- - n_+, \end{aligned} \tag{45}$$

where n_- and n_+ are the number of the left- and right-handed zero modes per flavor in the given background gluon field respectively. We see that Q must be an integer and cannot change under smooth variations of the background gluon field. That's why we named it as the topological charge. This is a special case of the celebrated Atiyah-Singer index theorem for the Euclidean Dirac operator.

[10] The operators $\mathcal{P}_+$ and $\mathcal{P}_-$ are also used to project a Dirac fermion to its right and left components respectively. This is because in Weyl basis

$$\mathcal{P}_+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{P}_- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Classification of the gluon fields with the topological charge

Since the topological charge Q can take on only integer values, and therefore cannot change under continuous deformation of the gauge fields, it is a very convenient quantum number to classify the gauge fields. How do we perform this assignment?

Again, we are interested in the gluon fields that have a finite Euclidean action. For this, the field strength $G_{\mu\nu}(x)$ should vanish faster than $1/|x|^2$ at infinity. This means that the gauge fields at infinity should approach to a pure gauge. Thus

$$\lim_{|x|^2\to\infty} A_\mu(x) = \frac{i}{g} U\partial_\mu U^{-1}. \tag{46}$$

This equation defines a map $\mathbb{S}^3 \to \mathcal{G}$, where for QCD, $\mathcal{G} = \mathsf{SU}(3)_c$. Again only the $\mathsf{SU}(2)$ subgroup is active so Eq. (46) is a map from $\mathbb{S}^3$ to $\mathbb{S}^3$. We already know that these maps can be classified by their homotopy classes and the homotopy group is

$$\pi_3(\mathbb{S}^3) = \mathbb{Z}. \tag{47}$$

This shows that all finite-action gluons fields in Euclidean QCD fall into topologically distinct equivalence classes labeled by the topological charge Q. This charge is also known as the Pontryagin index of the gauge field.

Topological charge of the instantons

It is not hard to convince ourselves that the instanton and anti-instanton solutions we have written previously satify Eq. (46) since they interpolate between the pure gauges with different winding numbers. We can even evaluate their topological charges very easily.

Since the 't Hooft index is self-dual, see Eq. (28), we have that

$$\widetilde{G}^{(I)}_{\mu\nu} = G^{(I)}_{\mu\nu}. \tag{48}$$

Then the topological charge of the instanton can be calculated as

$$\begin{aligned} Q_I =& \frac{g^2}{16\pi^2}\int \mathrm{d}^4x \,\mathrm{Tr}\, G^{(I)}_{\mu\nu}\widetilde{G}^{(I)}_{\mu\nu} \\ =& \frac{g^2}{16\pi^2}\int \mathrm{d}^4x \,\mathrm{Tr}\, G^{(I)}_{\mu\nu}G^{(I)}_{\mu\nu} \\ =& 1 \end{aligned} \tag{49}$$

In a similar way, one can show that for the anti-instanton solution

$$Q_{\bar{I}} = -1. \tag{50}$$

Therefore we get

$$Q = n_{\text{final}} - n_{\text{initial}}, \tag{51}$$

where n_{initial} and n_{final} are the winding numbers of the initial and final pure gauges respectively.

4 Hidden θ parameter via the Cluster Decomposition

After the discussion of the QCD vacuum and the classification of the gluon fields according to their topological charge, i.e. their Pontryagin indices, we know demonstrate that these non-trivial properties forces us to add another term to the QCD Lagrangian given in Eq. (1).

Let $\mathcal{O}[A, q, \overline{q}]$ be an operator consisting of quarks and gluons, i.e. the QCD fields. We assume that the vacuum expectation value (VEV) of this operator is strongly localized in a spacetime volume $\mathcal{V}$. This assumption is reasonable, since QCD confines at low energy. This strong localization implies that the VEV is non-zero only in a small sub-volume $\mathcal{V}_1$ inside $\mathcal{V}$. Using the Euclidean path integral, we can write an expression for this VEV in the most general form as

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\int \mathcal{D}q\,\mathcal{D}\overline{q}\,\mathcal{D}A\,\mathcal{O}[q,\overline{q},A]e^{-S_E[q,\overline{q},A]}}{\int \mathcal{D}q\,\mathcal{D}\overline{q}\,\mathcal{D}A\,e^{-S_E[q,\overline{q},A]}}, \tag{52}$$

where $|0\rangle$ is the vacuum state which is yet to be specified. The quark fields play no role in this discussion so we drop them from the path integral to simplify the notation.

In the previous section we have seen that the gluon fields for which the Euclidean action is finite are classified according to their topological charge Q. With this in mind, we can write the VEV as

$$\langle 0|\mathcal{O}|0\rangle = \frac{\sum_Q w(Q) \int \mathcal{D}A_Q\,\mathcal{O}[A]e^{-S_E[A]}}{\sum_Q w(Q) \int \mathcal{D}A_Q\,e^{-S_E[A]}}, \tag{53}$$

where the sum is from $Q = -\infty$ to $Q = \infty$, and $w(Q)$ is a so far unknown weight function. Let us split the Euclidean volume $\mathcal{V}$ into two volumes:

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2. \tag{54}$$

This way the action factorizes as

$$S_E[\mathcal{V}] = S_E[\mathcal{V}_1] + S_E[\mathcal{V}_2]. \tag{55}$$

The corresponding topological charges are also factorized:

$$Q = Q_1 + Q_2. \tag{56}$$

Strictly speaking, even though Q needs to be an integer, Q_1 and Q_2 do not have to be integers. However, the gluon fields with non-zero Q are the instantons so their topological charge densities are strongly localized too. Therefore we can approximately take Q_1 and Q_2 to be integers.

The factorization of the topological charge allows us to factorize the path integral measure as

$$\begin{aligned}\sum_Q w(Q) \int \mathcal{D}A_Q &= \sum_Q w(Q) \sum_{Q_1} \int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)} \sum_{Q_2} \int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)} \times \delta_{Q,Q_1+Q_2} \\ &= \sum_{Q_1}\sum_{Q_2} w(Q_1+Q_2) \int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)} \int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)}.\end{aligned} \tag{57}$$

So the VEV becomes

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\sum_{Q_1}\sum_{Q_2} w(Q_1+Q_2)\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}\mathcal{O}[A_{Q_1}]e^{-S_E[\mathcal{V}_1]}\int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)}e^{-S_E[\mathcal{V}_2]}}{\sum_{Q_1}\sum_{Q_2} w(Q_1+Q_2)\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}e^{-S_E[\mathcal{V}_1]}\int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)}e^{-S_E[\mathcal{V}_2]}} \tag{58}$$

Note that we took $\mathcal{O}[A_Q] = \mathcal{O}[A_{Q_1}]$ since $\mathcal{O}$ is assumed to be strongly localized in $\mathcal{V}_1$. This assumption also implies that the VEV should not depend on anything in $\mathcal{V}_2$ as a result of the principle of cluster decomposition. From Eq. (58), we can observe that this is possible if

$$w(Q_1+Q_2) = w(Q_1)w(Q_2). \tag{59}$$

This fixes the weight $w(Q)$ to be

$$w(Q) = e^{iQ\theta}, \quad \theta\in[0,2\pi) \quad \text{and} \quad \theta\in\mathbb{R}, \tag{60}$$

where we have introduced the θ-parameter. It should be real since the weight should be finite for all $Q\in(-\infty,+\infty)$. The condition in Eq. (59) ensures that all the Q_2 dependent terms cancel in Eq. (58) so that we are left with

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\sum_{Q_1} e^{iQ_1\theta}\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}\mathcal{O}[A_{Q_1}]e^{-S_E[\mathcal{V}_1]}}{\sum_{Q_1} e^{iQ_1\theta}\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}e^{-S_E[\mathcal{V}_1]}}. \tag{61}$$

By defining

$$\mathcal{D}A := \sum_Q \mathcal{D}A_Q \quad \text{and} \quad S'_E := S_E - iQ\theta, \tag{62}$$

we can write Eq. (61) as

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\int \mathcal{D}A\,\mathcal{O}[A]\,e^{-S'_E}}{\int \mathcal{D}A\,e^{-S'_E}} \tag{63}$$

Hence, summing over the topological charges is equivalent to adding a θ-dependent parameter into the action. By using the definition of Q, i.e. Eq. (34), the new action becomes explicitly

$$S'_E = S_E - \frac{i\theta g^2}{16\pi^2}\int \mathrm{d}^4x_E\,\mathrm{Tr}\,G_{\mu\nu}\widetilde{G}_{\mu\nu}. \tag{64}$$

Analytically continuing back to the Minkowski space via

$$\mathrm{d}^4x_E = i\,\mathrm{d}^4x_M\,, \quad G^a_{i4} = -iG^a_{i0}, \quad \epsilon^{0123} = -1, \tag{65}$$

we get in Minkowski space

$$S'_M = S_M + \frac{\theta g^2}{16\pi^2}\int \mathrm{d}^4x_M\,\mathrm{Tr}\,G_{\mu\nu}\widetilde{G}^{\mu\nu}. \tag{66}$$

At the end, we conclude that the *correct* QCD Lagrangian reads

$$\begin{aligned}\mathcal{L}_{\mathrm{QCD}} &= -\frac{1}{2}\,\mathrm{Tr}\,G_{\mu\nu}G^{\mu\nu} + \frac{\theta g_s^2}{16\pi^2}\,\mathrm{Tr}\,G_{\mu\nu}\widetilde{G}^{\mu\nu}\\ &= -\frac{1}{4}G^a_{\mu\nu}G^{a\,\mu\nu} + \frac{\theta g_s^2}{32\pi^2}G^a_{\mu\nu}\widetilde{G}^{a\,\mu\nu}\end{aligned} \tag{67}$$

The additional term is called the θ-term. This analysis shows that this term *has* to be included in the QCD Lagrangian, and θ appears as a hidden parameter of QCD.

5 Theta vacua

Finally, we will define the vacuum state(s) of the QCD. We say initially said that the "candidate" vacua are labeled by the winding number n. The true vacuum state of QCD depends on the θ parameter and given by[11]

[11] R. Jackiw (Oct. 1980). In: *Rev. Mod. Phys.* 52 (4), pp. 661–673.

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{i\theta n} |n\rangle . \tag{68}$$

This is called the θ vacua. Each θ value corresponds to another vacuum state. More importantly, all these vacua are completely secluded from each other. To see this consider the time ordered product $T(\mathcal{O}_1\mathcal{O}_2\cdots)$ of a set of gauge invariant operators. Let $|\theta_1\rangle$ and $|\theta_2\rangle$ are two vacua with $\theta_1 \neq \theta_2$. Then

$$\langle\theta_1|T(\mathcal{O}_1\mathcal{O}_2\cdots)|\theta_2\rangle = \sum_{m,n} e^{i(\theta_2 n-\theta_1 m)} \langle m|T(\mathcal{O}_1\mathcal{O}_2\cdots)|n\rangle . \tag{69}$$

The matrix element $\langle m|T(\mathcal{O}_1\mathcal{O}_2\cdots)|n\rangle$ depends only on the difference $Q = n - m$. Thus, Eq. (69) becomes

$$\begin{aligned}\langle\theta_1|T(\mathcal{O}_1\mathcal{O}_2\cdots)|\theta_2\rangle &= \sum_n e^{i(\theta_2-\theta_1)n} \sum_Q e^{-i\theta_1 Q} F(Q) \\ &= 2\pi\delta(\theta_2-\theta_1) \sum_Q e^{-\theta_1 Q} F(Q).\end{aligned} \tag{70}$$

This equation is zero if $\theta_1 \neq \theta_2$. This property describes a superselection rule which indicates that θ is a fundamental parameter labeling the Yang-Mills vacuum, and each value of θ corresponds to a different theory.

References

Belavin, A.A. et al. (1975). "Pseudoparticle solutions of the Yang-Mills equations". In: *Physics Letters B* 59.1, pp. 85–87. ISSN: 0370-2693. URL: `https://www.sciencedirect.com/science/article/pii/037026937590163X`.

Bott, Raoul (1956). "An application of the Morse theory to the topology of Lie-groups". eng. In: *Bulletin de la Société Mathématique de France* 84, pp. 251–281. URL: `http://eudml.org/doc/86904`.

Callan, Curtis G., Roger Dashen, and David J. Gross (May 1978). "Toward a theory of the strong interactions". In: *Phys. Rev. D* 17 (10), pp. 2717–2763. URL: `https://link.aps.org/doi/10.1103/PhysRevD.17.2717`.

Jackiw, R. (Oct. 1980). "Introduction to the Yang-Mills quantum theory". In: *Rev. Mod. Phys.* 52 (4), pp. 661–673. URL: https://link.aps.org/doi/10.1103/RevModPhys.52.661.