Prepared for submission to JHEP

The Strong CP Problem

Chang Hyeon Lee

Department of Physics, Chungnam National University, Daejeon 34134, South Korea

E-mail: changhyeonlee@cnu.ac.kr

Abstract: This is a my note for summary.

Contents

1 Symmetry description in (Quantum) field theory **1**
1.1 Noether's theorem 1
1.2 Ward Identity in Quantum Field Theory 2
1.3 Spontaneous Symmetry Breaking 5
1.3.1 Symmetry transformations and generators 5
1.3.2 Spontaneous Breaking of Discrete Symmetry 5
1.3.3 Spontaneous Breaking of Continuous Symmetry 6

2 Chiral anomaly **6**
2.1 Euclidean path integral 8
2.1.1 Fermion measure 9
2.2 Calculating the Jacobian 11
2.3 Non-conservation of the axial current 20
2.4 The Anomaly with Non-abelian gauge theories 23

3 Instantons and θ vacua **25**
3.1 Instantons in quantum mechanics 25
3.1.1 Double-well potential 27
3.2 Euclidean Yang-Mills action 29
3.3 Topology of the Yang-Mills vacuum 30
3.3.1 Winding number 32
3.3.2 Yang-Mills Instantons 34
3.3.3 Instanton action 35
3.4 Quark zero modes and index theorems 36
3.4.1 Classification of the gluon fields with the topological charge 38
3.4.2 Topological charge of the instantons 38
3.5 Hidden θ parameter via the Cluster Decomposition 39
3.6 Theta vacua 41

4 The Strong CP Problem **41**
4.1 The QCD Lagrangian 41
4.2 Quark masses in the Standard Model 42
4.2.1 Electroweak symmetry breaking 42
4.2.2 Fermions in the Standard Model 44
4.2.3 Fermion masses 44
4.2.4 Cabibbo-Kobayashi-Maskawa (CKM) matrix 46
4.3 Strong CP violation 47
4.3.1 Chiral rotations with multiple generations 47
4.3.2 Phase induced by the diagonalization of the quark mass matrix 49
4.3.3 Physical θ parameter 50

4.4 Strong CP Problem 51

A Grassmann numbers **51**

1 Symmetry description in (Quantum) field theory

1.1 Noether's theorem

Let the Lagrangian of n fields $\{\phi_n\}$ is a functional of filed and its first derivative

$$\mathcal{L} = \mathcal{L}(\phi_n, \partial_\mu \phi_n), \tag{1.1}$$

and the action is invariant under a continuous global symmetry parametrized by ϵ. Since the symmetry is continuous, we can consider an infinitesimal transformation,

$$\phi_n(x) \to \phi'_n(x) = \phi_n(x) + \epsilon \chi_n(\phi). \tag{1.2}$$

To derive Noether's theorem, we promote the constant parameter ϵ to a spacetime-dependent function $\epsilon(x)$,

$$\phi_n(x) \to \phi'_n(x) = \phi_n(x) + \epsilon(x) \chi_n(\phi), \tag{1.3}$$

where $\chi_n(\phi)$ depends on a specific transformation rule. Under this transformation, the Lagrangian is transform as $\mathcal{L} \to \mathcal{L} + \delta_\epsilon \mathcal{L}$,

$$\begin{aligned}
\delta_\epsilon \mathcal{L} &= \sum_n \left[\frac{\partial \mathcal{L}}{\partial \phi_n} \delta_\epsilon \phi_n + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \delta_\epsilon(\partial_\mu \phi_n(x)) \right] \\
&= \sum_n \left[\frac{\partial \mathcal{L}}{\partial \phi_n} \delta_\epsilon \phi_n + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \partial_\mu(\delta_\epsilon \phi_n(x)) \right], \qquad \delta_\epsilon(\partial_\mu \phi_n) = \partial_\mu(\delta_\epsilon \phi_n), \\
&= \sum_n \left[\frac{\partial \mathcal{L}}{\partial \phi_n} \epsilon(x) \chi_n(\phi) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \partial_\mu(\epsilon(x) \chi_n(\phi)) \right] \\
&= \sum_n \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} (\partial_\mu \epsilon) \chi_n(\phi) + \left\{ \frac{\partial \mathcal{L}}{\partial \phi_n} \chi_n(\phi) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \partial_\mu \chi_n(\phi) \right\} \epsilon(x) \right].
\end{aligned} \tag{1.4}$$

If this transformation is global symmetric transformation, i.e. ϵ is constant, then it can modify the Lagrangian only up to a total derivative,

$$\delta_{\epsilon_0} \mathcal{L} = \sum_n \left[\frac{\partial \mathcal{L}}{\partial \phi_n} \chi_n(\phi) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \partial_\mu \chi_n(\phi) \right] \epsilon_0 = \epsilon_0 \partial_\mu K^\mu. \tag{1.5}$$

Since this holds for arbitrary constant ϵ_0, we may identify the expression in brackets with $\partial_\mu K^\mu$,

$$\delta_\epsilon \mathcal{L} = \sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} (\partial_\mu \epsilon) \chi_n(\phi) + \partial_\mu K^\mu \epsilon(x) \tag{1.6}$$

Then the corresponding change in the action reads

$$\delta_\epsilon S = \int d^4x\, \delta_\epsilon \mathcal{L} = \int d^4x \left[\sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} (\partial_\mu \epsilon) \chi_n(\phi) + \partial_\mu K^\mu \epsilon(x) \right] \tag{1.7}$$
$$= - \int d^4x\, \epsilon(x) \partial_\mu \left[\sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \chi_n(\phi) - K^\mu \right] + \int d^4x\, \partial_\mu \left[\sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \epsilon(x) \chi_n(\phi) \right].$$

If the fields $\{\phi_n\}$ obey the classical equation of motion, then the arbitrary variation of action is vanished, including variation due to the for given symmetry transformation,

$$\delta S = 0 \quad \to \quad \delta_\epsilon S = 0 \quad \to \quad \partial_\mu J^\mu = 0, \tag{1.8}$$

where

$$J^\mu = \sum_n \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_n)} \chi_n(\phi) - K^\mu. \tag{1.9}$$

Note that there is no contribution of the total derivative term, if we assumed that $\epsilon(x)$ vanished sufficiently fast at spatially infinity.

1.2 Ward Identity in Quantum Field Theory

In this subsection we will derive the quantum version of Noether theorem. Consider the single scalar field ϕ. Then the generating functional is

$$Z[K] = \int \mathcal{D}\phi \, \exp \left[iS[\phi] + i \int d^4x\, K(x) \phi(x) \right], \tag{1.10}$$

where $K(x)$ is an external source for $\phi(x)$ and $\mathcal{D}\phi$ denotes integration over all field configurations. As in the derivation of Noether's theorem, we consider a continuous global symmetry parametrized by ϵ. Its infinitesimal form is

$$\phi(x) \to \phi'(x) = \phi(x) + \epsilon \chi(\phi). \tag{1.11}$$

Then promote the constant ϵ to a local parameter $\epsilon(x)$ as follows,

$$\phi(x) \to \phi'(x) = \phi(x) + \epsilon(x) \chi(\phi). \tag{1.12}$$

This is simply a change of integration variable; we may rewrite the generating functional in terms of the new field ϕ',

$$Z[K] = \int \mathcal{D}\phi' \, \exp \left[iS[\phi'] + i \int d^4x\, K(x) \phi'(x) \right]. \tag{1.13}$$

Recalling the formula for variation of action under the symmetry transformation (1.7),

$$S[\phi'] = S[\phi] - \int d^4x\, \epsilon(x) \partial_\mu J^\mu(x), \tag{1.14}$$

where $J^\mu(x)$ is Neother current. Then (1.13) becomes

$$Z[K] = \int \mathcal{D}\phi' \left[\exp\left\{iS[\phi] + i\int d^4x\, K(x)\phi(x)\right\}\right.$$
$$\left. \times \exp\left\{-i\int d^4x\, \epsilon(x)[\partial_\mu J^\mu(x) - K(x)\chi(\phi)]\right\}\right]. \qquad (1.15)$$

Since $\epsilon(x)$ is infinitesimal, one can expand the second exponential part,

$$Z[K] = \int \mathcal{D}\phi' \left[\exp\left\{iS[\phi] + i\int d^4x\, K(x)\phi(x)\right\}\right.$$
$$\left. \times \left\{1 - i\int d^4x\, \epsilon(x)[\partial_\mu J^\mu(x) - K(x)\chi(\phi)]\right\}\right]. \qquad (1.16)$$

We need an additional assumption that the integral measure $\mathcal{D}\phi$ does not change under the symmetry transformation,

$$\mathcal{D}\phi' = \mathcal{D}\phi. \qquad (1.17)$$

With this assumption, the first part of above expression becomes original generating functional,

$$Z[K] = Z[K] - i\int \mathcal{D}\phi \left[\exp\left\{iS[\phi] + i\int d^4x\, K\phi\right\} \times \int d^4x\, \epsilon(x)[\partial_\mu J^\mu - K\chi(\phi)]\right].$$

It gives

$$0 = \int \mathcal{D}\phi \left[\exp\left\{iS[\phi] + i\int d^4x\, K\phi\right\} \times \int d^4x\, \epsilon(x)[\partial_\mu J^\mu - K\chi(\phi)]\right]. \qquad (1.18)$$

Change the integral order as follows,

$$0 = \int d^4x\, \epsilon(x) \int \mathcal{D}\phi\, \exp\left[iS[\phi] + i\int d^4x\, K\phi\right][\partial_\mu J^\mu - K\chi(\phi)] \qquad (1.19)$$

This expression should be true for arbitrary $\epsilon(x)$,

$$0 = \int \mathcal{D}\phi\, \exp\left[iS[\phi] + i\int d^4x\, K\phi\right][\partial_\mu J^\mu - K\chi(\phi)]. \qquad (1.20)$$

By setting $K = 0$, we get

$$\int \mathcal{D}\phi\, \partial_\mu J^\mu e^{(iS[\phi])} = \langle \partial_\mu J^\mu \rangle = 0, \qquad (1.21)$$

where We normalize the generating functional $Z[0] = 1$.

Differentiating (1.20) multiple times with respect to K and then setting $K = 0$ gives

$$\partial_\mu \left\langle T\left\{J^\mu(x)\, \phi(x^1)\cdots\phi(x^n)\right\}\right\rangle = -i\sum_{i=1}^{n} \delta(x - x^i)\, \left\langle T\left\{\phi(x^1)\cdots\chi(\phi(x^i))\cdots\phi(x^n)\right\}\right\rangle$$
$$= 0, \qquad (x \neq x^i). \qquad (1.22)$$

It implies the divergence is zero away from contact points. Functional derivatives with respect to K commute with the path integral and bring down field insertions. These are known as the Ward identities and they mean that $\partial_\mu J^\mu$ vanishes inside any correlation function as long as its position does not coincide with the insertion point of other fields. This is the quantum version of the Noether theorem. If the functional measure is invariant, these Ward identities show that the classical symmetry is preserved in the quantum theory.

Derivation

We start from the identity (1.20) written with explicit spacetime dependence,

$$0 = \int \mathcal{D}\phi \; \exp\left(iS[\phi] + i \int d^4 z \, K(z)\phi(z) \right) \left[\partial_\mu J^\mu(x) - K(x)\chi(\phi(x)) \right]. \quad (1.23)$$

To generate correlation functions we differentiate with respect to the source K. The only two functional-derivative rules we need are

$$\frac{\delta}{\delta K(y)} \exp\left(i \int d^4 z \, K(z)\phi(z) \right) = i\phi(y) \exp\left(i \int d^4 z \, K(z)\phi(z) \right), \; \frac{\delta K(x)}{\delta K(y)} = \delta(x-y).$$

Now act with $\frac{\delta}{\delta K(x^1)} \cdots \frac{\delta}{\delta K(x^n)}$ on (1.23) and set $K = 0$. For the first term, all derivatives hit the source exponential and simply bring down field insertions:

$$\frac{\delta^n}{\delta K(x^1) \cdots \delta K(x^n)} \left(e^{i \int K\phi} \, \partial_\mu J^\mu(x) \right) \bigg|_{K=0} = i^n \, \phi(x^1) \cdots \phi(x^n) \, \partial_\mu J^\mu(x). \quad (1.24)$$

For the second term, notice an important simplification: after setting $K = 0$, the factor $K(x)$ kills every contribution except those in which *exactly one* derivative acts on $K(x)$. The specific calculation gives the contact terms

$$\begin{aligned} \frac{\delta^n \left(e^{i \int K\phi} \, K(x)\chi(\phi(x)) \right)}{\delta K(x^1) \cdots \delta K(x^n)} \Bigg|_{K=0} &= \sum_{j=1}^{n} \left(\frac{\delta^{n-1} \left(e^{i \int K\phi} \right)}{\delta K(x^1) \cdots \cancel{\delta K(x^j)} \cdots \delta K(x^n)} \right) \cdot \frac{\delta K(x)\chi(\phi(x))}{\delta K(x^j)} \\ &= i^{n-1} \sum_{j=1}^{n} \delta(x - x^j) \, \phi(x^1) \cdots \cancel{\phi(x^j)} \cdots \phi(x^n) \, \chi(\phi(x^j)). \end{aligned}$$

Putting back this results into the n-time differentiated version of (1.23) yields, at the level of the path integral,

$$\begin{aligned} i \int \mathcal{D}\phi \, e^{iS[\phi]} \, \phi(x^1) \cdots \phi(x^n) \, \partial_\mu J^\mu(x) = \sum_{j=1}^{n} \delta(x - x^j) \quad & (1.25) \\ \times \int \mathcal{D}\phi \, e^{iS[\phi]} \, \phi(x^1) \cdots \cancel{\phi(x^j)} \cdots \phi(x^n) \, \chi(\phi(x^j)). & \end{aligned}$$

Finally, the path integral with source K generates *time-ordered* vacuum correlators. Differentiating the functional identity (1.23) with respect to K therefore yields a relation for T-ordered correlators.

1.3 Spontaneous Symmetry Breaking

There are two common types of symmetry breaking in particle physics. In spontaneous symmetry breaking, the Lagrangian is invariant under a symmetry transformation, while the ground state (vacuum) is not. In explicit symmetry breaking, the Lagrangian contains terms that are not invariant under the symmetry.

1.3.1 Symmetry transformations and generators

A symmetry transformation U is defined by the invariance of the Hamiltonian,

$$U^\dagger H U = H, \tag{1.26}$$

where the symmetry transformation is $U(\theta) = \exp(i\theta^a T^a)$. A state is invariant under U (up to a phase) if

$$U|\psi\rangle = e^{i\alpha}|\psi\rangle, \qquad \alpha \in \mathbb{R}. \tag{1.27}$$

For simplicity set $\alpha = 0$. If the vacuum is invariant under U, then the (unbroken) symmetry generators annihilate the vacuum

$$\begin{aligned} U|0\rangle = |0\rangle \; &\rightarrow [I + i\theta^a T^a + \cdots]|0\rangle \approx |0\rangle + i\theta^a T^a|0\rangle \\ &\Rightarrow T^a|0\rangle = 0 \end{aligned} \tag{1.28}$$

If the vacuum is not invariant, then at least one generator does not annihilate the vacuum

$$U|0\rangle \neq |0\rangle \quad \Rightarrow \quad \exists\, a \text{ such that } T^a|0\rangle \neq 0. \tag{1.29}$$

1.3.2 Spontaneous Breaking of Discrete Symmetry

Let us consider a real scalar field ϕ with the following Lagrangian,

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi), \quad V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4, \tag{1.30}$$

where $\lambda > 0$. The Lagrangian is invariant under a discrete $\mathbb{Z}_2$ transformation,

$$\mathbb{Z}_2 : \phi \rightarrow -\phi. \tag{1.31}$$

To find whether this symmetry is spontaneously broken, we need to find the ground state. It corresponds to the point(s) in field space at which the potential in (1.30) is minimized. At the classical level, the stationary points satisfy $V'(\phi) = 0$,

$$\phi^2 = \begin{cases} 0, & m^2 > 0, \\ \frac{-m^2}{\lambda}, & m^2 < 0. \end{cases} \tag{1.32}$$

If $m^2 > 0$, then there is a single minimum given by

$$\phi = v_0 = 0, \quad (m^2 > 0). \tag{1.33}$$

The vacuum expectation value (VEV) in the vacuum state $|\Omega_0\rangle$ is defined by

$$v_0 \equiv \langle \Omega_0 | \phi | \Omega_0 \rangle. \tag{1.34}$$

This vacuum is invariant under $\mathbb{Z}_2$, so the symmetry is not spontaneously broken. On the other hand if $m^2 < 0$ they are two separate minima given by

$$\phi = v_{\pm} = \pm\sqrt{\frac{-m^2}{\lambda}}. \tag{1.35}$$

These correspond to two distinct ground states $|\Omega_+\rangle$ and $|\Omega_-\rangle$. They have the same energy and are therefore degenerate. A $\mathbb{Z}_2$ transformation maps $|\Omega_+\rangle$ to $|\Omega_-\rangle$ and vice versa. So the ground state is not invariant under the $\mathbb{Z}_2$ transformation anymore, and one says that the $\mathbb{Z}_2$ symmetry is spontaneously broken.

1.3.3 Spontaneous Breaking of Continuous Symmetry

Now we consider a Lagrangian of complex scalar field Φ,

$$\mathcal{L} = (\partial^\mu \Phi)^* (\partial_\mu \Phi) - V(\Phi), \quad V(\Phi) = m^2 |\Phi|^2 + \frac{\lambda}{4} |\Phi|^4. \tag{1.36}$$

This Lagrangian is invariant under the global $U(1)$ transformation

$$U(1) : \Phi \to \Phi' = e^{i\alpha} \Phi, \quad \alpha \in \mathbb{R}, \tag{1.37}$$

where α is spacetime-independent (a constant), and is therefore called a global parameter. This is a continuous transformation because varying α continuously connects any group element to the identity.

The potential in (1.36) is minimized at

$$|\Phi|^2 = \begin{cases} 0, & m^2 > 0, \\ \frac{-2m^2}{\lambda}, & m^2 < 0. \end{cases} \tag{1.38}$$

In the first case, there is a single ground state $|\Omega_0\rangle$ which is invariant under $U(1)$, so the symmetry is not spontaneously broken. In contrast, for $m^2 < 0$ the minima form a circle $|\Phi| = \sqrt{-2m^2/\lambda}$, leading to a continuum of degenerate vacua related by $U(1)$.

2 Chiral anomaly

When deriving the Ward identity, our only non-trivial assumption was the invariance of the path integral measure. Therefore, we expect that anomalies are tied to the (non-)invariance of the path integral measure.

Consider a massless Dirac fermion ψ coupled to electromagnetism. The action is

$$S = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\bar{\psi} \gamma^\mu D_\mu \psi \right], \tag{2.1}$$

where D_μ is the gauge-covariant derivative,

$$D_\mu = \partial_\mu - iqA_\mu, \tag{2.2}$$

where q is the EM charge of the fermion normalized such that $q = -1$ for the electron. In this notation, the fermion ψ transforms under the $U(1)_{\text{EM}}$ gauge transformation as

$$\psi \to e^{iq\alpha(x)}\psi. \tag{2.3}$$

This theory also has two global phase symmetries

$$V : \psi \to e^{i\epsilon}\psi, \qquad J^\mu_V = \bar{\psi}\gamma^\mu\psi \tag{2.4}$$
$$A : \psi \to e^{i\epsilon\gamma^5}\psi, \quad J^\mu_A = \bar{\psi}\gamma^\mu\gamma^5\psi. \tag{2.5}$$

Noether currents

Under an infinitesimal vector transformation with real parameter ϵ, ψ and $\bar{\psi}$ transform as

$$\psi' = e^{i\epsilon}\psi = [1 + i\epsilon + \mathcal{O}(\epsilon^2)]\psi \approx \psi + i\epsilon\psi, \tag{2.6}$$

and

$$\bar{\psi}' = (\psi')^\dagger\gamma^0 \approx [\psi + i\epsilon\psi]^\dagger\gamma^0 = [\psi^\dagger - i\epsilon\psi^\dagger]\gamma^0 = \bar{\psi} - i\epsilon\bar{\psi}. \tag{2.7}$$

An infinitesimal axial transformation is

$$\psi' = e^{i\epsilon\gamma^5}\psi = [1 + i\epsilon\gamma^5 + \mathcal{O}(\epsilon^2)]\psi \approx \psi + i\epsilon\gamma^5\psi, \tag{2.8}$$

and

$$\begin{aligned}\bar{\psi}' = (\psi')^\dagger\gamma^0 \approx [\psi + i\epsilon\gamma^5\psi]^\dagger\gamma^0 = [\psi^\dagger - i\epsilon\psi^\dagger(\gamma^5)^\dagger]\gamma^0 &= \psi^\dagger\gamma^0 - i\epsilon\psi^\dagger\gamma^5\gamma^0 \\ &= \psi^\dagger\gamma^0 + i\epsilon\psi^\dagger\gamma^0\gamma^5 \\ &= \bar{\psi} + i\epsilon\bar{\psi}\gamma^5,\end{aligned} \tag{2.9}$$

where we use $\{\gamma^\mu, \gamma^5\} = 0$ and $(\gamma^5)^\dagger = \gamma^5$.

Classically, (for a massless Dirac fermion), Noether theorem tell us

$$\partial_\mu J^\mu_V = 0, \quad \partial_\mu J^\mu_A = 0. \tag{2.10}$$

We want to study which of these classical statements survive in the quantum theory.

In the previous section, we assumed that the path integral measure is invariant under the symmetry transformation in the quantum theory. In this case we are interested in the transformation of the fermion measure

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi, \tag{2.11}$$

under the vector and axial transformation. From (2.6)–(2.9), the vector transformation acts with opposite phases on ψ and $\bar{\psi}$, while the axial transformation acts with the same sign. This difference is reflected in the Jacobian of the fermionic path-integral measure and is the origin of the axial anomaly.

2.1 Euclidean path integral

For anomaly calculations, it is convenient to Wick rotate to Euclidean time by defining

$$x_E^0 \equiv ix^0 = it, \tag{2.12}$$

where the subscript E denotes the Euclidean coordinates. Then this transformation requires the introduction of the corresponding components of the partial derivative and the vector potential,

$$\partial_{E0} \equiv \frac{\partial}{\partial x_E^0} = -i\frac{\partial}{\partial x^0} = -i\partial_0, \qquad A_{E0} \equiv -iA_0. \tag{2.13}$$

The metric becomes

$$\eta_{\mu\nu}dx^\mu dx^\nu \quad \to \quad -\delta_{\mu\nu}dx_E^\mu dx_E^\nu. \tag{2.14}$$

Then due to this transformation, the Clifford algebra should be modified,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}, \quad \{\gamma_E^\mu, \gamma_E^\nu\} = -2\delta^{\mu\nu}, \quad \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma_E^0\gamma_E^1\gamma_E^2\gamma_E^3 = \gamma_E^5. \tag{2.15}$$

With the convention in (2.15), the Euclidean gamma matrices are anti-hermitian,

$$[\gamma_E^\mu]^\dagger = -\gamma_E^\mu. \tag{2.16}$$

However by the definition, the γ_E^5 remains Hermitian,

$$[\gamma_E^5]^\dagger = [\gamma_E^0\gamma_E^1\gamma_E^2\gamma_E^3]^\dagger = \gamma_E^3\gamma_E^2\gamma_E^1\gamma_E^0 = \gamma_E^0\gamma_E^1\gamma_E^2\gamma_E^3 = \gamma_E^5. \tag{2.17}$$

Under the Wick rotation, the path-integral weight e^{iS} becomes e^{-S_E}, where S_E is the Euclidean action obtained by analytic continuation, $S_E \equiv -iS\big|_{x^0=-ix_E^0}$.

Gamma matrix Hermitianity and metric signature

Consider the Dirac equation as Schrodinger form

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \quad \to \quad i\frac{\partial\psi}{\partial t} = [-i\gamma^0\gamma^i\partial_i + m\gamma^0]\psi \equiv H_D\psi, \tag{2.18}$$

where we use $(\gamma^0)^2 = \eta^{00} = 1$. To preserve the inner product,

$$\frac{d}{dt}||\psi||^2 = \frac{d}{dt}\int d^3x\, \psi^\dagger(x)\psi(x) = i\int d^3x\, \psi^\dagger(H_D - H_D^\dagger)\psi \overset{!}{=} 0, \tag{2.19}$$

the Hamiltonian H_D is Hermitian. More specifically,

$$H_D^\dagger = -i(\gamma^0\gamma^i)^\dagger\partial_i + m(\gamma^0)^\dagger = -i\gamma^0\gamma^i\partial_i + m\gamma^0 \overset{!}{=} H_D, \tag{2.20}$$

where we use $(\partial_i)^\dagger = -\partial_i$. Then

$$(\gamma^0)^\dagger = \gamma^0, \quad (\gamma^0\gamma^i)^\dagger = \gamma^0\gamma^i. \tag{2.21}$$

Remark. In this note we use the mostly-minus convention $\eta^{00} = +1$. If one instead uses the mostly-plus convention $\eta^{00} = -1$, it is common to choose a different Hermiticity convention for the gamma matrices and/or a modified definition of the Dirac adjoint so that the Hamiltonian is Hermitian with respect to the chosen inner product.

2.1.1 Fermion measure

Remark. From here on we work in Euclidean space. Depending on the context, I will either keep the subscript E explicit or drop it for notational simplicity. I ask for the reader's kind understanding.

We want to know how the fermion measure is modified under the transformation given in (2.4) and (2.5). For practical calculation, we need to properly define the measure. Consider the following eigenvalue equation,

$$\not{D}_E|n\rangle = \lambda_n|n\rangle \quad \to \quad \not{D}_E\psi_n(x) = \lambda_n\psi_n(x), \tag{2.22}$$

where $\{\lambda_n\}$ are the eigenvalues and $\{\psi_n\}$ are the eigenspinors. Note that in Euclidean space, the Dirac operator

$$\not{D}_E \equiv \gamma_E^\mu D_{E\mu} \to \not{D}_E^\dagger = (D_{E\mu})^\dagger(\gamma_E^\mu)^\dagger = (\text{anti-Hermitian})\cdot(\text{anti-Hermitian}) = \not{D}_E \tag{2.23}$$

meaning $\not{D}_E$ is Hermitian. This implies that the eigenvalues $\{\lambda_n\}$ are real, and the eigenspinors are orthonormal,

$$\langle n|m\rangle = \delta_{nm} \quad \to \quad \int d^4x_E\,\langle n|x\rangle\langle x|m\rangle = \int d^4x_E\,\psi_n^\dagger(x_E)\psi_m(x_E) = \delta_{nm}, \tag{2.24}$$

and the complete basis

$$\sum_n |n\rangle\langle n| = 1 \quad \to \quad \delta^4(x-y)\mathbb{I}_{\text{spinor}} = \langle x|y\rangle = \sum_n \langle x|n\rangle\langle n|y\rangle = \sum_n \psi_n(x)\psi_n^\dagger(y). \tag{2.25}$$

Then a general Dirac spinor ψ can be expanded in the eigenspinor basis as

$$\psi(x) = \langle x|\psi\rangle = \sum_n \langle x|n\rangle\langle n|\psi\rangle = \sum_n \langle n|\psi\rangle\langle x|n\rangle = \sum_n a_n\psi_n(x), \tag{2.26}$$

$$\bar\psi(x) = \langle\bar\psi|x\rangle = \sum_n \langle\bar\psi|n\rangle\langle n|x\rangle = \sum_n \langle n|x\rangle\langle\bar\psi|n\rangle = \sum_n \psi_n^\dagger(x)\bar b_n \tag{2.27}$$

where $\{a_n\}$ and $\{\bar b_n\}$ are Grassmann-valued numbers so that the spinors can satisfy the anti-commutation relations. In Euclidean space, ψ and $\bar\psi$ are treated as independent Grassmann fields. Note that $\langle n|\psi\rangle$ is Grassmann-valued while $\langle x|n\rangle$ is a c-number, so they commute and no extra sign arises when reordering them. Therefore, we can write the Euclidean action for the Dirac fermions as

$$\begin{aligned} S_E = \int d^4x_E\, \bar{\psi}(x) \not{D}_E \psi(x) &= \int d^4x_E\, \langle \bar{\psi} | x_E \rangle \langle x_E | \not{D}_E | \psi \rangle \\ &= \langle \bar{\psi} | \not{D}_E | \psi \rangle. \end{aligned} \tag{2.28}$$

Substituting the complete basis (2.25),

$$\begin{aligned} S_E = \langle \bar{\psi} | \not{D}_E | \psi \rangle &= \sum_{n,m} \langle \bar{\psi} | m \rangle \langle m | \not{D}_E | n \rangle \langle n | \psi \rangle \\ &= \sum_{n,m} \lambda_n \langle \bar{\psi} | m \rangle \langle m | n \rangle \langle n | \psi \rangle \\ &= \sum_n \lambda_n \langle \bar{\psi} | n \rangle \langle n | \psi \rangle = \sum_n \lambda_n \bar{b}_n a_n. \end{aligned} \tag{2.29}$$

Also we can write the fermion measure as the coefficient a_n and $\bar{b}_n$,

$$\begin{aligned} \int \mathcal{D}\bar{\psi}\mathcal{D}\psi &= \int \left[\prod_x d\bar{\psi}(x)\, d\psi(x) \right] \\ &= \int \left[\frac{1}{\det U^\dagger \, \det U} \prod_n d\bar{b}_n\, da_n \right], \qquad U_{xn} \equiv \langle x | n \rangle, \\ &= \int \prod_n d\bar{b}_n\, da_n. \end{aligned} \tag{2.30}$$

Derivation

From the (2.26) and (2.27), one can define the unitary transformation between the position and spectral basis of $\not{D}_E$,

$$\psi(x) = \sum_n \langle n | \psi \rangle \langle x | n \rangle = \sum_n a_n U_{xn}, \tag{2.31}$$

$$\bar{\psi}(x) = \sum_n \langle n | x \rangle \langle \bar{\psi} | n \rangle = \sum_n (U_{xn})^\dagger \bar{b}_n, \tag{2.32}$$

where we discretize the spacetime coordinate. The unitarity directly comes from the completeness of each basis,

$$\begin{aligned} UU^\dagger &= \sum_n U_{xn} (U^\dagger)_{ny} \\ &= \sum_n U_{xn} (U_{yn})^\dagger = \sum_n \langle x | n \rangle \langle y | n \rangle^\dagger = \langle x | y \rangle = \delta_{xy}, \end{aligned} \tag{2.33}$$

$$\begin{aligned} U^\dagger U &= \sum_x (U^\dagger)_{nx} U_{xm} \\ &= \sum_x (U_{xn})^\dagger U_{xm} = \sum_x \langle x | n \rangle^\dagger \langle x | m \rangle = \langle n | m \rangle = \delta_{nm}. \end{aligned} \tag{2.34}$$

Thus, the matrix U is an unitary matrix

$$UU^\dagger = U^\dagger U = I. \tag{2.35}$$

Consider a linear transformation of Grassmann vector,[a]

$$\theta = M\xi \quad \Leftrightarrow \quad \theta_i = \sum_{j=1}^{N} M_{ij}\xi_j. \tag{2.36}$$

Then

$$\prod_{i=1}^{N} d\theta_i = (\det M)^{-1} \prod_{i=1}^{N} d\xi_i. \tag{2.37}$$

Thus in our case

$$\prod_x d\bar{\psi}(x)\, d\psi(x) = \frac{1}{\det U \cdot \det U^\dagger} \prod_n d\bar{b}_n\, da_n = \prod_n d\bar{b}_n\, da_n. \tag{2.38}$$

[a]More detail about a Grassmann number, see the Appendix.

Therefore the Euclidean partition function for the fermions is

$$\begin{aligned}
Z_E = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{-S_E} &= \int \prod_n d\bar{b}_n\, da_n \, \exp\left[-\sum_m \lambda_m \bar{b}_m a_m\right] \\
&= \int \prod_n d\bar{b}_n\, da_n \prod_m \exp\left[-\lambda_m \bar{b}_m a_m\right] \\
&= \int \prod_n d\bar{b}_n\, da_n \prod_m \left[1 - \lambda_m \bar{b}_m a_m\right] \\
&= \prod_n \lambda_n = \det[\not{D}_E].
\end{aligned} \tag{2.39}$$

Note that this expression is exact. See the Appendix of Grassmann variable.

2.2 Calculating the Jacobian

Let us compute how the fermionic path-integral measure changes under vector and axial rotations. We start with the axial rotation. It is sufficient to consider the transformation of ψ since the transformation of $\bar{\psi}$ is identical for axial rotation and differs by a sign for vector rotations.

Under an axial rotation

$$\delta\psi(x) = i\epsilon(x)\gamma^5\psi(x) \quad \Leftrightarrow \quad \sum_k (\delta a_k)\psi_k(x) = i\epsilon(x)\sum_m a_m \gamma^5 \psi_m(x). \tag{2.40}$$

By applying the orthogonality relation (2.24)

$$
\begin{aligned}
\int d^4x_E\, \psi_n^\dagger(x_E) \sum_k (\delta a_k)\psi_k(x_E) &= \sum_k (\delta a_k) \int d^4x_E\, \psi_n^\dagger(x_E)\psi_k(x_E) \\
&= \sum_k (\delta a_k)\delta_{nk} \\
&= \delta a_n = i \int d^4x_E\, \epsilon(x_E)\psi_n^\dagger(x_E) \sum_m \gamma^5 \psi_m(x_E) a_m \\
&= \sum_m i \int d^4x_E\, \epsilon(x_E)\psi_n^\dagger(x_E)\gamma^5\psi_m(x_E) a_m \equiv \sum_m X_{nm} a_m.
\end{aligned}
\tag{2.41}
$$

where we defined

$$
X_{nm} \equiv i \int d^4x_E\, \epsilon(x_E)\, \psi_n^\dagger(x_E)\gamma^5\psi_m(x_E). \tag{2.42}
$$

One can rewrite as a matrix form

$$
\mathbf{a} \to \mathbf{a}' = \mathbf{a} + \delta\mathbf{a} = (1+X)\mathbf{a}, \tag{2.43}
$$

where

$$
\psi(x) = \sum_n a_n\, \psi_n(x), \qquad \mathbf{a} \equiv (a_1, a_2, \ldots)^T. \tag{2.44}
$$

For Grassmann variables, the Jacobian $\mathcal{J}$ of this transformation is[1]

$$
\prod_n da_n' = \big[\det(1+X)\big]^{-1} \prod_n da_n \equiv \mathcal{J} \prod_n da_n. \tag{2.45}
$$

Equivalently,

$$
\prod_n da_n = \det(1+X) \prod_n da_n' = \mathcal{J}^{-1} \prod_n da_n'. \tag{2.46}
$$

At leading order in ϵ, we can approximate the Jacobian as

$$
\begin{aligned}
\mathcal{J} = [\det(1+X)]^{-1} = \det[(1+X)^{-1}] &= \det[1 - X + \mathcal{O}(X^2)] \\
&\approx \det(1-X).
\end{aligned}
\tag{2.47}
$$

Using the fact $\det(1+A) = \exp[\mathrm{Tr}\{\ln(1+A)\}]$,

$$
\begin{aligned}
\det(1-X) = \exp[\mathrm{Tr}\{\ln(1-X)\}] &= \exp[\mathrm{Tr}\{-X + \mathcal{O}(X^2)\}] \\
&= \exp[-\mathrm{Tr}(X)].
\end{aligned}
\tag{2.48}
$$

[1]For regular c-number, $\mathcal{J} = \det(1+X)$.

Proof of det and exp(Tr ln)

Let M be an $N \times N$ matrix with eigenvalues $\{\mu_i\}_{i=1}^N$. Then

$$\det M = \prod_{i=1}^{N} \mu_i, \qquad \mathrm{Tr}[\ln M] = \sum_{i=1}^{N} \ln \mu_i, \tag{2.49}$$

so

$$\det M = \exp\left(\sum_{i=1}^{N} \ln \mu_i\right) = \exp\bigl(\mathrm{Tr}\ \ln M\bigr). \tag{2.50}$$

In particular, if X is small (here $X = \mathcal{O}(\epsilon)$), we may use the series

$$\ln(1+X) = X - \frac{1}{2}X^2 + \cdots, \tag{2.51}$$

so to leading order

$$\mathrm{Tr}\ \ln(1+X) \simeq \mathrm{Tr}\, X. \tag{2.52}$$

Explicitly,

$$\begin{aligned} \mathcal{J} = e^{-\mathrm{Tr}(X)} &= \exp - \left[\sum_n X_{nn}\right] \\ &= \exp\left[-i \sum_n \int d^4x_E\, \epsilon(x_E) \psi_n^\dagger(x_E) \gamma^5 \psi_n(x_E)\right] \\ &= \exp\left[-i \int d^4x_E\, \epsilon(x_E) \sum_n \psi_n^\dagger(x_E) \gamma^5 \psi_n(x_E)\right]. \end{aligned} \tag{2.53}$$

For $\bar{\psi}$, the transformation is the same so

$$\mathcal{J}^2 \int \prod_n d\bar{b}_n da_n = \int \prod_n d\bar{b}'_n da'_n. \tag{2.54}$$

Transformation for $\bar{\psi}$

In Euclidean space we treat ψ and $\bar{\psi}$ as independent Grassmann fields. Therefore we *define* the infinitesimal axial rotation by

$$\delta\psi(x_E) = i\epsilon(x_E)\gamma^5\psi(x_E), \qquad \delta\bar{\psi}(x_E) = i\epsilon(x_E)\bar{\psi}(x_E)\gamma^5. \tag{2.55}$$

Then using the decomposition

$$\delta\bar{\psi}(x_E) = i\epsilon(x_E)\bar{\psi}(x_E)\gamma^5 \quad \Leftrightarrow \quad \sum_k (\delta\bar{b}_k)\psi_k^\dagger(x) = i\epsilon(x)\sum_m \bar{b}_m \psi_m^\dagger(x)\gamma^5. \tag{2.56}$$

By applying the orthogonality relation (2.24)

$$\begin{aligned}
\int d^4x_E \sum_k (\delta\bar{b}_k)\psi_k^\dagger(x)\psi_n(x) &= \sum_k (\delta\bar{b}_k)\int d^4x_E\, \psi_k^\dagger(x)\psi_n(x) \\
&= \sum_k (\delta\bar{b}_k)\delta_{kn} \\
&= \delta\bar{b}_n = i\int d^4x_E\,\epsilon(x)\sum_m \bar{b}_m\psi_m^\dagger(x)\gamma^5\psi_n(x) \\
&= \sum_m \bar{b}_m \int d^4x_E\, i\epsilon(x)\psi_m^\dagger(x)\gamma^5\psi_n(x) \equiv \sum_m \bar{b}_m X_{mn}.
\end{aligned} \tag{2.57}$$

One can rewrite as a matrix form

$$\bar{\mathbf{b}} \to \bar{\mathbf{b}}' = \bar{\mathbf{b}} + \delta\bar{\mathbf{b}} = \bar{\mathbf{b}}(1+X), \tag{2.58}$$

where

$$\bar{\psi}(x) = \sum_n \bar{b}_n\,\psi_n^\dagger(x), \qquad \bar{\mathbf{b}} \equiv (\bar{b}_1, \bar{b}_2, \ldots). \tag{2.59}$$

Change the column vector as follows

$$\begin{aligned}
[\bar{b}']^T = (1+X^T)\bar{b}^T \quad \to \quad \prod_n d\bar{b}'_n &= [\det{(1+X^T)}]^{-1}\prod_n d\bar{b}_n \\
&= [\det{(1+X)}]^{-1}\prod_n d\bar{b}_n.
\end{aligned} \tag{2.60}$$

Therefore

$$\mathcal{J}_{\bar{\psi}} = [\det{(1+X)}]^{-1}. \tag{2.61}$$

Let us briefly discuss what will be different for vector transformations. Following similar steps it is straightforward to show that for the vector transformation, the Jacobian factors for ψ and $\bar{\psi}$ are

$$\mathcal{J}_\psi^{(V)} = [\det{(1+Y)}]^{-1}, \quad \mathcal{J}_{\bar{\psi}}^{(V)} = [\det{(1-Y)}]^{-1}, \tag{2.62}$$

where Y is same as X except that the γ^5 term is absent,

$$Y_{nm} = i\int d^4x_E\,\epsilon(x)\psi_n^\dagger(x)\psi_m(x). \tag{2.63}$$

Vector transformation derivation

Under the vector rotation

$$\begin{cases} \delta\psi = i\epsilon\psi \\ \delta\bar{\psi} = -i\epsilon\bar{\psi} \end{cases} \quad \rightarrow \quad \begin{cases} \sum_k (\delta a_k)\psi_k(x) = i\epsilon(x)\sum_m a_m\psi_m(x) \\ \sum_k (\delta\bar{b}_k)\psi_k^\dagger(x) = -i\epsilon(x)\sum_m \bar{b}_m\psi_m^\dagger(x). \end{cases} \tag{2.64}$$

By applying the orthogonality relation (2.24), we do the same process without γ^5 factor, (2.41) for ψ and (2.57) for $\bar{\psi}$. Therefore we can define

$$Y_{nm} = i\int d^4x_E\, \epsilon(x)\psi_n^\dagger(x)\psi_m(x). \tag{2.65}$$

One can rewrite as a matrix form

$$\bar{\mathbf{b}} \rightarrow \bar{\mathbf{b}}' = \bar{\mathbf{b}} + \delta\bar{\mathbf{b}} = \bar{\mathbf{b}}(1-Y), \tag{2.66}$$

where

$$\bar{\psi}(x) = \sum_n \bar{b}_n\,\psi_n^\dagger(x), \qquad \bar{\mathbf{b}} \equiv (\bar{b}_1, \bar{b}_2, \ldots). \tag{2.67}$$

Change the column vector as follows

$$\begin{aligned} [\bar{b}']^T = (1-Y^T)\bar{b}^T \quad \rightarrow \quad \prod_n d\bar{b}'_n &= [\det(1-Y^T)]^{-1}\prod_n d\bar{b}_n \\ &= [\det(1-Y)]^{-1}\prod_n d\bar{b}_n. \end{aligned} \tag{2.68}$$

Therefore

$$\mathcal{J}_{\bar{\psi}}^{(V)} = [\det(1-Y)]^{-1}. \tag{2.69}$$

For ψ case, it follows the same logic from (2.43) to (2.45),then we get

$$\mathcal{J}_{\psi}^{(V)} = [\det(1+Y)]^{-1}. \tag{2.70}$$

The opposite sign is a direct consequence of the opposite sign in the infinitesimal vector transformation in (2.7). As a result, the Jacobian factors from ψ and $\bar{\psi}$ cancel at linear order in $\epsilon(x)$,

$$\mathcal{J}^{(V)} \equiv \mathcal{J}_{\psi}^{(V)}\,\mathcal{J}_{\bar{\psi}}^{(V)} = [\det(1+Y)]^{-1}[\det(1-Y)]^{-1} = 1 + \mathcal{O}(\epsilon^2). \tag{2.71}$$

Therefore the vector symmetry receives no anomalous contribution in the Ward-identity derivation, which keeps only terms linear in the arbitrary function $\epsilon(x)$.

Derivation

Using $\det A = \exp(\mathrm{Tr} \ln A)$ we can rewrite

$$\begin{aligned}
\mathcal{J}^{(V)} = [\det(1+Y)]^{-1}[\det(1-Y)]^{-1} &= [\det\{(1+Y)(1-Y)\}]^{-1} \\
&= [\det(1-Y^2)]^{-1} \\
&= \det[(1-Y^2)^{-1}] \\
&= \exp\left[\mathrm{Tr}\{\ln(1-Y^2)^{-1}\}\right] \\
&= \exp\left[-\mathrm{Tr}\{\ln(1-Y^2)\}\right]. \qquad (2.72)
\end{aligned}$$

Since $Y = \mathcal{O}(\epsilon)$, we have $Y^2 = \mathcal{O}(\epsilon^2)$ and may expand

$$\ln(1-Y^2) = -Y^2 - \frac{1}{2}Y^4 + \cdots . \qquad (2.73)$$

Substituting (2.73) into (2.72) gives

$$\mathcal{J}^{(V)} = \exp\left[\mathrm{Tr}(Y^2) + \mathcal{O}(Y^4)\right] = 1 + \mathrm{Tr}(Y^2) + \mathcal{O}(Y^4). \qquad (2.74)$$

In particular, there is no term linear in $\epsilon(x)$.

Let us return to the evaluation of the Jacobian given in (2.53). To calculate this term, we need to regularize the infinite sum. We need to do this in a gauge invariant way. Since the eigenvalues $\{\lambda_n\}$ of the Dirac operator are gauge invariant, we can use them for regularization. So we write

$$\begin{aligned}
\int d^4x_E\, \epsilon(x) \sum_n \psi_n^\dagger \gamma^5 \psi_n &= \lim_{\Lambda\to\infty} \int d^4x_E\, \epsilon(x) \sum_n \psi_n^\dagger(x_E)\gamma^5\psi_n(x_E) e^{-\lambda_n^2/\Lambda^2} \\
&= \lim_{\Lambda\to\infty} \int d^4x_E\, \epsilon(x) \sum_n \psi_n^\dagger(x_E)\gamma^5 e^{-\lambda_n^2/\Lambda^2}\psi_n(x_E) \\
&= \lim_{\Lambda\to\infty} \int d^4x_E\, \epsilon(x) \sum_n \psi_n^\dagger(x_E)\gamma^5 e^{-\not{D}_E^2/\Lambda^2}\psi_n(x_E) \\
&\equiv \lim_{\Lambda\to\infty} \int d^4x_E\, \epsilon(x_E)\mathcal{W}_\Lambda(x_E), \qquad (2.75)
\end{aligned}$$

where we use the eigenvalue equation (2.22). Now we take the Fourier transform of ψ_n via

$$\psi_n(x_E) = \int \frac{d^4k_E}{(2\pi)^4} e^{ik_E x}\psi_n(k_E), \qquad (2.76)$$

so that $\mathcal{W}_\Lambda$ becomes

$$\mathcal{W}_\Lambda = \sum_n \int \frac{d^4k_E}{(2\pi)^4}\frac{d^4k'_E}{(2\pi)^4} \psi_n^\dagger(k_E) e^{-ik_E x}\gamma^5 e^{-\not{D}_E^2/\Lambda^2} e^{ik'_E x}\psi_n(k'_E). \qquad (2.77)$$

Using the completeness of the Dirac-operator eigenmodes and inserting momentum eigenstates,

$$\sum_n \psi_n(k_E)\psi_n^\dagger(k'_E) = \sum_n \langle k_E|n\rangle\langle n|k'_E\rangle$$
$$= \langle k_E|k'_E\rangle = (2\pi)^4\delta^4(k_E - k'_E)\mathbb{I}_{\text{spinor}}. \tag{2.78}$$

Then we get

$$\begin{aligned} \mathcal{W}_\Lambda &= \int \frac{d^4k_E}{(2\pi)^4}\frac{d^4k'_E}{(2\pi)^4}\psi_n^\dagger(k_E)\underbrace{e^{-ik_Ex}\gamma^5 e^{-\not{D}_E^2/\Lambda^2}e^{ik'_Ex}}_{\hat{O}(k_E,k'_E)}\psi_n(k'_E) \\ &= \int \frac{d^4k_E}{(2\pi)^4}\frac{d^4k'_E}{(2\pi)^4}\text{tr}[\hat{O}(k_E,k'_E)](2\pi)^4\delta^4(k_E-k'_E) \\ &= \int \frac{d^4k_E}{(2\pi)^4}\text{tr}[e^{-ik_Ex}\gamma^5 e^{-\not{D}_E^2/\Lambda^2}e^{ik_Ex}]. \end{aligned} \tag{2.79}$$

Identity derivation

Consider the Euclidean momentum eigenstates $|k_E\rangle$. Using the completeness of the Dirac-operator eigenmodes and inserting momentum eigenstates,

$$\sum_n \psi_n(k_E)\psi_n^\dagger(k'_E) = \sum_n \langle k_E|n\rangle\langle n|k'_E\rangle$$
$$= \langle k_E|k'_E\rangle = (2\pi)^4\delta^4(k_E - k'_E)\mathbb{I}_{\text{spinor}}. \tag{2.80}$$

Therefore,

$$\begin{aligned} \sum_n \psi_n^\dagger(k_E)\hat{O}\,\psi_n(k'_E) &= \sum_n \text{tr}\Big[\psi_n^\dagger(k_E)\hat{O}\,\psi_n(k'_E)\Big] \\ &= \sum_n \text{tr}\Big[\hat{O}\,\psi_n(k'_E)\psi_n^\dagger(k_E)\Big] \\ &= \text{tr}\left[\hat{O}\sum_n \psi_n(k'_E)\psi_n^\dagger(k_E)\right] \\ &= (2\pi)^4\delta^{(4)}(k_E-k'_E)\ \text{tr}(\hat{O}), \end{aligned} \tag{2.81}$$

where tr is the trace over spinor indices. In the first line we inserted tr since the quantity is a scalar; this makes it possible to use cyclicity and the completeness relation.

Remark. More generally, consider arbitrary two vector u and v in vectorspace $\mathcal{V}$,

$$\begin{aligned} \text{Tr}_\mathcal{V}[\hat{O}|u\rangle\langle v|] = \sum_n \text{Tr}_\mathcal{V}[\hat{O}|u\rangle\langle v|n\rangle\langle n|] &= \sum_n \text{Tr}_\mathcal{V}[\langle n|\hat{O}|u\rangle\langle v|n\rangle] \\ &= \sum_n \langle n|\hat{O}|u\rangle\langle v|n\rangle \\ &= \sum_n \langle v|n\rangle\langle n|\hat{O}|u\rangle \\ &= \langle v|\hat{O}|u\rangle. \end{aligned} \tag{2.82}$$

Here $\mathrm{Tr}_{\mathcal{V}}$ denotes the trace over the vector space $\mathcal{V}$ on which $\hat{O}$ acts. We use tr for the trace over spinor indices, and $\mathrm{Tr}_{\mathcal{V}}$ for the trace over an abstract vector space $\mathcal{V}$.

One can calculate the trace part as follows,

$$
\begin{aligned}
e^{-ik_E x} e^{-\not{D}_E^2/\Lambda^2} e^{ik_E x} &= e^{-ik_E x} \exp\left[-\frac{1}{\Lambda^2}\left\{D_E^\mu D_{E\mu} - \frac{ie}{2}\gamma_E^\mu \gamma_E^\nu F_{\mu\nu}\right\}\right] e^{ik_E x} \\
&= \exp\left[-\frac{1}{\Lambda^2}\left\{(D_E^\mu + ik_E^\mu)(D_{E\mu} + ik_{E\mu}) - \frac{ie}{2}\gamma_E^\mu \gamma_E^\nu F_{\mu\nu}\right\}\right]. \quad (2.83)
\end{aligned}
$$

Identities

$$
\begin{aligned}
\not{D}_E^2 &= \gamma_E^\mu \gamma_E^\nu D_{E\mu} D_{E\nu} \\
&= \left(\frac{1}{2}\{\gamma_E^\mu, \gamma_E^\nu\} + \frac{1}{2}[\gamma_E^\mu, \gamma_E^\nu]\right) D_{E\mu} D_{E\nu} \\
&= \frac{1}{2}\{\gamma_E^\mu, \gamma_E^\nu\} D_\mu D_\nu + \frac{1}{2}[\gamma_E^\mu, \gamma_E^\nu] D_{E\mu} D_{E\nu} \\
&= D_E^\mu D_{E\mu} + \frac{1}{4}[\gamma_E^\mu, \gamma_E^\nu]\left([D_{E\mu}, D_{E\nu}] + \{D_{E\mu}, D_{E\nu}\}\right) \\
&= D_E^\mu D_{E\mu} + \frac{1}{4}[\gamma_E^\mu, \gamma_E^\nu][D_{E\mu}, D_{E\nu}], \qquad (\text{anti-sym} \cdot \text{sym} = 0) \\
&= D_E^\mu D_{E\mu} - \frac{ie}{4}\left(\gamma_E^\mu \gamma_E^\nu - \gamma_E^\nu \gamma_E^\mu\right) F_{\mu\nu} \\
&= D_E^\mu D_{E\mu} - \frac{ie}{4}\left(\gamma_E^\mu \gamma_E^\nu F_{\mu\nu} - \gamma_E^\nu \gamma_E^\mu F_{\mu\nu}\right) \\
&= D_E^\mu D_{E\mu} - \frac{ie}{4}\left(\gamma_E^\mu \gamma_E^\nu F_{\mu\nu} + \gamma_E^\nu \gamma_E^\mu F_{\nu\mu}\right), \qquad F_{\mu\nu} = -F_{\nu\mu} \\
&= D_E^\mu D_{E\mu} - \frac{ie}{2}\gamma_E^\mu \gamma_E^\nu F_{\mu\nu}. \qquad (2.84)
\end{aligned}
$$

And

$$
\begin{aligned}
e^{-ik_E x} D_\mu e^{ik_E x} &= e^{-ik_E x}(\partial_\mu - ieA_\mu) e^{ik_E x} \\
&= e^{-ik_E x}(\partial_\mu) e^{ik_E x} - ieA_\mu \\
&= \partial_\mu - ieA_\mu + ik_{E\mu} = D_\mu + ik_{E\mu}. \qquad (2.85)
\end{aligned}
$$

Introduce the dimensionless momentum $\tilde{k}_E^\mu \equiv -k_E^\mu/\Lambda$ with $k_E^2 = g_E^{\mu\nu} k_{E\mu} k_{E\nu} = -\delta^{\mu\nu} k_{E\mu} k_{E\nu}$.

Then the trace part becomes

$$
\begin{aligned}
e^{-ik_E x}e^{-\not{D}_E^2/\Lambda^2}e^{ik_E x} &= \exp\left[-\frac{1}{\Lambda^2}\left\{(D_E^\mu + ik_E^\mu)(D_{E\mu} + ik_{E\mu}) - \frac{ie}{2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right\}\right] \\
&= \exp\left[-\frac{(D_E^\mu D_{E\mu} + 2ik_E^\mu D_{E\mu} - k_E^\mu k_{E\mu})}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right] \\
&= \exp\left[\frac{k_E^\mu k_{E\mu}}{\Lambda^2} - \frac{2i\tilde{k}_E^\mu D_{E\mu}}{\Lambda} - \frac{D_E^\mu D_{E\mu}}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right] \\
&= \exp\left[-\tilde{k}_E^2 - \frac{2i\tilde{k}_E^\mu D_{E\mu}}{\Lambda} - \frac{D_E^\mu D_{E\mu}}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right] \\
&= e^{-\tilde{k}_E^2}\cdot\exp\left[-\frac{2i\tilde{k}_E^\mu D_{E\mu}}{\Lambda} - \frac{D_E^\mu D_{E\mu}}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right]. \qquad (2.86)
\end{aligned}
$$

The last line of the above equation since it is proportional to the identity matrix which commutes with all the other elements in exponent. Therefore (2.79) becomes

$$
\mathcal{W}_\Lambda(x_E) = \Lambda^4\int\frac{d^4\tilde{k}_E}{(2\pi)^4}\,e^{-\tilde{k}_E^2}\cdot\mathrm{tr}\left[\gamma_E^5\exp\left(-\frac{2i\tilde{k}_E^\mu D_{E\mu}}{\Lambda} - \frac{D_E^\mu D_{E\mu}}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}(x_E)\right)\right] \qquad (2.87)
$$

Expanding the exponential will give a various power of Λ, however we need to keep only the terms that will survive once we take the $\Lambda\to\infty$ limit. Since we have the overall Λ^4 factor, we only need to keep the terms up to order Λ^{-4}.

We will also take the trace, so we also need to be aware of the gamma matrix structure,

$$
\mathrm{tr}\left(\gamma_E^5\gamma_E^\mu\gamma_E^\nu\gamma_E^\rho\gamma_E^\sigma\right) = -4\epsilon^{\mu\nu\rho\sigma}, \qquad (2.88)
$$

with the convention $\epsilon^{0123} = -1$.

Derivation and lower power

One can also exploit this formula when the power of gamma matrix lower than 4. For example

$$
\mathrm{Tr}\left(\gamma_E^5\gamma_E^\mu\gamma_E^\nu\right) = \mathrm{Tr}\left(-\gamma_E^5\gamma_E^\mu\gamma_E^\nu\gamma_E^\rho\gamma_E^\rho\right) = 4\epsilon^{\mu\nu\rho\rho} = 0. \qquad (2.89)
$$

Thus, only term from (2.87) that will survive the $\Lambda\to\infty$ limit are the second order expansion from the last term in (2.87). In the end we get

$$
\begin{aligned}
\mathcal{W} &\equiv \lim_{\Lambda\to\infty} \mathcal{W}_\Lambda \\
&= \lim_{\Lambda\to\infty} \Lambda^4 \int \frac{d^4\tilde{k}_E}{(2\pi)^4}\, e^{-\tilde{k}_E^2} \cdot \mathrm{Tr}\left[\gamma^5 \exp\left(-\frac{2i\tilde{k}_E^\mu D_{E\mu}}{\Lambda} - \frac{D_E^\mu D_{E\mu}}{\Lambda^2} + \frac{ie}{2\Lambda^2}\gamma_E^\mu\gamma_E^\nu F_{\mu\nu}\right)\right] \\
&= \lim_{\Lambda\to\infty} \mathcal{W}_\Lambda \\
&= \lim_{\Lambda\to\infty} \Lambda^4 \int \frac{d^4\tilde{k}_E}{(2\pi)^4}\, e^{-\tilde{k}_E^2} \cdot \mathrm{Tr}\left[\gamma^5 \left\{1 + \left(\frac{\sim}{\Lambda} + \frac{\sim}{\Lambda^2} + \frac{\sim}{\Lambda^2}\right)\right.\right. \\
&\qquad\qquad \left.\left. + \frac{1}{2}\left(\frac{\sim}{\Lambda^2} + \frac{\sim}{\Lambda^3} + \frac{\sim}{\Lambda^4} - \frac{e^2}{4\Lambda^4}\gamma_E^\mu\gamma_E^\nu\gamma_E^\rho\gamma_E^\sigma F_{\mu\nu}F_{\rho\sigma}\right) + \cdots\right\}\right] \\
&= \lim_{\Lambda\to\infty} \Lambda^4 \int \frac{d^4\tilde{k}_E}{(2\pi)^4}\, e^{-\tilde{k}_E^2} \cdot \left[0 \cdot \frac{1}{\Lambda^{\neq 4}} + \sim \cdot \frac{1}{\Lambda^{\neq 4}} - \frac{e^2}{8\Lambda^4} \cdot (-4\epsilon^{\mu\nu\rho\sigma})F_{\mu\nu}F_{\rho\sigma}\right] \\
&= \lim_{\Lambda\to\infty} \Lambda^4 \int \frac{d^4\tilde{k}_E}{(2\pi)^4}\, e^{-\tilde{k}_E^2} \cdot \left(\frac{e^2}{2\Lambda^4}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right) \\
&= \frac{e^2}{32\pi^4}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} \int d^4\tilde{k}_E\, e^{-\tilde{k}_E^2} \\
&= \frac{e^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}.
\end{aligned}
\tag{2.90}
$$

Therefore from (2.53) and (2.75),

$$
\begin{aligned}
\mathcal{J} &= \exp\left[-i\int d^4x_E\, \epsilon(x) \sum_n \psi_n^\dagger(x)\gamma^5\psi_n(x)\right] \\
&= \exp\left[-i \lim_{\Lambda\to\infty} \int d^4x_E\, \epsilon(x)\mathcal{W}_\Lambda\right] \\
&= \exp\left[-i\int d^4x_E\, \epsilon(x)\mathcal{W}\right] \\
&= \exp\left[-i\int d^4x_E\, \epsilon(x_E)\frac{e^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}(x_E)F_{\rho\sigma}(x_E)\right].
\end{aligned}
\tag{2.91}
$$

In the last line we wrote $\epsilon(x_E)$ explicitly to emphasize that the transformation parameter is a function of Euclidean coordinates.

2.3 Non-conservation of the axial current

Using the same logic of (1.16),

$$
\begin{aligned}
Z[\bar{K}, K] = \int \mathcal{D}\bar{\psi}'\mathcal{D}\psi' \exp&\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E\,(\bar{K}\psi + \bar{\psi}K)\right] \\
&\times\left[1 - \int d^4x_E\, \epsilon(x)\left\{\partial_{E\mu}J^\mu_{(E)A} - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right],
\end{aligned}
\tag{2.92}
$$

where K and $\bar{K}$ are external sources for ψ and $\bar{\psi}$ respectively.

Using the same logic as in (1.16), we perform an infinitesimal axial rotation and expand to $\mathcal{O}(\epsilon)$. After rewriting the integrand in terms of the unprimed fields using the inverse transformation,

$$\begin{aligned} Z_E[\bar{K}, K] = \int \mathcal{D}\bar{\psi}' \, \mathcal{D}\psi' \, \exp\Big[-S_E[\bar{\psi}, \psi] + \int d^4x_E \, (\bar{K}\psi + \bar{\psi}K)\Big] \\ \times \left[1 - \int d^4x_E \, \epsilon(x_E) \left\{\partial_{E\mu} J^\mu_{(E)A} - i\big(\bar{K}\gamma^5_E\psi + \bar{\psi}\gamma^5_E K\big)\right\}\right]. \end{aligned} \tag{2.93}$$

Here K and $\bar{K}$ are external sources for ψ and $\bar{\psi}$.

Derivation

One can consider the Dirac action with external sources η and $\bar{\eta}$,

$$Z[\bar{K}, K] = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[iS[\bar{\psi}, \psi] + i\int d^4x \, (\bar{K}\psi + \bar{\psi}K)\right]. \tag{2.94}$$

Under the chiral rotation we get,

$$\begin{aligned} Z[\bar{K}, K] = \int \mathcal{D}\bar{\psi}'\mathcal{D}\psi' \, \exp\left[iS[\bar{\psi}, \psi] + i\int d^4x \, (\bar{K}\psi + \bar{\psi}K)\right] \\ \times \exp\left[-i\int d^4x \, \epsilon \left\{\partial_\mu J^\mu_A - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right]. \end{aligned} \tag{2.95}$$

Do the Wick rotation

$$x^0 = -ix^0_E, \quad x^i = x^i_E, \quad \partial_0 = i\partial_{E0}, \quad \partial_i = \partial_{Ei}, \quad \gamma^0 = -i\gamma^0_E, \quad \gamma^i = \gamma^i_E,$$

then

$$\begin{aligned} Z_E[\bar{K}, K] = \int \mathcal{D}\bar{\psi}'\mathcal{D}\psi' \, \exp\left[-S_E[\bar{\psi}, \psi] + \int d^4x_E \, (\bar{K}\psi + \bar{\psi}K)\right] \\ \times \exp\left[-\int d^4x_E \, \epsilon \left\{\partial_{E\mu} J^\mu_{(E)A} - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right], \end{aligned} \tag{2.96}$$

where we identified the Euclidean action as

$$S \quad \to \quad iS_E = \int d^4x_E \, \bar{\psi} \not{D}_E \psi, \tag{2.97}$$

with

$$\psi(x_E) \to i\psi(x_E) \quad \text{and} \quad \bar{\psi}(x_E) \to \bar{\psi}(x_E). \tag{2.98}$$

We use

$$\gamma^\mu_E D_{E\mu} = \gamma^0_E D_{E0} + \gamma^i_E D_{Ei} = \gamma^0 D_0 + \gamma^i D_i = \gamma^\mu D_\mu, \tag{2.99}$$

which follows easily from the Wick rotation rule.

From the (2.54) and (2.91) we know that the modified fermion measure reads

$$\begin{aligned}\int \mathcal{D}\bar{\psi}'\mathcal{D}\psi' &= \mathcal{J}^2 \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \\ &= \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[-i\int d^4x_E\, \epsilon(x_E)\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right], \end{aligned} \tag{2.100}$$

where we use the fact that $\mathcal{J}$ is constant with respect to the integral variables. By substituting this results into (2.93), expanding the exponential, and setting $K = 0$ and $\bar{K} = 0$ gives

$$\langle \partial_\mu J^\mu_A \rangle = \frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}. \tag{2.101}$$

We see that the axial current is no longer conserved, which means that the axial symmetry is anomalous. This is the chiral anomaly or ABJ anomaly which is named after Adler, Bell, and Jackiw.

Derivation

$$\begin{aligned} Z[\bar{K},K] = &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \mathcal{J}^{-2}\cdot \exp\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E\,(\bar{K}\psi + \bar{\psi}K)\right] \\ &\qquad \times \left[1 - \int d^4x_E\,\epsilon(x)\left\{\partial_{E\mu}J^\mu_{(E)A} - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right], \\ = &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \exp\left[i\int d^4x_E\,\epsilon(x_E)\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right] \\ &\qquad \times \exp\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E\,(\bar{K}\psi + \bar{\psi}K)\right] \\ &\qquad \times \left[1 - \int d^4x_E\,\epsilon(x)\left\{\partial_{E\mu}J^\mu_{(E)A} - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right] \\ = &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \left[1 + i\int d^4x_E\,\epsilon(x_E)\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} + \mathcal{O}(\epsilon^2)\right] \\ &\qquad \times \exp\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E\,(\bar{K}\psi + \bar{\psi}K)\right] \\ &\qquad \times \left[1 - \int d^4x_E\,\epsilon(x)\left\{\partial_{E\mu}J^\mu_{(E)A} - i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}\right] \\ = &\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \exp\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E\,(\bar{K}\psi + \bar{\psi}K)\right] \\ &\times\Bigg[1 + \int d^4x_E\,\epsilon(x_E) \\ &\qquad \times\left\{\frac{ie^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} - \partial_{E\mu}J^\mu_{(E)A} + i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\} + \mathcal{O}(\epsilon^2)\Bigg], \end{aligned}$$

where $J^\mu_{(E)A} = i\bar{\psi}\gamma^\mu_E\gamma^5_E\psi$. Then we require

$$0 = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[-S_E[\psi,\bar{\psi}] + \int d^4x_E \, (\bar{K}\psi + \bar{\psi}K)\right] \times \int d^4x_E \, \epsilon(x_E) \left\{\frac{ie^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} - \partial_{E\mu}J^\mu_{(E)A} + i(\bar{K}\gamma^5\psi + \bar{\psi}\gamma^5 K)\right\}. \tag{2.102}$$

Take $K = \bar{K} = 0$,

$$\langle \partial_{E\mu}J^\mu_{(E)A}\rangle = i\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}. \tag{2.103}$$

Until now, we omit the Euclidean index E to the $A_\mu(x)$ but the above field strength tensor is defined on the Euclidean space. So we need to inverse Wick rotation to Minkowski space,

$$A_0 = iA_{E0} \quad \to \quad F_{E0i} = \partial_{E0}A_{Ei} - \partial_{Ei}A_{E0} = -i(\partial_0 A_i - \partial_i A_0) = -iF_{0i} \tag{2.104}$$

Then

$$\epsilon^{\mu\nu\rho\sigma}F_{E\mu\nu}F_{E\rho\sigma} = 4\epsilon^{0ijk}F_{E0i}F_{Ejk} = -4i\epsilon^{0ijk}F_{0i}F_{jk} = -i\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}, \tag{2.105}$$

where $\epsilon^{\mu\nu\rho\sigma}_E = \epsilon^{\mu\nu\rho\sigma}_M$. Therefore

$$\langle \partial_{E\mu}J^\mu_{(E)A}\rangle = i\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{E\mu\nu}F_{E\rho\sigma} \quad \to \quad \langle \partial_\mu J^\mu_A\rangle = \frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}. \tag{2.106}$$

We use

$$\partial_{E\mu}J^\mu_{(E)A} = \partial_\mu J^\mu_A, \tag{2.107}$$

which drive easily from the Wick rotation rule.

2.4 The Anomaly with Non-abelian gauge theories

In the previous example the gauge symmetry under which the fermions transform was an abelian gauge symmetry. We will be eventually interested in QCD where the fermions transform under the fundamental representation of $SU(3)_C$ which is a non-abelian symmetry. Fortunately, we only need to modify (2.84),

$$\not{D}^2_E = D^2_E - \frac{ig}{2}\gamma^\mu_E\gamma^\nu_E F^a_{\mu\nu}T^a, \tag{2.108}$$

where $D_{E\mu} = \partial_{E\mu} - igA^a_{E\mu}T^a$ so that $[D_{E\mu}, D_{E\nu}] = -igF^a_{E\mu\nu}T^a$. The g is the gauge coupling constant $\{T^a\}$ are the generators in the fundamental representation, and $F^a_{\mu\nu}$ is the gauge field strength tensor. Recall that the main contribution to the anomaly comes from squaring the second term and taking the trace together with γ^5. Now we also need to take a trace over the group generators. So all the results in the abelian case can be applied to the non-abelian case by replacing

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \quad \to \epsilon^{\mu\nu\rho\sigma} F^a_{\mu\nu} F^b_{\rho\sigma} \mathrm{Tr}\left(T^a T^b\right) = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F^a_{\mu\nu} F^a_{\rho\sigma}, \tag{2.109}$$

where we have used the convention that the generators in the fundamental representation are normalized by $\mathrm{Tr}\left(T^a T^b\right) = \delta_{ab}/2$. Note that for abelian case, $\mathrm{Tr}\left(T^a T^b\right) = \mathrm{Tr}(1) = 1$.

We can directly apply this to the massless QCD Lagrangian. By making an axial rotation given in (2.5) to a single quark the path integral measure changes by

$$\begin{aligned}\int \mathcal{D}\bar{q}\mathcal{D}q \to & \int \mathcal{D}\bar{q}\mathcal{D}q \, \exp\left[i\int d^4x_E \, \epsilon(x_E)\frac{g_s^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma} G^a_{\mu\nu} G^a_{\rho\sigma}\right] \\ = & \int \mathcal{D}\bar{q}\mathcal{D}q \, \exp\left[i\int d^4x_E \, \epsilon(x_E)\frac{g_s^2}{16\pi^2} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right],\end{aligned} \tag{2.110}$$

where

$$\tilde{G}^{a\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} G^a_{\rho\sigma}. \tag{2.111}$$

Note that the first line is 32 not 16, compare to the (2.100). This modification corresponds to a change in the QCD Lagrangian given by

$$\mathcal{L} \quad \to \quad \mathcal{L}' = \mathcal{L} + \epsilon\frac{g_s^2}{16\pi^2} G^a_{\mu\nu}\tilde{G}^{a\mu\nu}. \tag{2.112}$$

Note that this modification is purely a quantum effect and cannot be seen from a classical analysis.

Derivation

The chiral anomaly effect in Euclidean space is

$$\begin{aligned}\int \mathcal{D}\bar{q}\mathcal{D}q \, e^{-\Delta S_E} \equiv & \int \mathcal{D}\bar{q}\mathcal{D}q \, \exp\left[i\int d^4x_E \, \epsilon(x_E)\frac{g_s^2}{16\pi^2} G^a_{E\mu\nu}\tilde{G}^a_{E\mu\nu}\right] \\ = & \int \mathcal{D}\bar{q}\mathcal{D}q \, \exp\left[i\int (id^4x) \, \epsilon(x)\frac{g_s^2}{16\pi^2}\left(-iG^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right)\right] \\ = & \int \mathcal{D}\bar{q}\mathcal{D}q \, \exp\left[i\int d^4x \, \epsilon\frac{g_s^2}{16\pi^2} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right] \equiv \int \mathcal{D}\bar{q}\mathcal{D}q \, e^{i\Delta S}.\end{aligned} \tag{2.113}$$

Therefore,

$$e^{i\Delta S} = \exp\left[i\int d^4x \, \Delta\mathcal{L}\right] = \exp\left[i\int d^4x \, \epsilon\frac{g_s^2}{16\pi^2} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right]. \tag{2.114}$$

Also it turns out that $G^a_{\mu\nu}\tilde{G}^{a\mu\nu}$ is a total derivative. It can be written as

$$G^a_{\mu\nu}\tilde{G}^{a\mu\nu} = \partial_\mu\left[\epsilon^{\mu\alpha\beta\gamma}\left(A^a_\alpha G^a_{\beta\gamma} - \frac{g_s}{3} f^{abc} A^a_\alpha A^b_\beta A^c_\gamma\right)\right] \equiv \partial_\mu \mathcal{K}^\mu, \tag{2.115}$$

where $\mathcal{K}^\mu$ is the Chern-Simons current. So, at first sight it seems that it should not contribute to the local equations of motion, and sholud be irrelevant. This is not true though. The QCD has a complicated vacuum structure, and this term will play a vital role.

3 Instantons and θ vacua

In the last subsection, we have seen that a chiral rotation $q \to e^{i\epsilon\gamma_5} q$ of a single quark flavor modifies the QCD Lagrangian as shown in (2.112). We also mentioned that the induced term can be written as a total derivative of the so-called Chern–Simons current $\mathcal{K}^\mu$, cf. (2.115). One might therefore think that this term is irrelevant because it does not affect the classical equations of motion. However, this statement is true only at the perturbative level. This term crucially affects the vacuum structure of QCD, albeit non-perturbatively. Furthermore, this effect does not depend on whether the gauge coupling constant g_s is small or not. Non-perturbative effects can be present even if the theory is weakly coupled.

In quantum mechanics, we have already studied tunneling through a potential barrier. In the WKB approximation, the transmission amplitude behaves as

$$|\tau(E)| \sim \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x_2} dx\,\sqrt{2m\,(V(x)-E)}\right][1+\mathcal{O}(\hbar)]\,, \tag{3.1}$$

where $E < V_{\max}$ and x_1, x_2 are the classical turning points defined by $V(x_{1,2}) = E$. No matter how small the coupling in the potential V is, this effect can not be seen in any order of perturbation theory. Hence, it is strictly a non-perturbative phenomenon. There are also phenomena in quantum field theory that are analogs of barrier penetration in quantum mechanics.

3.1 Instantons in quantum mechanics

We consider a particle of unit mass in one dimension with the Hamiltonian,

$$H = \frac{p^2}{2} + V(x). \tag{3.2}$$

Then the transition amplitude is defined as

$$\left\langle x_f, \frac{T}{2}\middle| x_i, -\frac{T}{2}\right\rangle = \langle x_f|e^{-\frac{i}{\hbar}HT}|x_i\rangle = N\int \mathcal{D}x\, e^{\frac{i}{\hbar}S[x]}, \tag{3.3}$$

where

$$S = \int_{-T/2}^{T/2} dt\,\left[\frac{1}{2}\left(\frac{dx}{dt}\right)^2 - V(x)\right]. \tag{3.4}$$

By performing a Wick rotation, we can write the Euclidean time,

$$\tau = it \quad \to \quad \langle x_f|e^{-HT/\hbar}|x_i\rangle = N\int \mathcal{D}x\, e^{-S_E/\hbar}, \tag{3.5}$$

where

$$S_E = \int_{-T/2}^{T/2} d\tau\,\left[\frac{1}{2}\left(\frac{dx}{d\tau}\right)^2 + V\right]. \tag{3.6}$$

The Euclidean functional integral represents a sum over all paths $x(\tau)$ that obey the boundary conditions

$$x(-T/2) = x_i, \qquad x(T/2) = x_f. \tag{3.7}$$

We see that the Euclidean action has the same structure as the Minkowski action S_M in real time, but it leads to classical motion in the inverted potential $-V(x)$.

In the semi-classical limit where $\hbar \to 0$, the Euclidean path integral is dominated by the stationary points of the Euclidean action. The corresponding classical equation of motion is

$$\left.\frac{\delta S_E[x]}{\delta x}\right|_{x=x_{cl}} = 0 \quad \to \quad \frac{d^2 x_{cl}(\tau)}{d\tau^2} - \left.\frac{dV}{dx}\right|_{x=x_{cl}(\tau)} = 0. \tag{3.8}$$

Derivation

Consider a expansion of non-stationary point x_0,

$$S_E(x_0+\epsilon) \approx S_E(x_0) + S_E'(x_0)\epsilon, \tag{3.9}$$

where

$$\epsilon \equiv x - x_0 \in [x_0 - \Delta, x_0 + \Delta]. \tag{3.10}$$

Then

$$\begin{aligned} J[x_0] \equiv \int \mathcal{D}x\, e^{-S_E/\hbar} &\approx \int_{-\Delta}^{\Delta} d\epsilon \, \exp\left[-\frac{1}{\hbar} S_E(x_0+\epsilon)\right] \\ &= e^{-S_E(x_0)/\hbar} \int_{-\Delta}^{\Delta} d\epsilon \, e^{-\frac{S_E'(x_0)}{\hbar}\epsilon} \\ &= -e^{-S_E(x_0)/\hbar} \cdot \frac{\hbar}{S_E'(x_0)} \left[e^{-\frac{S_E'(x_0)}{\hbar}\epsilon} \right]_{-\Delta}^{\Delta} \\ &= e^{-S_E(x_0)/\hbar} \cdot \frac{\hbar}{S_E'(x_0)} \left[e^{S_E'(x_0)\Delta/\hbar} - e^{-S_E'(x_0)\Delta/\hbar} \right] \\ &= e^{-S_E(x_0)/\hbar} \cdot \frac{2\hbar}{S_E'(x_0)} \cdot \sinh\left[S_E'(x_0)\Delta/\hbar\right] \\ &= \frac{\hbar}{S_E'(x_0)} \cdot \left[e^{-S_E(x_0-\Delta)/\hbar} - e^{-S_E(x_0+\Delta)/\hbar} \right]. \end{aligned} \tag{3.11}$$

The expansion of minimum point x_*,

$$S_E(x_*+\epsilon) \approx S_E(x_*) + \frac{1}{2!} S_E''(x_*)\epsilon^2, \tag{3.12}$$

where

$$\epsilon \equiv x - x_* \in [x_* - \Delta, x_* + \Delta]. \tag{3.13}$$

Note that that Δ is not same as above Δ in (3.10). Then

$$\begin{aligned} J[x_*] \equiv \int dx\, e^{-S_E/\hbar} &\approx \int_{-\Delta}^{\Delta} d\epsilon \, \exp\left[-\frac{1}{\hbar} S_E(x_* + \epsilon)\right] \\ &= e^{-S_E(x_*)/\hbar} \int_{-\Delta}^{\Delta} d\epsilon \, e^{-\frac{S''_E(x_*)}{2\hbar}\epsilon^2} \\ &\approx e^{-S_E(x_*)/\hbar} \int_{-\infty}^{\infty} d\epsilon \, e^{-\frac{S''_E(x_*)}{2\hbar}\epsilon^2} \\ &= e^{-S_E(x_*)/\hbar} \cdot \sqrt{\frac{2\pi\hbar}{S''_E(x_*)}}. \end{aligned} \tag{3.14}$$

Without lose generality, one can assume $S'_E(x_0) > 0$, then calculate

$$\begin{aligned} J[x_0] &= \frac{\hbar}{S'_E(x_0)} \cdot \left[e^{-S_E(x_0-\Delta)/\hbar} - e^{-S_E(x_0+\Delta)/\hbar}\right] \qquad (3.15) \\ &= \frac{\hbar}{S'_E(x_0)} \cdot e^{-S_E(x_*)/\hbar} \cdot \left[e^{-\frac{S_E(x_0-\Delta)-S_E(x_*)}{\hbar}} - e^{-\frac{S_E(x_0+\Delta)-S_E(x_*)}{\hbar}}\right] \\ &\approx \frac{\hbar}{S'_E(x_0)} \cdot e^{-S_E(x_*)/\hbar} \cdot e^{-\frac{S_E(x_0-\Delta)-S_E(x_*)}{\hbar}} \equiv \frac{\hbar}{S'_E(x_0)} \cdot e^{-S_E(x_*)/\hbar} \cdot e^{-\delta/\hbar}. \end{aligned}$$

Note that one can always find the $\delta > 0$, for sufficiently small Δ, which does not contain the global minimum x_*. Therefore,

$$\frac{J[x_0]}{J[x_*]} \approx \frac{\frac{\hbar}{S'_E(x_0)} e^{-S_E(x_*)/\hbar} e^{-\delta/\hbar}}{e^{-S_E(x_*)/\hbar}\sqrt{\frac{2\pi\hbar}{S''_E(x_*)}}} = \frac{\sqrt{S''_E(x_*)}}{S'_E(x_0)} \cdot \sqrt{\frac{\hbar}{2\pi}} \cdot e^{-\delta/\hbar} \xrightarrow{\hbar\to 0} 0. \tag{3.16}$$

3.1.1 Double-well potential

Consider a potential that has two minima located at $x = \pm a$, corresponding to the classical ground states of the system. In Euclidean time, the equation of motion can be interpreted as classical motion in the inverted potential $-V(x)$, so the minima of V become maxima of $-V$. Suppose the particle starts at one of these maxima. Then there are two qualitatively different solutions

- The particle can remain at the maximum where it started.
- It can move from one maximum to the other.

In Minkowski time, the first solution corresponds to a particle initially sitting in one of the minima and remaining there, as expected classically. The second corresponds to quantum tunneling between the two minima. Thus, tunneling is represented by a classical solution of the *Euclidean* equations of motion (an instanton), even though it has no counterpart as a real-time classical trajectory.

Consider a particle that starts at the left maximum of the Euclidean (inverted) potential and ends at the right maximum,

$$x(-\infty) = -a \quad \to \quad x(+\infty) = +a. \tag{3.17}$$

The conserved Euclidean energy is

$$E_E = \frac{1}{2}\left(\frac{dx}{d\tau}\right)^2 - V(x). \tag{3.18}$$

For an instanton satisfying $x(\tau \to \pm\infty) = \pm a$ with $\frac{dx}{d\tau} \to 0$ and with the convention $V(\pm a) = 0$, we have $E_E = 0$. With $E_E = 0$, we obtain

$$\frac{dx}{d\tau} = \pm\sqrt{2V(x)}, \tag{3.19}$$

and hence

$$\tau - \tau_0 = \pm \int_{x(\tau_0)}^{x(\tau)} \frac{dx'}{\sqrt{2V(x')}}. \tag{3.20}$$

For the quartic double-well potential

$$V(x) = \frac{\lambda}{4}(x^2 - a^2)^2, \tag{3.21}$$

the condition $E_E = 0$ gives, for the trajectory from $-a$ to $+a$,

$$\frac{dx}{d\tau} = \sqrt{2V(x)} = \sqrt{\frac{\lambda}{2}}\,(a^2 - x^2), \tag{3.22}$$

where we used $|x^2 - a^2| = a^2 - x^2$ along $x \in [-a, a]$. Separating variables,

$$\tau - \tau_0 = \sqrt{\frac{2}{\lambda}} \int_0^{x(\tau)} \frac{dx'}{a^2 - x'^2} = \frac{1}{a}\sqrt{\frac{2}{\lambda}}\ \mathrm{arctanh}\left(\frac{x(\tau)}{a}\right). \tag{3.23}$$

Therefore,

$$x(\tau) = a\ \tanh\left[a\sqrt{\frac{\lambda}{2}}\,(\tau - \tau_0)\right]. \tag{3.24}$$

This solution is known as an instanton with the center at τ_0. Similarly, we can construct solutions that start at the right maximum to the left maximum,

$$x(-\infty) = +a \quad \to \quad x(+\infty) = -a. \tag{3.25}$$

These are called anti-instantons.

3.2 Euclidean Yang-Mills action

The action for a generic Yang-Mills theory in Minkowski space is

$$S_{YM} = -\frac{1}{2}\int d^4x\, \mathrm{Tr}(G_{\mu\nu}G^{\mu\nu}) = -\frac{1}{4}\int d^4x\, G^a_{\mu\nu}G^{a\mu\nu}. \tag{3.26}$$

Here we use the standard normalization $\mathrm{Tr}(T^aT^b) = \frac{1}{2}\delta^{ab}$. The $G_{\mu\nu}$ is the Lie-algebra-valued field strength defined by

$$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \equiv G^a_{\mu\nu}T^a, \tag{3.27}$$

with A_μ being the Lie-algebra valued gauge potential,

$$A_\mu = A^a_\mu T^a, \tag{3.28}$$

where $\{T^a\}$ are the generators of $\mathcal{G}$. The components of $G_{\mu\nu}$ are

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu, \tag{3.29}$$

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

$$[T^a, T^b] = if^{abc}T^c. \tag{3.30}$$

For QCD, the gauge group is $\mathcal{G} = SU(3)_c$, the gauge potentials $\{A^a_\mu\}^8_{a=1}$ 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

$$A_\mu \quad \rightarrow \quad UA_\mu U^{-1} + \frac{i}{g}U\partial_\mu U^{-1}, \tag{3.31}$$

while the transformation law for the field strength is

$$G_{\mu\nu} \quad \rightarrow \quad UG_{\mu\nu}U^{-1}. \tag{3.32}$$

One can perform the Wick rotation as,

$$x^0_E \equiv ix^0, \qquad \partial_{E0} = -i\partial_0, \qquad A_{E0} \equiv -iA_0. \tag{3.33}$$

From (3.29) we also get

$$\begin{aligned} G^a_{Ei0} &= \partial_{Ei}A^a_{E0} - \partial_{E0}A^a_{Ei} + gf^{abc}A^b_{Ei}A^c_{E0} = -iG^a_{i0} \\ G^{ai0}_E &= \partial^i_E A^{a0}_E - \partial^0_E A^{ai}_E + gf^{abc}A^{bi}_E A^{c0}_E = iG^{ai0}, \end{aligned} \tag{3.34}$$

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

$$S_{YM} = \frac{i}{4}\int d^4x_E\, G^a_{E\mu\nu}G^{a\mu\nu}_E \equiv iS^EYM. \tag{3.35}$$

Note that the metric in Euclidean spacetime is $-\delta_{\mu\nu}$, so raising and lowering indices only introduces an overall minus sign.

3.3 Topology of the Yang-Mills vacuum

We want to explore the vacuum structure of Yang–Mills theory within the semiclassical approximation. The first step is to identify field configurations (classical solutions) that minimize the Euclidean action. From the Euclidean Yang–Mills action (3.35), one finds that it is positive definite, and hence it is minimized when the field strength vanishes,

$$G^a_{E\mu\nu}\Big|_{\text{VAC}} = 0. \tag{3.36}$$

However, this does not mean that the gauge potential must vanish. Using the gauge transformation laws (3.31) and (3.32), one sees that there exist gauge potentials with vanishing field strength,

$$A^{\text{vac}}_{E\mu}(x_E) = \frac{i}{g}\,U(x_E)\,\partial_{E\mu}U^{-1}(x_E), \tag{3.37}$$

for which $G_{E\mu\nu} = 0$. In particular, choosing $U = \mathbb{I}$ gives $A_{E\mu} = 0$. Gauge fields of this form are called *pure gauges* and represent classical vacuum configurations. Distinct vacua are classified by the homotopy classes of the gauge transformations U at spatial infinity (winding number).

Each gauge transformation $U(x_E)$ is, in general, a map

$$U(x_E):\ \mathbb{R}^4(\text{Euclidean spacetime}) \quad \rightarrow \quad \mathcal{G}(\text{gauge group}). \tag{3.38}$$

We can employ the gauge freedom in the Yang-Mills theory and choose the temporal gauge, where

$$A_{E0}(x_E) = 0, \qquad (\text{temporal gauge}). \tag{3.39}$$

In this gauge, vacuum configurations may be taken to be time-independent,

$$\mathbf{A}_E\Big|_{\text{VAC}} = \frac{i}{g}U(\mathbf{x}_E)\nabla U^{-1}(\mathbf{x}_E). \tag{3.40}$$

Furthermore, we restrict attention to tunneling configurations for which the Euclidean Yang–Mills action is finite, $S_E < \infty$. This restriction is natural for two reasons:

- In the semiclassical approximation, the path integral is dominated by configurations with finite S_E, since contributions are weighted by e^{-S_E} (up to conventions).

- Finite S_E forces the field strength to vanish at Euclidean infinity, so the fields approach vacuum configurations asymptotically; this is precisely the setting in which tunneling between vacua is well-defined.

Derivation

Consider

$$\mathcal{L} = -\frac{1}{2g^2}\mathrm{Tr}(G_{\mu\nu}G^{\mu\nu}) = -\frac{1}{2g^2}\mathrm{Tr}\left[-2(G_{0i})^2 + G_{ij}G_{ij}\right]$$
$$= \frac{1}{g^2}\mathrm{Tr}(E_iE_i - B_iB_i), \tag{3.41}$$

where

$$E_i = G_{0i} = \dot{A}_i - D_iA_0, \quad B_i = -\frac{1}{2}\epsilon_{ijk}G_{jk}, \quad D_i = \partial_i - ig[A_i, \cdot], \tag{3.42}$$

with

$$\frac{\delta S}{\delta A_0} = 0 \quad \rightarrow \quad (D_iE_i)^a = J^{0a} = 0 \qquad \text{(Gauss constraints)}. \tag{3.43}$$

The canonical momentum is

$$\Pi_i^a = \frac{\partial\mathcal{L}}{\partial\dot{A}_i^a} = \frac{1}{g^2}E_i^a, \qquad \Pi_0^a = 0. \tag{3.44}$$

Then the Hamiltonian is

$$\mathcal{H} = \Pi_i^a\dot{A}_i^a - \mathcal{L} = \frac{2}{g^2}\mathrm{Tr}\left[E_i(E_i + D_iA_0)\right] - \frac{1}{g^2}\mathrm{Tr}(E_i^2 - B_i^2)$$
$$= \frac{1}{g^2}\mathrm{Tr}(E_i^2 + B_i^2) + \frac{2}{g^2}\mathrm{Tr}(E_iD_iA_0). \tag{3.45}$$

Using integration by parts,

$$\int d^3x\,\mathrm{Tr}(E_iD_iA_0) = -\int d^3x\,\mathrm{Tr}[(D_iE_i)A_0] + \int d^3x\,\mathrm{Tr}[D_i(E_iA_0)]. \tag{3.46}$$

From the constraints (3.43) and temporal gauge (3.39), we get

$$\mathcal{H} = \frac{1}{g^2}\mathrm{Tr}(E_iE_i + B_iB_i) \tag{3.47}$$

The Euclidean Lagrangian is

$$\mathcal{L}_E = \frac{1}{2g^2}\mathrm{Tr}(G_{E\mu\nu}G_E^{\mu\nu}) = \frac{1}{2g^2}\mathrm{Tr}\left[2(G_{E0i})^2 + G_{Eij}G_{Eij}\right]$$
$$= \frac{1}{g^2}\mathrm{Tr}(E_{Ei}^2 + B_{Ei}^2), \tag{3.48}$$

where

$$E_{Ei} = G_{E0i}, \qquad B_{Ei} = -\frac{1}{2}\epsilon_{ijk}G_{Ejk}. \tag{3.49}$$

Note that

$$E_i = iE_{Ei}, \qquad B_i = B_{Ei}. \tag{3.50}$$

It turns out this condition requires us to choose the gauge transformations to that approach to the identity at spatial infinity,

$$U(\mathbf{x}_E) \quad \rightarrow \quad \mathbb{I} \qquad \text{as} \qquad |\mathbf{x}_E| \quad \rightarrow \infty. \tag{3.51}$$

This means that we can treat the spatial infinity as a same point. Then one can do the one-point compactification of $\mathbb{R}^3$ to S^3. Finally, the gauge transformation parameter $U(x_E)$ as a map

$$U(\mathbf{x}_E): \; S^3 \quad \rightarrow \quad \mathcal{G}. \tag{3.52}$$

3.3.1 Winding number

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

In order to gain some intuition, let us start with a much simpler case which is the classification of the maps from S^1 to S^1. The map is given by

$$f: S^1 \rightarrow S^1, \qquad f(e^{i\theta}) = e^{i\phi(\theta)}, \tag{3.53}$$

with a continuous lift $\phi: \mathbb{R} \rightarrow \mathbb{R}$ satisfying $\phi(\theta + 2\pi) = \phi(\theta) + 2\pi n$ for some integer n, and

$$S^1 = \{e^{i\theta} \,|\, \theta \in [0, 2\pi)\}. \tag{3.54}$$

We have the following options:

- Mapping S^1 to a single point. This is the trivial map, which sends every point to a constant element of S^1,

  $$f_0(e^{i\theta}) = 1, \qquad \forall\theta, \tag{3.55}$$

  with zero winding number, $n = 0$.

- Mapping S^1 into a finite subset of S^1,

  $$f_0(e^{i\theta}) = \{e^{i\phi_1}, e^{i\phi_2}, \dots e^{i\phi_k}\}, \qquad \forall\theta. \tag{3.56}$$

  However this map can transform smoothly to trivial map, thus it is in the same winding number $n = 0$.

- Mapping S^1 to itself once with the same orientation,

  $$f_{+1}(e^{i\theta}) = e^{i\theta} \quad \leftrightarrow \quad \phi(\theta) = \theta. \tag{3.57}$$

  This map cannot homotopic(continuously transform) to trivial map. Then the winding number is

  $$n = \frac{\Delta\phi}{2\pi} = \frac{\phi(2\pi) - \phi(0)}{2\pi} = 1. \tag{3.58}$$

- Mapping S^1 into the full S^1 identically but opposite direction of the increasing angle,

$$f_{-1}(e^{i\theta}) = e^{-i\theta} \quad \leftrightarrow \quad \phi(\theta) = -\theta. \tag{3.59}$$

  The winding number is

$$n = \frac{\Delta\phi}{2\pi} = \frac{-2\pi - 0}{2\pi} = -1. \tag{3.60}$$

- Mapping S^1 into the S^1 at the target range more than one times,

$$f_n(e^{i\theta}) = e^{in\theta} \quad \leftrightarrow \quad \phi(\theta) = n\theta. \tag{3.61}$$

  Then the winding number is

$$n = \frac{\Delta\phi}{2\pi} = \frac{2n\pi - 0}{2\pi}. \tag{3.62}$$

Equivalently, the winding number can be written as

$$n = \frac{1}{2\pi}\int_0^{2\pi} d\theta\, \frac{d\phi(\theta)}{d\theta}. \tag{3.63}$$

This example shows that we can classify the maps from S^1 to S^1 by a countable infinite number of equivalence classes labeled by an integer winding number.

If we define a product of maps by pointwise multiplication in the target S^1,

$$(f_n \cdot f_m)(e^{i\theta}) \equiv f_n(e^{i\theta})\, f_m(e^{i\theta}), \tag{3.64}$$

then

$$(f_n \cdot f_m)(e^{i\theta}) = e^{in\theta} e^{im\theta} = e^{i(n+m)\theta} = f_{n+m}(e^{i\theta}), \tag{3.65}$$

so the winding numbers add. More generally, the set of homotopy classes of maps $S^m \to X$ forms the m-th homotopy group $\pi_m(X)$. In particular, for $X = S^n$ we write $\pi_m(S^n)$, where m and n are the dimensions of the domain and target spheres, respectively. In other words, the homotopy group of the maps from S^m to S^n. For example, $m = 1$ case is corresponding to loop concatenation.

For the maps from S^1 to S^1 we have shown that

$$\pi_1(S^1) = \mathbb{Z}. \tag{3.66}$$

This extends to higher dimensions, for every $d \geq 1$ one has

$$\pi_d(S^d) = \mathbb{Z}. \tag{3.67}$$

In particular, our interested case is

$$\pi_3(S^3) = \mathbb{Z}. \tag{3.68}$$

With this results we can classify the zero-energy states, meaning the pure gauges denoted by the winding number of U by defining

$$\mathbf{A}_{E(n)}\Big|_{\text{VAC}} = \frac{i}{g} U_n(\mathbf{x}_E) \nabla U^{-1}_{(n)}(\mathbf{x}_E), \tag{3.69}$$

where $U_{(n)}$ has the winding number n. This analysis shows that Yang–Mills theory (and in particular QCD) admits a countably infinite family of classically degenerate vacuum configurations labeled by an integer n. The true quantum vacuum is not any single $|n\rangle$ state, but rather an appropriate superposition (the θ vacuum), as we shall see.

3.3.2 Yang-Mills Instantons

Consider that we start at the pure gauge configuration $A^{(n)}_{E\mu}|_{\text{VAC}}$ at $\tau = -\infty$ with the winding number n, and will end at the pure gauge $A^{(m)}_{E\mu}|_{\text{VAC}}$ at $\tau = +\infty$ with the winding number $m = n + 1$. From the Euclidean Yang–Mills action (3.35), the equation of motion is

$$D_\mu G_{\mu\nu} = \partial_\mu G_{\mu\nu} - ig[A_\mu, G_{\mu\nu}] = 0. \tag{3.70}$$

The solution to this equation $A^{(I)}_{E\mu}(x_E)$ subject to the boundary conditions mentioned above is called a Yang-Mills instanton. Explicitly, the solution is

$$A^{(I)}_{E\mu}(x_E) = \frac{2}{g} \frac{\eta^a_{\mu\nu}(x_{E\nu} - z_{E\nu})\tau^a}{(x_{E\mu} - z_{E\mu})^2 + \rho^2}, \tag{3.71}$$

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

$$\eta^a_{\mu\nu} = \delta_{a\mu}\delta_{0\nu} - \delta_{a\nu}\delta_{0\mu} + \epsilon_{a\mu\nu 0}, \qquad \epsilon_{0123} = +1. \tag{3.72}$$

Some properties can be read directly from the solution:

- Its spin (Lorentz index) is coupled with the color orientation via the 't Hooft symbol.

- Its contribution is genuinely non-perturbative, being weighted by e^{-S_I} with $S_I \sim 8\pi^2/g^2$, and thus it is invisible at any finite order in perturbation theory.

- The configuration is localized around z_E with characteristic size ρ. For $r \equiv |x_E - z_E| \gg \rho$ (in regular gauge),

$$A^a_{E\mu}(x_E) \sim \frac{1}{r}. \tag{3.73}$$

With the same logic, one can get a Yang-Mills anti-instanton by choosing $m = n - 1$. The solution is identical to the instanton solution except one uses the anti-self-dual 't Hooft symbol,

$$\bar{\eta}^a_{\mu\nu} = -\delta_{a\mu}\delta_{0\nu} + \delta_{a\nu}\delta_{0\mu} + \epsilon_{a\mu\nu 0}. \tag{3.74}$$

Duality

Consider a rank-two antisymmetric tensor $X_{\mu\nu}$. Its Hodge dual is

$$\tilde{X}_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}X_{\rho\sigma}. \tag{3.75}$$

The definition of (anti-) self-dual is

$$X_{\mu\nu} = \pm\tilde{X}_{\mu\nu}. \tag{3.76}$$

From the 't Hooft symbol, one can find

$$\eta^a_{\mu\nu} = +\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\eta^a_{\rho\sigma}, \qquad \bar{\eta}^a_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\bar{\eta}^a_{\rho\sigma}. \tag{3.77}$$

3.3.3 Instanton action

From the instanton solution given in (3.70), we can easily get the corresponding instanton field strength as

$$G^{(I)}_{E\mu\nu}(x_E) = -\frac{4\rho^2}{g}\frac{\eta^a_{\mu\nu}\tau^a}{[(x_{E\mu} - z_{E\mu})^2 + \rho^2]^2}. \tag{3.78}$$

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

$$S_I = \frac{1}{2}\int d^4x \,\mathrm{Tr}\left[G^{(I)}_{E\mu\nu}G^{(I)}_{E\mu\nu}\right] = \frac{8\pi^2}{g^2}. \tag{3.79}$$

The anti-instanton has the same action. More generally, a configuration that interpolates between vacua with winding numbers n and m, the minimal action in that sector is

$$S = \frac{8\pi^2}{g^2}\,|m - n|. \tag{3.80}$$

We see that the instanton action is finite and strictly positive. Moreover, S_I does not depend on:

- the position z_E of the instanton, by translational invariance;
- the global $SU(2)$ color orientation (embedding) of the solution, by gauge symmetry;
- the size ρ, due to the classical scale invariance of the Yang–Mills Lagrangian.

Note that this scale invariance is a classical symmetry; in the quantum theory it is broken by the trace anomaly (equivalently by the running of the gauge coupling), which makes the instanton size distribution ρ-dependent.

3.4 Quark zero modes and index theorems

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

$$\partial_\mu J_A^\mu = \frac{g^2}{8\pi^2} \mathrm{Tr}\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right]. \tag{3.81}$$

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

$$G^a_{\mu\nu}\tilde{G}^{a\mu\nu} = \partial_\mu\left[\epsilon^{\mu\alpha\beta\gamma}\left(A^a_\alpha G^a_{\beta\gamma} - \frac{g}{3} f^{abc} A^a_\alpha A^b_\beta A^c_\gamma\right)\right] \equiv \partial_\mu \mathcal{K}^\mu. \tag{3.82}$$

This implies that the integral of (3.81) over the Euclidean spacetime should yield a quantity that depends only on the boundary, meaning topological. Now we demonstrate this explicitly and, in the process, connect it to fermionic zero modes and an index theorem.

We define the following quantity in Euclidean space,

$$Q \equiv \frac{g^2}{16\pi^2}\int d^4x_E\, \mathrm{Tr}\left[G_{E\mu\nu}\tilde{G}_E^{\mu\nu}\right], \tag{3.83}$$

which is called the topological charge. From the anomaly section, we can see that this term can be expressed as (2.75)

$$Q = -\lim_{\Lambda\to\infty}\sum_m e^{-\lambda_m^2/\Lambda^2}\int d^4x_E\, \psi_m^\dagger(x_E)\gamma_E^5\psi_m(x_E), \tag{3.84}$$

where λ_m and ψ_m are the eigenvalues and the eigenfunctions of the Euclidean Dirac operator,

$$\not{D}_E\psi_m(x_E) = \lambda_m\psi_m(x_E). \tag{3.85}$$

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

$$\not{D}_E\gamma_E^5\psi_n(x_E) = -\gamma_E^5\not{D}_E\psi_n(x_E) = -\lambda_n\gamma_E^5\psi_n(x_E). \tag{3.86}$$

Derivation

From the definition of Euclidean gamma matrices (2.15), one can find its anticommutation. Without loss of generality consider $\mu = 0$,

$$\begin{aligned}\{\gamma_E^0,\gamma_E^5\} &= \gamma_E^0\gamma_E^5 + \gamma_E^5\gamma_E^0\\ &= \gamma_E^0\gamma_E^0\gamma_E^1\gamma_E^2\gamma_E^3 + \gamma_E^0\gamma_E^1\gamma_E^2\gamma_E^3\gamma_E^0\\ &= -\gamma_E^1\gamma_E^2\gamma_E^3 + \gamma_E^1\gamma_E^2\gamma_E^3 = 0.\end{aligned} \tag{3.87}$$

Therefore,

$$\{\gamma_E^\mu,\gamma_E^5\} = 0. \tag{3.88}$$

Using this result, the Euclidean Dirac operator is

$$\not{D}_E \gamma_E^5 = \gamma_E^\mu D_{E\mu} \gamma_E^5 = \gamma_E^\mu \gamma_E^5 D_{E\mu} = -\gamma_E^5 \gamma_E^\mu D_{E\mu} = -\gamma_E^5 \not{D}_E, \tag{3.89}$$

where we use $[D_{E\mu}, \gamma_E^5] = 0$, because covariant derivative is simple derivative in spinor space and γ_E^5 is constant matrix in spinor space. Therefore,

$$\{\not{D}_E, \gamma_E^5\} = 0. \tag{3.90}$$

Thus each eigenfunction $\psi_n(x_E)$ with a positive eigenvalue $\lambda_n > 0$, there exists another eigenfunction

$$\psi_{-n}(x_E) \equiv \gamma_E^5 \psi_n(x_E), \tag{3.91}$$

with the opposite sign eigenvalue

$$\not{D}_E \psi_{-n}(x_E) = -\lambda_n \psi_{-n}(x_E), \qquad \lambda_{-n} = -\lambda_n. \tag{3.92}$$

Therefore, non-vanishing eigenvalues appear in the spectrum are paired and cancel.

The remaining eigenfunctions with $\lambda = 0$ are called *zero modes*. Let us denote them by $\psi_{0,k}(x_E)$, $k = 1, \ldots, n_0$. They can be decomposed into definite chirality components using the projectors

$$P_\pm \equiv \frac{1 \pm \gamma_E^5}{2}, \qquad \psi_{0,k}^\pm(x_E) \equiv P_\pm \psi_{0,k}(x_E). \tag{3.93}$$

In particular,

$$\gamma_E^5 \psi_{0,k}^\pm(x_E) = \pm \psi_{0,k}^\pm(x_E). \tag{3.94}$$

For $\lambda_m \neq 0$,

$$\int d^4x_E\, \psi_m^\dagger \gamma_E^5 \psi_m = \int d^4x_E\, \psi_m^\dagger \psi_{-m} = 0, \tag{3.95}$$

since ψ_m and ψ_{-m} have different eigenvalues of the Hermitian operator $\mathcal{D}_E$ and are therefore orthogonal. Therefore only zero modes contribute to (3.84), and we obtain

$$\begin{aligned} Q &= -\sum_k \int d^4x_E\, \psi_{0,k}^\dagger \gamma_E^5 \psi_{0,k} \\ &= -\sum_k \int d^4x_E \left[(\psi_{0,k}^+)^\dagger \psi_{0,k}^+ - (\psi_{0,k}^-)^\dagger \psi_{0,k}^- \right] \\ &= n_- - n_+, \end{aligned} \tag{3.96}$$

where $n_\pm$ is the number of zero modes of chirality $\pm$ in the given background gauge field, and we used orthonormality of the eigenfunctions. We see that Q must be an integer and cannot change under smooth variations of the background gluon field. This is why Q is called the topological charge. This is a special case of the celebrated Atiyah-Singer index theorem for the Euclidean Dirac operator.

Derivation

Consider

$$\begin{aligned}\psi^{\dagger}_{E0,k}\gamma^5_E\psi_{E0,k} &= (\psi^+_{E0,k}+\psi^-_{E0,k})^{\dagger}\gamma^5_E(\psi^+_{E0,k}+\psi^-_{E0,k})\\ &= (\psi^+_{E0,k}+\psi^-_{E0,k})^{\dagger}(\psi^+_{E0,k}-\psi^-_{E0,k})\\ &= \left(\psi^+_{E0,k}\right)^{\dagger}\psi^+_{E0,k}-\left(\psi^-_{E0,k}\right)^{\dagger}\psi^-_{E0,k},\end{aligned} \tag{3.97}$$

where $P_\pm$ are projection operators satisfying $P_+ + P_- = \mathbb{I}$ and $P_\pm^2 = P_\pm$. This is simply the standard chiral decomposition.

3.4.1 Classification of the gluon fields with the topological charge

Since the topological charge Q takes only integer values and is invariant under continuous deformations of the gauge field, it provides a convenient topological label for classifying gauge-field configurations. How is this classification implemented explicitly?

We restrict attention to gluon fields with finite Euclidean action. Finite action requires the field strength to fall off sufficiently fast at infinity so that

$$\int d^4x_E\,\mathrm{Tr}\big(G_{E\mu\nu}G_E^{\mu\nu}\big) < \infty, \tag{3.98}$$

which in particular implies $G_{E\mu\nu}(x_E) = \mathcal{O}(1/r^2)$ as $r \equiv |x_E| \to \infty$. As a result, the gauge field approaches a pure gauge configuration at infinity. Therefore

$$\lim_{|\mathbf{x}_E|\to\infty} A_{E\mu}(\mathbf{x}_E) = \frac{i}{g}U\partial_{E\mu}U^{-1}. \tag{3.99}$$

This is precisely the winding-number classification of the asymptotic pure gauge U (cf. (3.69)): at infinity one has a map $U : S^3 \to SU(3)$, whose homotopy class is labeled by an integer. We already know that the homotopy group is

$$\pi_3(S^3) = \mathbb{Z}. \tag{3.100}$$

This shows that finite-action Euclidean gauge fields in QCD decompose into topologically distinct sectors labeled by an integer $Q \in \mathbb{Z}$. This integer is the Pontryagin index (topological charge) of the gauge field.

3.4.2 Topological charge of the instantons

Since the instanton and anti-instanton solutions (3.70) approach a pure gauge at infinity (the boundary condition (3.99)), we can evaluate their topological charges. Since the instanton field strength is self-dual, its topological charge can be computed directly

$$\begin{aligned}Q_I &= \frac{g^2}{16\pi^2}\int d^4x_E\,\mathrm{Tr}\left[G^{(I)}_{E\mu\nu}\tilde{G}^{(I)}_{E\mu\nu}\right]\\ &= \frac{g^2}{16\pi^2}\int d^4x_E\,\mathrm{Tr}\left[G^{(I)}_{E\mu\nu}G^{(I)}_{E\mu\nu}\right] = 1,\end{aligned} \tag{3.101}$$

where we use the instanton action (3.79). Similarly, for the anti-instanton (which is anti-self-dual), one finds $Q_{\bar{I}} = -1$.

3.5 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, we now demonstrate that these non-trivial properties force us to add another term to the QCD Lagrangian.

Let $\mathcal{O}[A, q, \bar{q}]$ be an operator consisting of quarks and gluons. 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 $\mathcal{V}_1 \subset \mathcal{V}$. Using the Euclidean path integral, we can write an expression for this VEV in the most general form as

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\int \mathcal{D}q\mathcal{D}\bar{q}\mathcal{D}A\, \mathcal{O}[A, q, \bar{q}]e^{-S_E[A,q,\bar{q}]}}{\int \mathcal{D}q\mathcal{D}\bar{q}\mathcal{D}A\, e^{-S_E[A,q,\bar{q}]}}, \tag{3.102}$$

where $|0\rangle$ is the vacuum state which is yet to be specified. Since the following argument concerns only the gauge-field topological sectors, we suppress the quark integrations and write the path integral schematically over A only.

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. Then one can write the VEV as

$$\langle 0|\mathcal{O}|0\rangle = \frac{\sum_Q \omega(Q) \int \mathcal{D}A_Q\, \mathcal{O}[A]e^{-S_E[A]}}{\sum_Q \omega(Q) \int \mathcal{D}A_Q\, e^{-S_E[A]}}, \tag{3.103}$$

where the sum is from $Q = -\infty$ to $Q = +\infty$, and $\omega(Q)$ is an unknown weight function. Let us split the Euclidean spacetime volume into two widely separated regions,

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2. \tag{3.104}$$

For configurations in which $\mathcal{V}_1$ and $\mathcal{V}_2$ are well separated, we may write

$$S_E[A; \mathcal{V}] \simeq S_E[A; \mathcal{V}_1] + S_E[A; \mathcal{V}_2], \qquad Q[A; \mathcal{V}] = Q[A; \mathcal{V}_1] + Q[A; \mathcal{V}_2]. \tag{3.105}$$

Although $Q[A; \mathcal{V}]$ is an integer for finite-action configurations on all of $\mathbb{R}^4$, the partial integrals $Q[A; \mathcal{V}_1]$ and $Q[A; \mathcal{V}_2]$ need not be integers because $\mathcal{V}_1$ and $\mathcal{V}_2$ have an artificial boundary between them. Nevertheless, for instanton-like configurations whose topological charge density is localized, and for widely separated regions, $Q[A; \mathcal{V}_1]$ and $Q[A; \mathcal{V}_2]$ are well approximated by integers. The factorization of the topological charge allows us to factorize the path integral measure as

$$\begin{aligned}\sum_Q \omega(Q) \int \mathcal{D}A_Q &= \sum_Q \omega(Q) \left(\sum_{Q_1} \int \mathcal{D}A^{(\mathcal{V}_1)}_{Q_1} \sum_{Q_2} \int \mathcal{D}A^{(\mathcal{V}_2)}_{Q_2} \right) \cdot \delta_{Q,Q_1+Q_2} \\ &= \sum_{Q_1} \sum_{Q_2} \omega(Q_1 + Q_2) \int \mathcal{D}^{(\mathcal{V}_1)}_{A_{Q_1}} \int \mathcal{D}^{(\mathcal{V}_2)}_{A_{Q_2}}. \end{aligned} \tag{3.106}$$

So the VEV becomes

$$\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} = \frac{\sum_{Q_1}\sum_{Q_2}\omega(Q_1+Q_2)\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}\mathcal{O}[A_{Q_1}]e^{-S_E[\mathcal{V}_1]}\int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)}e^{-S_E[\mathcal{V}_2]}}{\sum_{Q_1}\sum_{Q_2}\omega(Q_1+Q_2)\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}e^{-S_E[\mathcal{V}_1]}\int \mathcal{D}A_{Q_2}^{(\mathcal{V}_2)}e^{-S_E[\mathcal{V}_2]}}. \tag{3.107}$$

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 (3.107), one can observe that this is possible if

$$\omega(Q_1+Q_2) = \omega(Q_1)\cdot\omega(Q_2). \tag{3.108}$$

This functional equation on $\mathbb{Z}$ implies

$$\omega(Q) = e^{cQ}. \tag{3.109}$$

Requiring $\omega(Q)$ to remain bounded for both $Q \to +\infty$ and $Q \to -\infty$ forces $\mathrm{Re}(c) = 0$, hence $c = i\theta$ with $\theta \in \mathbb{R}$. Since $Q \in \mathbb{Z}$, the parameter is 2π-periodic, $\theta \sim \theta + 2\pi$,

$$\omega(Q) = e^{i\theta Q}, \qquad \theta \in \mathbb{R} \bmod 2\pi. \tag{3.110}$$

where we have introduced the θ-parameter. This θ-parameter should be real since the weight function $\omega(Q)$ should be finite for all $Q \in (-\infty, +\infty)$. Then the Q_2 dependent terms cancel in (3.107),

$$\begin{aligned}\langle 0|\mathcal{O}|0\rangle_{\mathcal{V}} &= \frac{\sum_{Q_1} e^{iQ_1\theta}\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}\,\mathcal{O}[A_{Q_1}]e^{-S_E[\mathcal{V}_1]}}{\sum_{Q_1} e^{iQ_1\theta}\int \mathcal{D}A_{Q_1}^{(\mathcal{V}_1)}\,e^{-S_E[\mathcal{V}_1]}} \\ &= \frac{\int \mathcal{D}A\,\mathcal{O}[A]e^{-S'_E}}{\int \mathcal{D}A\,e^{-S'_E}},\end{aligned} \tag{3.111}$$

where $S'_E \equiv S_E[A] - i\theta\,Q[A]$. It implies that summing over the topological charges is equivalent to adding a θ-dependent parameter into the action. By using the definition of Q (3.83), the new action becomes explicitly

$$S'_E = S_E - \frac{i\theta g^2}{16\pi^2}\int d^4x_E\,\mathrm{Tr}\left[G_{E\mu\nu}\tilde{G}_E^{\mu\nu}\right]. \tag{3.112}$$

Rotating back to Minkowski space using our Wick-rotation conventions (2.105).

$$S = iS_E \quad\to\quad S' = S + \frac{\theta g^2}{16\pi^2}\int d^4x\,\mathrm{Tr}\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right]. \tag{3.113}$$

We conclude that the QCD Lagrangian must allow an additional CP-violating term,

$$\begin{aligned}\mathcal{L}_{QCD} &= -\frac{1}{2}\mathrm{Tr}\left[G_{\mu\nu}G^{\mu\nu}\right] + \frac{\theta g^2}{16\pi^2}\mathrm{Tr}\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right] \\ &= -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + \frac{\theta g^2}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a\mu\nu}.\end{aligned} \tag{3.114}$$

The additional term is called the θ-term. This analysis shows that QCD contains a hidden parameter θ (defined modulo 2π), which weights the contributions of different topological sectors.

3.6 Theta vacua

Finally, we define the vacuum states of QCD. We consider a set of "candidate" vacua $|n\rangle$ labeled by the winding number $n \in \mathbb{Z}$. The physical vacuum depends on the parameter θ and is given by

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle. \tag{3.115}$$

This state is called a θ vacuum. Each θ value corresponds to another vacuum state. More importantly, different values of θ label distinct superselection sectors: no gauge-invariant local operator can connect $|\theta_1\rangle$ to $|\theta_2\rangle$ for $\theta_1 \neq \theta_2$. To see this, consider the time-ordered product $T(\mathcal{O}_1\mathcal{O}_2\cdots)$ of gauge-invariant operators. Let $|\theta_1\rangle$ and $|\theta_2\rangle$ be two θ vacua with $\theta_1 \neq \theta_2$. Let $|\theta_1\rangle$ and $|\theta_2\rangle$ are two vacua with $\theta_1 \neq \theta_2$. Then

$$\langle\theta_1|T(\mathcal{O}_1\mathcal{O}_2\cdots)|\theta_2\rangle = \sum_{m,n} e^{i(n\theta_2 - m\theta_1)}\langle m|T(\mathcal{O}_1\mathcal{O}_2\cdots)|n\rangle. \tag{3.116}$$

Since $\langle m|T(\mathcal{O}_1\mathcal{O}_2\cdots)|n\rangle$ depends only on the difference $Q \equiv n - m$, we may write

$$\langle m|T(\mathcal{O}_1\mathcal{O}_2\cdots)|n\rangle \equiv F(Q). \tag{3.117}$$

Then (3.116) becomes

$$\begin{aligned}\langle\theta_1|T(\mathcal{O}_1\mathcal{O}_2\cdots)|\theta_2\rangle &= \sum_{n\in\mathbb{Z}} e^{in(\theta_2-\theta_1)} \sum_{Q\in\mathbb{Z}} e^{iQ\theta_1}\, F(Q) \\ &= 2\pi\,\delta(\theta_2 - \theta_1) \sum_{Q\in\mathbb{Z}} e^{iQ\theta_1}\, F(Q).\end{aligned} \tag{3.118}$$

This equation is zero if $\theta_1 \neq \theta_2$. This superselection rule indicates that θ is a genuine parameter labeling the Yang–Mills vacuum: different θ correspond to distinct superselection sectors (equivalently, different choices of the θ vacuum).

4 The Strong CP Problem

4.1 The QCD Lagrangian

We start by writing the QCD Lagrangian in the flavor basis where the quark mass matrix is diagonal,

$$\mathcal{L}_{QCD} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + \sum_q (i\bar{q}\not{D}q - m_q\bar{q}q). \tag{4.1}$$

Here each q denotes a Dirac spinor field for a given quark flavor. Its Dirac adjoint is $\bar{q} \equiv q^\dagger\gamma^0$, where $\{\gamma^\mu\}$ are the gamma matrices. The $G^a_{\mu\nu}$ is the gluon field strength given by

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu, \tag{4.2}$$

where g_s is the gauge coupling constant. The eight gluons are described by the gauge fields A^a_μ, where the adjoint index $a = 1, \ldots, 8$ labels the color generators. The structure constants appear in the non-Abelian term because the gauge field is Lie-algebra valued and the generators satisfy $[T^a, T^b] = if^{abc}T^c$. Equivalently, in the adjoint representation one has $(T^a_{\text{adj}})^{bc} = -if^{abc}$. The kinetic term for the quarks is

$$\begin{aligned} \bar{q}\not{D}q &= \bar{q}\,\gamma^\mu D_\mu q \\ &= \bar{q}_i\,\gamma^\mu (D_\mu)_{ij}\, q_j \\ &= \bar{q}_i\,\gamma^\mu \left(\delta_{ij}\partial_\mu - ig_s A^a_\mu (T^a)_{ij}\right) q_j, \end{aligned} \tag{4.3}$$

where i and j are the color indices of the quarks. Here $\{T^a\}$ are the generators of $SU(3)_c$ in the fundamental representation. This reflects the fact that quarks transform in the fundamental representation of $SU(3)_c$. We choose $T^a = \lambda^a/2$, where $\{\lambda^a\}$ are the Gell–Mann matrices.

4.2 Quark masses in the Standard Model

4.2.1 Electroweak symmetry breaking

Recall the Standard Model (SM) is based on the gauge group

$$SU(3)_c \otimes SU(2)_{EW} \otimes U(1)_Y. \tag{4.4}$$

The fermion masses are generated via the Electroweak symmetry breaking where the subgroup $SU(2)_{EW} \otimes U(1)_Y$ is spontaneously broken to $U(1)_{EM}$.

The gauge group $SU(2)_{EW} \otimes U(1)_Y$ describes the Electroweak Theory which unifies the weak interactions and the electromagnetism. The $SU(2)$ part consists of three gauge bosons $\{W^a_\mu\}^3_{a=1}$ and $U(1)_Y$ has B_μ. For the electroweak symmetry breaking to occur, one also needs the Higgs multiplet H which is a complex doublet that transforms as an $SU(2)$ doublet with hypercharge $Y = +1/2$. The Lagrangian is

$$\mathcal{L}_{EW} = -\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + (D^\mu H)^\dagger (D_\mu H) - V(H^\dagger H), \tag{4.5}$$

where $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are field strength tensors for the $SU(2)_{EW}$ and $U(1)_Y$ respectively. The covariant derivative is given by

$$D_\mu H = \left(\partial_\mu - ig\, W^a_\mu \tau^a - i\frac{g'}{2}B_\mu\right) H, \tag{4.6}$$

where g and g' are the gauge couplings of $SU(2)$ and $U(1)_Y$, respectively. The $\tau^a = \sigma^a/2$ are the $SU(2)$ generators and the coefficient in front of the B_μ term comes from the hypercharge of the Higgs doublet.

The potential $V(H^\dagger H)$ is such that the Higgs doublet obtains a VEV which without loss of generality can be chosen as

$$\langle H \rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.7}$$

where $v \sim 250\,\text{GeV}$. One can redefine the Higgs field around this VEV as

$$H = \frac{1}{\sqrt{2}} \exp i\frac{\pi^a \sigma^a}{v} \begin{pmatrix} 0 \\ v+h \end{pmatrix}, \tag{4.8}$$

where h is a scalar excitation which will be identified with the Higgs particle. Note that in the case of spontaneous breaking of a global symmetry, π^a were the Goldstone bosons. However, in this case, the broken symmetry is a gauge symmetry. So the transformation law for $\{\pi^a\}$ describes a gauge redundancy and we can take this gauge freedom to set $\pi^a = 0$, called the Unitary gauge. This gauge fixing is very convenient in the study of spontaneous breaking of gauge symmetries since it removes the kinetic mixing between π and the gauge bosons.

In the unitary gauge, the kinetic term becomes

$$(D^\mu H)^\dagger (D_\mu H) = \frac{g^2 v^2}{8} \left[W^1_\mu W^{1\mu} + W_{\mu 2} W^{2\mu} + \left(\frac{g'}{g} B^\mu - W^{3\mu} \right) \left(\frac{g'}{g} B_\mu - W^3_\mu \right) + h\text{-terms} \right].$$

This result shows that the gauge bosons W^3_μ and B_μ mix with each other. The $W^{1,2}_\mu$ and certain combination of W^3_μ and B_μ are massive. To find the spectrum we need to diagonalize the mass terms. This can be achieved by defining

$$Z_\mu \equiv \cos\theta_W\, W^3_\mu - \sin\theta_W\, B_\mu, \qquad A_\mu \equiv \sin\theta_W\, W^3_\mu + \cos\theta_W\, B_\mu, \qquad \tan\theta_W \equiv \frac{g'}{g}. \tag{4.9}$$

Then one can show that the Lagrangian contains terms like

$$\mathcal{L}_{EW} \supset -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu, \qquad m_Z = \frac{gv}{2\cos\theta_w}, \tag{4.10}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$. Here, A_μ is the photon of electromagnetism, while Z_μ is a massive spin-1 boson known as the Z-boson.

Redefine W_μ field as eigenstates which are charged under the electromagnetism,

$$W^\pm_\mu \equiv \frac{1}{2} \left(W^1_\mu \mp W^2_\mu \right). \tag{4.11}$$

Then the Lagrangian in this field basis,

$$\mathcal{L}_{EW} \supset -\frac{1}{2} W^+_{\mu\nu} W^{-\,\mu\nu} + m_W^2\, W^+_\mu W^{-\,\mu}, \qquad W^\pm_{\mu\nu} \equiv \partial_\mu W^\pm_\nu - \partial_\nu W^\pm_\mu, \tag{4.12}$$

with

$$m_W = \frac{gv}{2}. \tag{4.13}$$

These are the W bosons of the Standard Model.

In summary, when the gauge symmetry is spontaneously broken, the W and the Z bosons should acquire an additional degree of freedom. This comes from the complex Higgs doublet degree of freedom before the symmetry breaking. Its three degrees of freedom are transferred to the gauge bosons while the remaning degree of freedom is the scalar excitation h which is nothing but the Higgs boson of the Standard Model.

$$\text{D.O.F} : \{W^a_\mu,\, B_\mu,\, H\} = (3\times 2,\ 2,\ 4) \ \rightarrow\ \{W^\pm_\mu,\, Z_\mu,\, A_\mu,\, h\} = (2\times 3,\ 3,\ 2,\ 1). \tag{4.14}$$

This is often summarized by saying that the gauge bosons “eat” the Goldstone modes and become massive.

4.2.2 Fermions in the Standard Model

The electroweak theory is chiral and maximally parity-violating, since the $SU(2)_{EW}$ gauge bosons couple only to left-handed fermions. Within the Standard Model this chiral structure is an input, the model does not predict why Nature chooses it.

The above statement means that the left-handed fermions are arranged into $SU(2)_{EW}$ doublets that transform under the fundamental representation, whereas the right-handed fermions are $SU(2)_{EW}$ singlets. The lepton doublets

$$L_i = \left\{ \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \right\}, \tag{4.15}$$

where e, μ, τ denote the electron, muon, and tau, respectively, and ν's are the corresponding neutrinos. The index i labels the generation. The right-handed fermions are represented by

$$e_R^i = \{e_R, \mu_R, \tau_R\}, \quad u_R^i = \{u_R, c_R, t_R\}, \quad d_R^i = \{d_R, s_R, b_R\}. \tag{4.16}$$

One needs to specify the hypercharges. These are given by

$$\begin{array}{lll} L: \ -\dfrac{1}{2} & e_R: \ -1 & Q: \ \dfrac{1}{6} \\ u_R: \ \dfrac{2}{3} & d_R: \ -\dfrac{1}{3} & H: \ \dfrac{1}{2}, \end{array}$$

where Q is the left-handed quarks doublet.

4.2.3 Fermion masses

Now recall that a Dirac mass term for a fermion can be written as

$$\mathcal{L}_{Dirac} \supset m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L). \tag{4.17}$$

However, such a term cannot be written as a renormalizable gauge-invariant operator, because ψ_L and ψ_R transform differently under $SU(2)_{EW} \times U(1)_Y$.

For the electron one can write the Yukawa interaction

$$\mathcal{L}_{\text{Yukawa}} = -y_e \, \bar{L}_e \, H \, e_R + \text{h.c.}, \tag{4.18}$$

where y_e is a dimensionless coupling called Yukawa coupling constant. After the electroweak symmetry breaking, H gets a VEV and this term becomes

$$\mathcal{L}_{\text{Yukawa}} \supset -\frac{v y_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) = -m_e \bar{e} e, \qquad m_e = \frac{y_e v}{\sqrt{2}}. \tag{4.19}$$

This is a mass term for the electron where $m_e = yv/\sqrt{2}$. With terms like these, the charged leptons e, μ, τ and the down-type quarks d, s, b get their masses.

To give the up-type quarks u, c, t their masses, one uses an interaction of the form

$$\mathcal{L}_{\text{Yuakawa}} \supset -y_u \, \bar{Q} \, \tilde{H} \, u_R + \text{h.c.}, \qquad \tilde{H} \equiv i\sigma^2 H^*. \tag{4.20}$$

Therefore, all the quark masses can be generated via the Yukawa interaction

$$\begin{aligned}\mathcal{L}_{\text{Yuakawa}} &= -y^d_{ij}\bar{Q}^i H d^j_R - y^u_{ij}\bar{Q}^i \tilde{H} u^j_R + h.c. \\ &\to -\frac{v}{\sqrt{2}}\left[y^d_{ij}\bar{Q}^i \begin{pmatrix}0\\1\end{pmatrix} d^j_R + y^u_{ij}\bar{Q}^i \begin{pmatrix}1\\0\end{pmatrix} u^j_R + h.c.\right],\end{aligned} \tag{4.21}$$

where there is implicit summation over the generation indices i and j. One can rewrite this interaction by introducing vectors,

$$\mathbf{u}_L = \{u_L, c_L, t_L\}, \qquad \mathbf{d}_L = \{d_L, s_L, b_L\} \tag{4.22}$$

$$\mathbf{u}_R = \{u_R, c_R, t_R\}, \qquad \mathbf{d}_R = \{d_R, s_R, b_R\}, \tag{4.23}$$

then

$$\mathcal{L}_{\text{mass}} = -\frac{v}{\sqrt{2}}\left(\mathbf{d}^\dagger_L y_d \mathbf{d}_R + \mathbf{u}^\dagger_L y_u \mathbf{u}_R\right) + h.c., \tag{4.24}$$

where y_u and y_d are called up and down Yukawa matrices respectively. Note that these matrices are not diagonal in general which implies that the mass matrix is also not diagonal. The expression in (4.24) is written in the flavour basis since it is written in terms of quark flavours u, d, s, c, b and t.

It is possible to diagonalize the mass matrix though. In general, the Yukawa matrices are not hermitian. However, $y_d y_d^\dagger$ and $y_u y_u^\dagger$ are. So as a consequence of the finite dimensional spectral theorem, they can be diagonalized via the unitary matrices, and the resulting diagonal matrix has only real entries. Therefore, one can write

$$y_d y_d^\dagger = U_d M_d^2 U_d^\dagger, \qquad y_u y_u^\dagger = U_u M_u^2 U_u^\dagger, \tag{4.25}$$

where U_u and U_d are unitary, and M_u and M_d are diagonal matrices. Using this diagonalization, the Yukawa matrices as

$$y_d = U_d M_d K_d^\dagger, \qquad y_u = U_u M_u K_u^\dagger, \tag{4.26}$$

where K_u and K_d are another unitary matrices. With these definitions, the mass term (4.24) takes the form

$$\mathcal{L}_{\text{mass}} = -\frac{v}{\sqrt{2}}\left(\mathbf{d}^\dagger_L U_d M_d K_d^\dagger \mathbf{d}_R + \mathbf{u}^\dagger_L U_u M_u K_u^\dagger \mathbf{u}_R\right) + h.c.\ . \tag{4.27}$$

To get a Lagrangian in the so-called mass basis, we make a change of basis as

$$\mathbf{d}_R \to K_d \mathbf{d}_R, \quad \mathbf{d}_L \to U_d \mathbf{d}_L, \quad \mathbf{u}_R \to K_u \mathbf{u}_R, \quad \mathbf{u}_L \to U_u \mathbf{u}_L. \tag{4.28}$$

In this basis, the mass matrix is diagonal,

$$\mathcal{L}_{\text{mass}} = -\sum_{i=1}^{3}\left(m_i^d d^\dagger_{L,i} d_{R,i} + m_i^u u^\dagger_{L,i} u_{R,i}\right) + h.c., \tag{4.29}$$

where $m_i^{d,u}$ is the elements of the diagonal matrices of $m^{u,d} = M_{u,d}/\sqrt{2}$.

4.2.4 Cabibbo-Kobayashi-Maskawa (CKM) matrix

The interactions of the quarks with the electroweak gauge bosons are flavour-diagonal which means that the gauge interactions do not mix the flavours. However, by performing the change of basis given in (4.28), these couplings are modified by this change of basis, and the modifications is encoded in the Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$V \equiv U_u^\dagger U_d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{4.30}$$

This matrix is a 3×3 complex unitary matrix, so it has 9 free parameters, 3 angles and 6 phases. However, many of these parameters can be eliminated by noting that there is still a $U(1)^6$ global symmetry which corresponds to separate $U(1)_V$ rotations for each of the six quarks,

$$\begin{aligned} d_L^j &\to e^{i\alpha_j} d_L^j, \qquad & u_L^j &\to e^{i\beta_j} u_L^j, \\ d_R^j &\to e^{i\alpha_j} d_R^j, \qquad & u_R^j &\to e^{i\beta_j} u_R^j, \end{aligned} \tag{4.31}$$

where there is no summation over j.This would have eliminated all the 6 phases, however if all the angles are equal $\alpha_j = \beta_j$, then V does not change. Thus, the 5 phases can be eliminated this way but overall phase is survived. In the end we are left with 3 angles, and one phase.

The standard parametrization for the CKM matrix is

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{4.32}$$

where $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. The three phases $\{\theta_{12}, \theta_{23}, \theta_{13}\}$ are rotation angles in the ij-flavour plane. The numerical values of the three angles and the phase are

$$\begin{aligned} \sin\theta_{12} &= 0.22500, \\ \sin\theta_{23} &= 0.01482, \\ \sin\theta_{13} &= 0.00369, \\ \delta &= 1.144 \pm 0.027. \end{aligned} \tag{4.33}$$

If the CKM matrix were real, there would be no CP violation. Therefore, the phase δ measures the amount of CP violation.

CKM matrix real

In the mass basis

$$\mathcal{L} = -\frac{g}{\sqrt{2}}\left(\bar{u}_{iL}\gamma^\mu V_{ij} d_{jL} W_\mu^\dagger + \bar{d}_{jL}\gamma^\mu V_{ij}^* u_{iL} W_\mu^-\right). \tag{4.34}$$

The CP transformation is

$$(CP)\psi_L(t,\mathbf{x})(CP)^{-1} = e^{i\delta_\psi} i\gamma^0\gamma^2\psi_L^*(t,-\mathbf{x}) \tag{4.35}$$

$$(CP)[\bar{u}_{iL}\gamma^\mu d_{jL}](t,\mathbf{x})(CP)^{-1} = \bar{d}_{jL}\gamma_\mu u_{iL}(t,-\mathbf{x}), \tag{4.36}$$

$$(CP)W_\mu^\pm(t,\mathbf{x})(CP)^{-1} = -W^{\mp\mu}(t,-\mathbf{x}). \tag{4.37}$$

Then

$$\begin{aligned}
&(CP)\left[\bar{u}_{iL}\gamma^\mu V_{ij} d_{jL} W_\mu^+\right](t,x)(CP)^{-1} \\
&\quad = V_{ij}^*(CP)[\bar{u}_{iL}\gamma^\mu d_{jL}](t,x)(CP)^{-1}(CP)W_\mu^+(t,x)(CP)^{-1} \\
&\quad = V_{ij}^*[\bar{d}_{jL}\gamma_\mu u_{iL}](t,-\mathbf{x})[-W^{-\mu}(t,-\mathbf{x})].
\end{aligned} \tag{4.38}$$

It implies

$$\bar{u}_L\gamma^\mu V d_L W^+ \quad \leftrightarrow \quad \bar{d}_L\gamma^\mu V^* u_L W^-. \tag{4.39}$$

If V is real number,

$$V_{ij}^* = V_{ij} \quad \rightarrow \quad (CP)\mathcal{L}(t,\mathbf{x})(CP)^{-1} = \mathcal{L}(t,-\mathbf{x}). \tag{4.40}$$

4.3 Strong CP violation

In the previous section, we have seen how to diagonalize the quark mass matrix, but we haven't talked about the consequences of this operation apart from the CKM matrix. Notice that the transformations that we performed, in particular the ones in (4.28) are chiral. In particular, their axial components are anomalous: the path-integral measure is not invariant under axial rotations, so the Lagrangian is shifted by the chiral anomaly. In this section, our goal is to compute the induced shift of the QCD θ term that arises when we diagonalize the quark mass matrix.

4.3.1 Chiral rotations with multiple generations

Consider the right-handed $\{\psi_R^i\}$ and left-handed $\{\psi_L^i\}$ Weyl fermions where i is the generation index. Now consider the chiral transformations

$$\boldsymbol{\psi}_R \quad \rightarrow \quad R\boldsymbol{\psi}_R, \qquad \boldsymbol{\psi}_L \quad \rightarrow \quad L\boldsymbol{\psi}_L. \tag{4.41}$$

In the component form

$$\psi_R^i \quad \rightarrow \quad R^{ij}\psi_R^j, \qquad \psi_L^i \quad \rightarrow \quad L^{ij}\psi_L^j. \tag{4.42}$$

Since L and R are unitary, each of them can be diagonalized by a unitary similarity transformation. Then we can write

$$\boldsymbol{\psi}_R \quad \rightarrow \quad W_R R_d W_R^\dagger \boldsymbol{\psi}_R, \qquad \boldsymbol{\psi}_L \quad \rightarrow \quad W_L L_d W_L^\dagger \boldsymbol{\psi}_L, \tag{4.43}$$

where W_R and W_L are unitary, R_d and L_d are unitary and diagonal matrices. Now perform a change of basis,

$$\psi'_L \equiv W_L^\dagger \psi_L, \qquad \psi'_R \equiv W_R^\dagger \psi_R. \tag{4.44}$$

In this basis, the chiral transformations in (4.41) act diagonally in generation space,

$$\psi'_R \quad \to \quad R_d \psi'_R, \qquad \psi'_L \quad \to \quad L_d \psi'_L. \tag{4.45}$$

One can easily find that the transformations do not mix the generations anymore. The moral of the story is we can always go to a basis where a general chiral transformation parametrized by L and R becomes diagonal. So by introducing the Dirac fermions

$$\Psi_i \equiv \begin{pmatrix} \psi_L^i \\ \psi_R^i \end{pmatrix}, \tag{4.46}$$

for each generation, we can write the transformation law as

$$\Psi_i \;\to\; \begin{pmatrix} (L_d)_{ii} & 0 \\ 0 & (R_d)_{ii} \end{pmatrix} \Psi_i = \begin{pmatrix} e^{il_i} & 0 \\ 0 & e^{ir_i} \end{pmatrix} \Psi_i = \exp\left[i \begin{pmatrix} l_i & 0 \\ 0 & r_i \end{pmatrix} \right] \Psi_i, \tag{4.47}$$

where $(L_d)_{ii}$ and $(R_d)_{ii}$ are the i-th components of the L_d and R_d diagonal matrices, with l_i and r_i being real. We can write this transformation as

$$\Psi_j \quad \to \quad \exp\left(i\frac{\alpha_j}{2}\gamma_5\right) \exp\left(i\frac{\beta_j}{2}\right) \Psi_j, \tag{4.48}$$

with

$$\alpha_j = r_j - l_j, \qquad \beta_j = r_j + l_j, \tag{4.49}$$

This decomposition separates a vector-like phase rotation from an axial (chiral) rotation. Since we are interested in the anomalous axial part, we can ignore the vector-like rotation in what follows.

Using the single-flavor anomaly result (cf. (2.112)), the net shift from independent axial rotations of all generations adds up,

$$\Delta\mathcal{L} = \left(\sum_i \alpha_i \right) \frac{g_s^2}{32\pi^2} G^a_{\mu\nu} \tilde{G}^{a\mu\nu}. \tag{4.50}$$

The phase coefficient can be written as

$$\sum_i \alpha_i = \sum_i (r_i - l_i) = \arg\left[\prod_i e^{ir_i} e^{-il_i} \right] = \arg\left[\det\left(L_d^\dagger R_d\right)\right] = \arg\left[\det\left(L^\dagger R\right)\right]. \tag{4.51}$$

Here we used (4.47), and the fact that $\det(WXW^\dagger) = \det(X)$ for unitary W, so the determinant is unchanged by the diagonalization in (4.43). Therefore we have found that the chiral rotation modifies the Lagrangian by

$$\Delta\mathcal{L} = \frac{g_s^2\,\theta_q}{32\pi^2} G^a_{\mu\nu} \tilde{G}^{a\mu\nu}, \qquad \theta_q \equiv \arg\left[\det\left(L^\dagger R\right)\right] \in \mathbb{R}. \tag{4.52}$$

4.3.2 Phase induced by the diagonalization of the quark mass matrix

We have seen that the Yukawa matrices can be expressed as (4.26)

$$y_d = U_d M_d K_d^\dagger, \qquad y_u = U_u M_u K_u^\dagger, \tag{4.53}$$

where M_d and M_u are diagonal and real, and the matrices $U_{u,d}$ and $K_{u,d}$ are unitary. Without loss of generality, we can also express them as

$$y_d = U_d M_d U_d^\dagger \tilde{K}_d^\dagger, \qquad y_u = U_u M_u U_u^\dagger \tilde{K}_u^\dagger, \tag{4.54}$$

where $\tilde{K}_{u,d}$ are

$$\tilde{K}_d^\dagger \equiv U_d K_d^\dagger, \qquad \tilde{K}_u^\dagger \equiv U_u K_u^\dagger. \tag{4.55}$$

Then, the mass term in (4.24) takes the form

$$\begin{aligned}\mathcal{L}_{\text{mass}} &= -\frac{v}{\sqrt{2}}\left(\mathbf{d}_L^\dagger y_d \mathbf{d}_R + \mathbf{u}_L^\dagger y_u \mathbf{u}_R\right) + h.c. \\ &= -\frac{v}{\sqrt{2}}\left(\mathbf{d}_L^\dagger U_d M_d U_d^\dagger \tilde{K}_d^\dagger \mathbf{d}_R + \mathbf{u}_L^\dagger U_u M_u U_u^\dagger \tilde{K}_u^\dagger \mathbf{u}_R\right) + h.c..\end{aligned} \tag{4.56}$$

In order to go to the mass basis, we need to perform a chiral rotation (only rotation right-handed) given by

$$\begin{array}{rclcrcl} \mathbf{u}_L & \to & \mathbf{u}_L, & & \mathbf{d}_L & \to & \mathbf{d}_L \\ \mathbf{d}_R & \to & \tilde{K}_d \mathbf{d}_R, & & \mathbf{u}_R & \to & \tilde{K}_u \mathbf{u}_R, \end{array} \tag{4.57}$$

and then a non-chiral rotation with

$$\mathbf{d}_{L,R} \to U_d \mathbf{d}_{L,R}, \qquad \mathbf{u}_{L,R} \to U_u \mathbf{u}_{L,R}. \tag{4.58}$$

The non-chiral rotation is not anomalous so it doesn't modify the Lagrangian.

Chiral and vector rotation

Consider a Dirac fermion $\mathbf{q}$. Then

$$\mathbf{q}_L = P_L \mathbf{q}, \qquad \mathbf{q}_R = P_R \mathbf{q}, \qquad P_{L,R} = \frac{1 \mp \gamma_5}{2}. \tag{4.59}$$

General transformations for flavor space act as

$$\mathbf{q}_L \to L \mathbf{q}_L, \qquad \mathbf{q}_R \to R \mathbf{q}_R. \tag{4.60}$$

The transformation (4.57) is

$$L = 1, \qquad R = \tilde{K}. \tag{4.61}$$

Define $\tilde{K} = e^{iA}$, with A is Hermitian in flavor space (and commuting with γ_5). Then

$$\mathbf{q} \to (P_L + \tilde{K} P_R)\mathbf{q} = (P_L + e^{iA} P_R)\mathbf{q} = e^{i\frac{A}{2}} e^{i\frac{A}{2}\gamma_5} \mathbf{q}, \tag{4.62}$$

where we used

$$e^{i\frac{A}{2}\gamma_5} = e^{-i\frac{A}{2}}P_L + e^{i\frac{A}{2}}P_R. \tag{4.63}$$

This is a chiral transformation in flavor space.
Now consider the transformation (4.58). This transformation corresponding to

$$L = R = U. \tag{4.64}$$

Then

$$\mathbf{q} \quad \rightarrow \quad (UP_L + UP_R)\mathbf{q} = U(P_L + P_R)\mathbf{q} = U\mathbf{q}. \tag{4.65}$$

Since the above transformation contains no γ_5, it is a non-chiral (vector-like) transformation.

From the (4.51), the phase induced by the chiral rotations is corresponding to

$$L = 1, \quad R = \begin{pmatrix} \tilde{K}_u & \\ & \tilde{K}_d \end{pmatrix}, \qquad \Psi_{L,R} = \begin{pmatrix} \mathbf{u}_{L,R} \\ \mathbf{d}_{L,R} \end{pmatrix}. \tag{4.66}$$

Therefore, the chiral anomaly part is

$$\begin{aligned} \theta_q = \arg\big[\det(L^\dagger R)\big] = \arg\big[\det(\tilde{K}_d)\cdot\det(\tilde{K}_u)\big] &= \arg\big[\det(K_d U_d^\dagger)\cdot\det(K_u U_u^\dagger)\big] \\ &= \arg\big[\det(K_d U_d^\dagger)\cdot\det(K_u U_u^\dagger)\big] + \arg\big[\det(M_d M_u)\big] \\ &= \arg\big[\det(K_d M_d U_d^\dagger)\cdot\det(K_u M_u U_u^\dagger)\big] \\ &= \arg\big[\det\big\{(y_d y_u)^\dagger\big\}\big] \\ &= -\arg\big[\det(y_d y_u)\big], \end{aligned} \tag{4.67}$$

where we have used the fact that the elements of M_d and M_u are real and positive.

4.3.3 Physical θ parameter

By using our recently derived result, we can write our final expression for the θ term in the QCD,

$$\mathcal{L}_\theta = \frac{\bar{\theta} g_s^2}{32\pi^2} G^a_{\mu\nu}\tilde{G}^{a\mu\nu}, \qquad \bar{\theta} \equiv \theta_{QCD} - \theta_q = \theta_{QCD} + \arg\big[\det(y_d y_u)\big]. \tag{4.68}$$

Here θ_{QCD} is the θ parameter of QCD that comes from the θ vacua and θ_q comes from the diagonalization of the quark mass matrix. The significance of the $\bar{\theta}$ term is that it cannot be removed by a chiral transformation. To see this explicitly we note that in QCD the current associated to the chiral transformation is not conversed,

$$\partial_\mu J^\mu_A = 2m_q \bar{q} i\gamma_5 q + \frac{g_s^2}{32\pi^2} G^a_{\mu\nu}\tilde{G}^{a\mu\nu}. \tag{4.69}$$

The first term arises due to the non-zero quark mass, while the second term comes from the chiral anomaly. Under a chiral rotation $q \to e^{i\alpha\gamma_5} q$, the θ term shifts as

$$\Delta\mathcal{L}_\theta = -\frac{2\alpha\, g_s^2}{32\pi^2}\, G^a_{\mu\nu}\tilde{G}^{a\mu\nu}, \tag{4.70}$$

while the mass term becomes

$$-\, m_q \bar{q} q \;\to\; -\, m_q \bar{q} e^{2i\alpha\gamma_5} q. \tag{4.71}$$

Choosing $\alpha = \bar{\theta}/2$ removes the $\bar{\theta}\, G\tilde{G}$ term from the Lagrangian, but then the same phase reappears in the mass term,

$$-\, m_q \bar{q} q \;\to\; -\, m_q \bar{q} e^{i\bar{\theta}\gamma_5} q = -\, m_q \cos\bar{\theta}\, \bar{q} q - i m_q \sin\bar{\theta}\, \bar{q}\gamma_5 q. \tag{4.72}$$

For this reason $\bar{\theta}$ is a physical parameter: an axial rotation can shift the coefficient of $G\tilde{G}$, but it simultaneously changes the phase of the quark mass term. Physical observables depend only on the invariant combination $\bar{\theta} = \theta_{\mathrm{QCD}} - \arg\det m_q$.

4.4 Strong CP Problem

We are now ready to state the Strong CP problem. The electric dipole moment of the neutron is defined by the Hamiltonian,

$$H = -d_n \mathbf{E} \cdot \hat{\mathbf{S}}. \tag{4.73}$$

A non-zero dipole moment has not been measured by any experiment yet. The current experimental limit is

$$|d_n| < 1.8 \times 10^{-26}\ e\,\mathrm{cm}. \tag{4.74}$$

The most precise theoretical calculation for d_n based on the QCD sum rules is

$$d_n = 2.4 \times 10^{-16}\,\bar{\theta}\ e\,\mathrm{cm}. \tag{4.75}$$

We see that the neutron electric dipole moment is proportional to the $\bar{\theta}$ parameter, hence it should be physical. Comparing experimental result and theoretical calculation yields the bound

$$|\bar{\theta}| \lesssim 10^{-10}. \tag{4.76}$$

The question of why the dimensionless parameter $\bar{\theta}$ is so small is called the strong CP problem.

A Grassmann numbers

In this section we review Grassmann numbers, also known as anticommuting numbers or Grassmann variables.

A set of Grassmann numbers $\{\theta_i\}_{i=1}^{n}$ obeys the anticommutation relations

$$\{\theta_i, \theta_j\} \equiv \theta_i\theta_j + \theta_j\theta_i = 0, \tag{A.1}$$

but they commute under addition,

$$\theta_i + \theta_j = \theta_j + \theta_i. \tag{A.2}$$

They generate the Grassmann algebra $\mathcal{G}$ over a field, which is usually taken to be $\mathbb{C}$. This algebra contains the additive identity 0 (so that $\theta_i + 0 = \theta_i$) and the multiplicative identity 1. We can multiply Grassmann elements by complex scalars, and the result is again an element of $\mathcal{G}$; in other words, $\mathcal{G}$ is a $\mathbb{C}$-algebra.

Let us consider the simplest case with $n = 1$. Since $\theta^2 = 0$, any polynomial (and more generally any analytic function) of the Grassmann variable truncates after the linear term:

$$f(\theta) = a + b\theta, \qquad a, b \in \mathbb{C}. \tag{A.3}$$

In particular, any element of the Grassmann algebra can be written in the form (A.3).

In physics, Grassmann numbers are used to represent fermionic fields. Since path integrals over fermions must yield ordinary complex numbers, we define the Grassmann integral as a linear map from $\mathcal{G}$ to $\mathbb{C}$. We also want it to share basic properties of ordinary integration, in particular linearity,

$$\int d\theta \; [c\, f(\theta) + d\, g(\theta)] = c \int d\theta\, f(\theta) + d \int d\theta\, g(\theta), \tag{A.4}$$

and invariance under a constant shift of the variable,

$$\int d\theta\, f(\theta) = \int d\theta\, f(\theta + \eta), \tag{A.5}$$

where η is a Grassmann constant (independent of θ). These requirements imply

$$\int d\theta\, 1 = 0. \tag{A.6}$$

The remaining normalization is a convention; we choose

$$\int d\theta\, \theta = 1. \tag{A.7}$$

In summary,

$$\int d\theta\, 1 = 0, \qquad \int d\theta\, \theta = 1, \tag{A.8}$$

which defines the Berezin integral.

Derivation

Suppose the linearity (A.4) holds. Then, for $f(\theta) = a + b\theta$,

$$\begin{aligned}\int d\theta\, f(\theta+\eta) &= \int d\theta\, [a + b(\theta+\eta)] \\ &= \int d\theta\, a + b\int d\theta\,\theta + b\int d\theta\,\eta \\ &= \int d\theta\, a + b\int d\theta\,\theta - b\,\eta\int d\theta\, 1,\end{aligned} \tag{A.9}$$

where we used that η anticommutes with $d\theta$, so $\int d\theta\,\eta = -\eta\int d\theta\,1$. Shift invariance demands $\int d\theta\, f(\theta+\eta) = \int d\theta\, f(\theta)$, hence $\int d\theta\,1 = 0$. Therefore,

$$\int d\theta\, f(\theta) = \int d\theta\,(a + b\theta) = b. \tag{A.10}$$

We can generalize these definitions to n Grassmann variables. A general function $f(\{\theta_i\})$ can be expanded as

$$f(\{\theta_i\}) = a + b_i\theta_i + \frac{1}{2!}c_{i_1 i_2}\theta_{i_1}\theta_{i_2} + \cdots + \frac{1}{n!}d_{i_1\cdots i_n}\theta_{i_1}\cdots\theta_{i_n}, \tag{A.11}$$

with implicit summation over repeated indices. Since the products of θ's are totally antisymmetric, the coefficients are antisymmetric as well. In particular,

$$d_{i_1\cdots i_n} = \epsilon_{i_1\cdots i_n}\, d, \qquad d \in \mathbb{C}, \tag{A.12}$$

where $\epsilon_{i_1\cdots i_n}$ is the Levi–Civita symbol with $\epsilon_{12\cdots n} = +1$. The Berezin integral over n variables is defined by extracting the coefficient of $\theta_1\cdots\theta_n$:

$$\int d^n\theta\, f(\{\theta_i\}) = d, \tag{A.13}$$

with the convention

$$d^n\theta \equiv d\theta_n \cdots d\theta_1. \tag{A.14}$$

Derivation

One readily finds

$$\int d\theta_i\, 1 = 0, \qquad \int d\theta_i\,\theta_j = \delta_{ij}. \tag{A.15}$$

Moreover, by antisymmetry,

$$\theta_{i_1}\cdots\theta_{i_n} = \epsilon_{i_1\cdots i_n}\,\theta_1\cdots\theta_n, \tag{A.16}$$

so that

$$\int d^n\theta\,\theta_{i_1}\cdots\theta_{i_n} = \epsilon_{i_1\cdots i_n}\int d^n\theta\,\theta_1\cdots\theta_n = \epsilon_{i_1\cdots i_n}. \tag{A.17}$$

Therefore,

$$\int d^n\theta\, f(\{\theta_i\}) = \frac{1}{n!}\,(\epsilon_{i_1\cdots i_n} d) \int d^n\theta\, \theta_{i_1}\cdots\theta_{i_n}$$
$$= \frac{1}{n!}\,(\epsilon_{i_1\cdots i_n} d)\, \epsilon_{i_1\cdots i_n} = d, \tag{A.18}$$

using $\epsilon_{i_1\cdots i_n}\epsilon_{i_1\cdots i_n} = n!$.

Consider a change of variables from $\{\theta_i\}$ to $\{\theta'_i\}$ defined by

$$\theta_i = X_{ij}\theta'_j, \tag{A.19}$$

where X is a matrix of complex numbers. Then the top component transforms as

$$\epsilon_{i_1\cdots i_n}(X_{i_1 j_1}\cdots X_{i_n j_n}) = (\det X)\,\epsilon_{j_1\cdots j_n}, \tag{A.20}$$

so the coefficient of $\theta'_1\cdots\theta'_n$ becomes $d' = (\det X)\,d$. Hence

$$\int d^n\theta\, f(\{\theta_i\}) = d = (\det X)^{-1} d' = (\det X)^{-1}\int d^n\theta'\, f(\{\theta'_i\}). \tag{A.21}$$

Equivalently, the Berezin measure transforms as $d^n\theta = (\det X)^{-1} d^n\theta'$, i.e. the Jacobian appears with the inverse determinant, in contrast to ordinary commuting variables.

References

[1] C. Eröncel, *Axions and Axion-Like Particles*, (2023).

[2] M. Srednicki, "Quantum field theory," Cambridge University Press, 2007, ISBN 978-0-521-86449-7, 978-0-511-26720-8 doi:10.1017/CBO9780511813917

[3] M. Nakahara, "Geometry, topology and physics,"

[4] A. Altland and B. Simons, "Condensed Matter Field Theory," Cambridge University Press, 2023, ISBN 978-1-108-78124-4, 978-1-108-49460-1 doi:10.1017/9781108781244

[5] M. Fecko, "Differential geometry and Lie groups for physicists," Cambridge University Press, 2011, ISBN 978-0-521-18796-1, 978-0-521-84507-6, 978-0-511-24296-0