22

Instantons of QCD

In Chapter 5 we saw that solutions of the classical equations of motion, which are characterized by a topological number, play an important role in two-dimensional QFT. Derick's theorem (5.36) forbids scalar field soliton solutions in higher than two-dimensional space-time. However, for gauge fields one can bypass the theorem, and indeed, as we have seen in Chapter 21, there are solitons in the form of magnetic monopoles in four-dimensional gauge theories. The topic of this chapter will be solutions of the Yang–Mills theory defined on a Euclidean space-time which have finite action and are topological in their nature, the *instantons*. We will start with a description of the basic properties of one instanton solution including the topological charge that characterizes it. We then describe the construction of multi-instanton solutions and the moduli space of instantons including its dimension, complex nature, singularities and symmetries. When Wick rotated to Minkowski space-time the instanton describes a tunneling process between different vacua. We will elaborate on this phenomenon in the context of the four-dimensional YM theory. Various properties of QCD and hadron physics were thought to be related to instantons. In certain cases like confinement, the relation to instantons is still a mystery. One case where the role of instantons is clear is the $U(1_A)$ problem. This will be described in the last section of this chapter.

The one instanton solution was derived in [32]. The basic properties of instantons were worked out by many authors including [125], [57]. The instantons of $SU(N)$ gauge symmetry were derived in [218]. There are several review papers such as [65], [208] and [189] and books [182] and [188] that describe the basic instanton solutions.

22.1 The basic properties of the instanton

The action of the four-dimensional YM theory in Euclidean space-time can be rewritten in the following form,

$$\begin{aligned} S &= \frac{1}{2g^2}\int \mathrm{d}^4x\, \mathrm{Tr}\,[F^{\mu\nu}F_{\mu\nu}] = \frac{1}{4g^2}\int \mathrm{d}^4x F^{a\,\mu\nu}F^a_{\mu\nu} \\ S &= \frac{1}{2g^2}\int \mathrm{d}^4x\left[\pm\,\mathrm{Tr}\,[F^{\mu\nu\,*}F_{\mu\nu}] + \frac{1}{2}\,\mathrm{Tr}\,[(F_{\mu\nu}\mp{}^*F_{\mu\nu})(F^{\mu\nu}\mp^* F^{\mu\nu})]\right], \end{aligned} \quad (22.1)$$

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

where as usual $F_{\mu\nu} = F^a_{\mu\nu} T^a$[1] and ${}^*F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma}$ is the dual field strength, and g is the YM coupling. The first term in the second line is a topological invariant, or a topological charge[2] since,

$$Q = \frac{1}{16\pi^2} \int \mathrm{d}^4 x \, \mathrm{Tr}\,[F^{\mu\nu\,*}F_{\mu\nu}] = \int \mathrm{d}^4 x \partial_\mu K_\mu = \oint d\sigma_\mu K_\mu, \tag{22.2}$$

where,

$$K_\mu = \frac{1}{8\pi^2} \epsilon_{\mu\nu\rho\sigma} \, \mathrm{Tr}\left[A_\nu \left(\partial_\rho A_\sigma + \frac{3}{2} A_\rho A_\sigma\right)\right]. \tag{22.3}$$

In fact Q is the Pontryagin index or the winding number of maps from the sphere at space-time infinity to the $SU(2)$ group manifold which is also the three sphere, namely $S^3_s \to S^3_g$. This topological invariant is the homotopy $\pi_3(S^3)$ which is an integer $\pi_3(S^3) \in \mathcal{Z}$. This can be shown as follows. Since the self-dual field is asymptotically a pure gauge, namely on σ_μ $A_\mu = U\partial_\mu U^{-1}$ and $F_{\mu\nu} = 0$ hence,

$$\begin{aligned} Q &= \frac{1}{24\pi^2} \oint \mathrm{d}\sigma_\mu \epsilon_{\mu\nu\rho\sigma} \, \mathrm{Tr}\,[A_\nu A_\rho A_\sigma] \\ Q &= \frac{1}{24\pi^2} \oint \mathrm{d}\sigma_\mu \epsilon_{\mu\nu\rho\sigma} \, \mathrm{Tr}\,[(\partial_\nu U)U^{-1}(\partial_\rho U)U^{-1}(\partial_\sigma U)U^{-1}]. \end{aligned} \tag{22.4}$$

If we take for U the following ansatz that will be shown below to correspond to the one instanton solution,

$$U(x) = \hat{x}_\mu \sigma_\mu = x_0 + i x_i \sigma_i, \tag{22.5}$$

we found,

$$\begin{aligned} Q &= \oint \mathrm{d}\sigma_\mu K_\mu = -\frac{1}{24\pi^2} \int \mathrm{d}\sigma_\mu \left(\frac{-12 x_\mu}{|x^4|}\right) \\ &= \frac{1}{2\pi^2} \int \mathrm{d}\Omega x_\mu |x|^2 \frac{x_\mu}{|x|^4} = \frac{1}{2\pi^2} \int d\Omega = 1. \end{aligned} \tag{22.6}$$

Thus we have shown that Q is indeed the winding number that measures how many times we wind S^3_g when we integrate over S^3_s.

Let us now return to (22.1). Since it is a sum of a topological charge and a positive semi-definite quantity, it is clear that it is minimized when the latter vanishes namely,

$$F_{\mu\nu} = \pm{}^* F_{\mu\nu}. \tag{22.7}$$

The corresponding gauge fields A_μ (with a $+$ sign) will be referred to as *instanton* or *self-dual* gauge field and those with a $-$ sign as *anti instanton* or *anti self-dual*

[1] In this chapter we denote the $SU(N)$ adjoint indices with $a = 1, \ldots, N^2 - 1$ whereas in Chapter 19 we used A and not a.

[2] Recall in analogy the topological charge defined in two-dimensional scalar field theories (5.3).

gauge field. It is straightforward to show that the self-duality condition implies the equation of motion,

$$D_\mu F_{\mu\nu} = 0. \tag{22.8}$$

A solution of the equation of motion is not necessary self-dual but it can be shown that the non-self-dual configurations are saddle points and not local minima of the action.

Comparing the expression of the action (22.1) and the topological charge (22.2) it is clear that a (anti) self-dual configuration that carries an instanton number $(Q = -k), Q = k$ has an action of,

$$S = \frac{8\pi^2}{g^2}|k|. \tag{22.9}$$

One can add the topological charge as an additional term to the action. To be more precise one adds a θ term,

$$S_\theta = i\frac{\theta}{16\pi^2}\int \mathrm{d}^4x\, \mathrm{Tr}\,[F^{\mu\nu\,*}F_{\mu\nu}] = i\theta k. \tag{22.10}$$

Whereas the ordinary YM action is the same for the instanton and anti-instanton, the θ term obviously distinguishes between them by assigning opposite charges to them. We will further discuss the *theta* term in Section 22.5. For the self-dual solution up to a constant the action is equal to the topological charge which by definition does not depend on the metric. This exhibits the topological nature of the instanton. Another indication of this nature is the fact that it has vanishing energy-momentum tensor as follows from,

$$T_{\mu\nu} = -\frac{2}{g^2}\,\mathrm{Tr}\left[F_{\mu\rho}F_{\nu\rho} - \frac{1}{4}\delta_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right] = 0. \tag{22.11}$$

This clearly implies that instantons do not curve the space-time they reside in.

The one instanton solution for the $SU(2)$ can be constructed from $U(x)$ given in (22.5) via,

$$A^a_\mu(x) = U^{-1}\partial_\mu U\frac{\rho^2}{(x-X)^2+\rho^2}, \tag{22.12}$$

which yields the explicit form,

$$A^a_\mu(x) = 2\frac{\eta^a_{\mu\nu}(x-X)^\nu}{(x-X)^2+\rho^2}, \tag{22.13}$$

where $\eta^a_{\mu\nu}$ is the 't Hooft antisymmetric symbol defined by,

$$\begin{aligned} \eta^a_{\mu\nu} &= \epsilon^a_{\mu\nu} \quad \mu,\nu = 1,2,3 \quad \eta^a_{\mu 4} = -\eta^a_{4\mu} = \delta^a_\mu \\ \bar\eta^a_{\mu\nu} &= \epsilon^a_{\mu\nu} \quad \mu,\nu = 1,2,3 \quad \bar\eta^a_{\mu 4} = -\bar\eta^a_{4\mu} = -\delta^a_\mu. \end{aligned} \tag{22.14}$$

It is easy to check that $\eta^a_{\mu\nu}$ and $\bar{\eta}^a_{\mu\nu}$ are self-dual and anti self-dual respectively. The corresponding field strength takes the form,

$$F^a_{\mu\nu} = -4\eta^a_{\mu\nu}\frac{\rho^2}{[(x-X)^2+\rho^2]^2}. \tag{22.15}$$

Obviously since $\eta^a_{\mu\nu}$ is self-dual so is $F^a_{\mu\nu}$.

The one instanton solution is characterized by eight parameters, four correspond to the center of the instanton X_μ, one to the size of the instanton ρ and three to three global $SU(2)$ gauge transformations. Recall that fixing a guage we fix only the local gauge transformations. The space of parameters of the k instanton solutions will be further addressed in Section 22.3 where it will be shown that in general for $SU(N)$ the dimension of the moduli space is $4kN$.

The instanton solution (22.13) falls off asymptotically as $\frac{1}{x}$ and hence it contributes a finite amount to the integral of the topological charge (winding number) which is of the form $\int A^3x^3d\Omega$. The field strength falls off as $1/x^4$ such that indeed the corresponding action is finite. However due to the $\frac{1}{x}$ asymptotic behavior it is difficult to form square integrable expressions that contain it. For that purpose one can use the following singular gauge transformation,

$$U = \frac{\sigma^\dagger_\mu(x-X)^\mu}{|x-X|}, \tag{22.16}$$

which renders the instanton to have a $\frac{1}{x^3}$ fall off as can be seen from,

$$A^a_\mu = \frac{1}{g}\frac{2\rho^2(x-X)^\nu\bar{\eta}^a_{\nu\mu}}{(x-X)^2[(x-X)^2+\rho^2]}. \tag{22.17}$$

The singular instanton is obviously singular at the location of the instanton $x_\mu = X_\mu$. This singularity is not physical and can be removed by a gauge transformation or by puncturing the Euclidean space with the singular point being removed.

Instantons of $SU(N)$ gauge theory can be constructed by embedding $SU(2)$ instantons in $SU(N)$ for instance,

$$A^{SU(N)}_\mu = \begin{pmatrix} 0 & 0 \\ 0 & A^{SU(2)}_\mu \end{pmatrix}, \tag{22.18}$$

where the instanton is the 2×2 matrix on the lower right. The most general $SU(N)$ one instanton configuration can be derived from (22.18) by the following transformation,

$$A^{SU(N)}_\mu = U^\dagger \begin{pmatrix} 0 & 0 \\ 0 & A^{SU(2)}_\mu \end{pmatrix} U \quad U \in \frac{SU(N)}{SU(N-2)\times U(1)}. \tag{22.19}$$

Operating with $U \in SU(N-2)\times U(1)$ obviously leaves the basic configuration invariant and thus only transformations with elements of the coset are relevant. This is in accordnace with the fact that for a k instanton solution the stability group is $S(U(N-2k)\times U1))$ as will be shown in Section 22.3. The dimension

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

of the coset is $N^2 - 1 - ((N-2)^2 = 4N - 5$. Together with the 5 parameters of the location and the size we have $4N$ collective coordinates. Indeed in Section 22.3 we will see that in general the dimension of the moduli space is $4Nk$ for instanton number equals k solution. To demonstrate this counting consider the case of $SU(3)$ for which the generators are the Gell–Mann matrices $\{\lambda^a\}$, $a = 1, \ldots, 8$. The first three generators λ_a, $a = 1, 2, 3$ form the $SU(2)$ $k = 1$ instanton. $\lambda_4, \ldots, \lambda_7$ form two doublets under this $SU(2)$ so they can generate new solutions while λ_8 commutes with the $SU(2)$ and hence leaves the basic $SU(2)$ solution invariant.

One can express the instanton solution in terms of quaternionic notation. This will turn out to be convenient for the ADHM construction of multi-instanton solutions (see Section 22.2). The idea is to make use of the representations of the covering group $SU(2)_{\mathrm{L}} \times SU(2)_{\mathrm{R}}$ of the Lorentz group of four-dimensional Euclidean space-time rather than the $SO(4)$ Lorentz group itself. In particular we represent any four vector of $SO(4)$ as a $(2, 2)$ representation of $SU(2)_{\mathrm{L}} \times SU(2)_{\mathrm{R}}$, for instance the four vector x_μ is denoted by $x_{\alpha\dot{\alpha}}$ or $x^{\dot{\alpha}\alpha}$ defined as follows,

$$x_{\alpha\dot{\alpha}} = x_\mu \sigma^\mu_{\alpha\dot{\alpha}} \quad x^{\dot{\alpha}\alpha} = x_\mu \bar{\sigma}^{\dot{\alpha}\alpha}_\mu, \tag{22.20}$$

where $\sigma^\mu_{\alpha\dot{\alpha}}$ is a 2×2 matrix of $(i\vec{\sigma}, 1)$ and $\bar{\sigma}^\mu_{\dot{\alpha}\alpha} = (\sigma^\mu_{\alpha\dot{\alpha}})^\dagger$. In terms of the quaternionic notation the one instanton solution (22.13) for $SU(2)$ gauge theory is given by,

$$A_\mu = \frac{1}{g} \frac{2(x-X)^\nu \Sigma_{\nu\mu}}{(x-X)^2 + \rho^2}, \tag{22.21}$$

where $\Sigma_{\mu\nu}$, which were introduced in Section 17.1, are the part of the Lorentz generators that do not act on the space-time coordinates but only on the internal degrees of freedom. Here using the $SU(2)_{\mathrm{L}} \times SU(2)_{\mathrm{R}}$ notation we define them as follows,

$$\Sigma_{\mu\nu} = \frac{1}{4}(\sigma_\mu \bar{\sigma}_\nu - \bar{\sigma}_\nu \sigma_\mu) \quad \bar{\Sigma}_{\mu\nu} = \frac{1}{4}(\bar{\sigma}_\mu \sigma_\nu - \sigma_\nu \bar{\sigma}_\mu). \tag{22.22}$$

The self-duality property of the instanton configuration follows trivially from the fact that $\Sigma_{\mu\nu}$ is self-dual, namely,

$$\Sigma_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\rho}\Sigma^{\lambda\rho} \quad \bar{\Sigma}_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\lambda\rho}\bar{\Sigma}^{\lambda\rho}. \tag{22.23}$$

The corresponding field strength reads,

$$F_{\mu\nu} = \frac{1}{g} \frac{4\rho\Sigma_{\mu\nu}}{((x-X)^2 + \rho^2)^2}. \tag{22.24}$$

As was discussed above, the instanton solution (22.21) falls off asymptotically as $\frac{1}{x}$ and hence it is difficult to form square integrable expressions that contain it. The solution in the singular gauge now reads,

$$A_\mu = \frac{1}{g} \frac{2\rho^2 (x-X)^\mu \bar{\Sigma}_{\nu\mu}}{((x-X)^2 (x-X)^2 + \rho^2)}. \tag{22.25}$$

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

22.2 The ADHM construction of instantons

The vacua of the YM theory is given by pure gauge configurations (which can be written in terms of a double index notation,

$$A^{\mu}_{ij} = \frac{1}{g} {U^{\dagger}}^{l}_{i} \partial^{\mu} U_{lj}, \tag{22.26}$$

where $i, j, l = 1, \ldots, N$. It is straightforward to check that this gauge field obeys the self-duality condition. The idea of the ADHM construction[3] is to generalize this configuration also to the k instanton case by taking now the matrices U to be of the form $U_{\mathcal{I}i}$ where $\mathcal{I} = 1, \ldots, N + 2k$ with the orthonormality condition,

$${U^{\dagger}}^{\mathcal{I}}_{i} U_{\mathcal{I}j} = \delta_{ij}. \tag{22.27}$$

The U matrices are the basis vectors of a null space,

$${\Delta^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} U_{\mathcal{I}i} = 0 = {U^{\dagger}}^{\mathcal{I}}_{i} \Delta_{\mathcal{I}I\dot{\alpha}}, \tag{22.28}$$

where $I = 1, \ldots, k$ and $\Delta_{\mathcal{I}I\dot{\alpha}}$ is a $(N + 2k) \times 2k$ complex valued matrix which is taken to be linear in the space-time coordinate x_{μ}, namely takes the form,

$$\Delta_{\mathcal{I}I\dot{\alpha}} = a_{\mathcal{I}I\dot{\alpha}} + b^{\alpha}_{\mathcal{I}I} x_{\alpha\dot{\alpha}} \quad {\Delta^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} = {a^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} + x^{\dot{\alpha}\alpha} b^{\mathcal{I}}_{I\alpha}, \tag{22.29}$$

and with ${\Delta^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} \equiv (\Delta_{\mathcal{I}I\dot{\alpha}})^{*}$.

The ADHM k instanton solution of the form (22.26) is self-dual if one further requires that $\Delta_{\mathcal{I}I\dot{\alpha}}$ obeys the following condition,

$${\Delta^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} \Delta_{\mathcal{I}J\dot{\beta}} = \delta^{\dot{\alpha}}_{\dot{\beta}} (f^{-1})_{IJ}, \tag{22.30}$$

where f is an arbitrary x-dependent $k \times k$ dimensional Hermitian matrix. Since $f_{IJ}(x)$ is arbitrary there are three "ADHM constraints" on a, $a^{\dagger}$, b and $b^{\dagger}$,

$$\begin{aligned}
{a^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} a_{\mathcal{I}J\dot{\beta}} &= \left(\frac{1}{2} a^{\dagger} a\right)_{IJ} \delta^{\dot{\alpha}}_{\dot{\beta}} \\
{a^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{I} b_{\mathcal{I}I\dot{\beta}} &= {b^{\dagger}}^{\beta\mathcal{I}}_{I} a^{\dot{\alpha}}_{\mathcal{I}} J \\
{b^{\dagger}}^{\mathcal{I}}_{\alpha I} b^{\beta}_{\mathcal{I}I} &= \left(\frac{1}{2} a^{\dagger} a\right)_{IJ} \delta^{\beta}_{\alpha}.
\end{aligned} \tag{22.31}$$

It is straightforward to realize that the ADHM construction is invariant under,

$$\Delta \to \mathcal{A} \Delta \mathcal{B}^{-1} \quad U \to \mathcal{A} U \quad f \to \mathcal{B} f \mathcal{B}^{\dagger}, \tag{22.32}$$

where $\mathcal{A} \in U(N + 2k)$ and $\mathcal{B} \in GL(k, \mathcal{C})$. Thus by construction there is a redundancy in a and b. One can choose a simple canonical form for b and a as follows,

$$\begin{aligned}
b^{\beta}_{\mathcal{I}J} &= \begin{pmatrix} 0 \\ \delta^{\beta}_{\alpha} \delta_{IJ} \end{pmatrix} \quad {b^{\dagger}}^{\mathcal{I}}_{\beta J} = (0\ , \delta^{\alpha}_{\beta} \delta_{ji}) \\
a_{\mathcal{I}J\dot{\alpha}} &= \begin{pmatrix} \hat{\alpha}_{iJ\dot{\alpha}} \\ (a'_{\alpha\dot{\alpha}})_{IJ} \end{pmatrix} \quad {a^{\dagger}}^{\dot{\alpha}\mathcal{I}}_{J} = (\hat{a}^{\dot{\alpha}}_{Ji},\ ({a'^{\dagger}}^{\dot{\alpha}al})_{IJ}).
\end{aligned} \tag{22.33}$$

where $\mathcal{I} = i + I\alpha$.

[3] Multi-instanton solutions were presented in [220], [132] and other papers. Our discussion of the construction of multi-instanton is based on the paper of ADHM [20]. This approach was further discussed in [70]. We follow the description of the construction given in [81].

In this parametrization the third ADHM constraint is automatically obeyed while the other two take the form,

$$\vec{\sigma}^{\dot{\alpha}}_{\dot{\beta}}(a^{\dagger\dot{\beta}} a_{\dot{\alpha}}) = \vec{\sigma}^{\dot{\alpha}}_{\dot{\beta}}((\hat{a}^{\dagger})^{\dot{\beta}} \hat{a}_{\dot{\alpha}} + a'^{\dagger\dot{\beta}} a'_{\dot{\alpha}}) = 0, \tag{22.34}$$

where we made use of the fact that a' must be Hermitian. In this canonical parametrization the matrix f reads,

$$f = 2((\hat{a}^{\dagger})^{\dot{\alpha}} \hat{a}_{\dot{\alpha}} + (a'_\mu + x_\mu 1_{[k]\times[k]})^2)^{-1}. \tag{22.35}$$

The field strength $F_{\mu\nu}$ that corresponds to the ADHM k instanton configuration (22.26) can be written in the following form,

$$\begin{aligned}
F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + g[A_\mu, A_\nu] = \frac{1}{g}\partial_{[\mu}(U^\dagger \partial_{\nu]} U) + \frac{1}{g}(U^\dagger \partial_{[\mu} U)(U^\dagger \partial_{\nu]} U) \\
&= \frac{1}{g}\partial_{[\mu} U^\dagger (1 - U^\dagger U)\partial_{\nu]} U) = \frac{1}{g}\partial_{[\mu} U^\dagger \Delta f \Delta^\dagger \partial_{\nu]} U \\
&= \frac{1}{g} U^\dagger \partial_{[\mu} \Delta f \partial_{\nu]} \Delta^\dagger U = \frac{1}{g} U^\dagger b \sigma_{[\mu} \bar{\sigma}_{\nu]} f b^+ U = 4\frac{1}{g} U^\dagger b \sigma_{\mu\nu} f b^+ U,
\end{aligned} \tag{22.36}$$

where we have made use of,

$$\mathcal{P}^{\mathcal{J}}_{\mathcal{I}} \equiv U_{\mathcal{I}i} U^{\dagger\mathcal{J}} = \delta^{\mathcal{J}}_{\mathcal{I}} - \Delta_{\mathcal{I}I\dot{\alpha}} f_{IJ} \Delta^{\dagger\dot{\alpha}\mathcal{J}}_{J}. \tag{22.37}$$

To get an explicit expression for A_μ we make use of the decomposition

$$U_{\mathcal{I}i} = \begin{pmatrix} \hat{U}_{ij} \\ (U'_\alpha)_{Ii} \end{pmatrix} \quad \Delta_{\mathcal{I}J\dot{\alpha}} = \begin{pmatrix} \hat{a}_{iJ\dot{\alpha}} \\ (\Delta'_{\alpha\dot{\alpha}})_{IJ} \end{pmatrix}. \tag{22.38}$$

From the completeness condition (22.37) $\hat{U}$ can take the form,

$$U = \sqrt{(i_{[N]\times[N]} - \hat{a}_{\dot{\alpha}} f (\hat{a})^{\dagger\dot{\alpha}})} \quad U' = \Delta'_{\dot{\alpha}} f (\hat{a})^{\dagger\dot{\alpha}} (\hat{u})^{\dagger -1}. \tag{22.39}$$

We next show that for the particular case of $k = 1$ the ADHM solution (22.26) is identical to (22.25). For $k = 1$ we have to drop the indices I, J. One can verify that in that case the parameters a'_μ can be identified with the center of the instanton X_μ. From the ADHM constraint (22.34) we get that,

$$(\hat{a}^\dagger)^{\dot{\beta}} \hat{a}_{\dot{\alpha}} = \rho^2 \delta^{\dot{\beta}}_{\dot{\alpha}}, \tag{22.40}$$

and,

$$f = \frac{1}{(x_\mu - X_\mu)^2 + \rho^2}, \tag{22.41}$$

where ρ will naturally be the size of the instanton.

From the relation (22.39) we deduce that,

$$\begin{aligned}
\hat{U} &= 1_{[N]\times[N]} + \frac{1}{\rho^2}\left(\sqrt{\frac{(x-X)^2}{(x-X)^2 + \rho^2}} - 1\right)(\hat{a}^\dagger)^{\dot{\beta}} \hat{a}_{\dot{\alpha}}, \\
U' &= -\frac{(x-X)_{\alpha\dot{\alpha}} (\hat{a}^\dagger)^{\dot{\alpha}}}{|x-X|\sqrt{(x-X)^2 + \rho^2}}.
\end{aligned} \tag{22.42}$$

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

Plugging these expressions into U we finally get the singular form of the one instanton solution (22.25).

In general for $k \neq 1$ one can show that the gauge configuration given in (22.36) indeed carries an instanton of charge k. To derive this result one makes use of the following relation,

$$g^2 \, \mathrm{Tr}\,[F_{\mu\nu}^2] = \partial_\mu \partial^\mu tr_k[\log f]. \tag{22.43}$$

This relation can be proven by expanding the two sides of the equations using the explicit expression for $F_{\mu\nu}$ (22.36). Upon integrating (22.43) over the whole Euclidean space-time divided by $\frac{1}{16\pi^2}$ and making use of the fact that asymptotically $f(x) \to \frac{1}{x^2}$, we find that indeed it is equal to the instanton charge k.

22.3 On the moduli space of instantons

The moduli space of instantons[4] $\mathcal{M}$ is the space of inequivalent self-dual Yang–Mills configurations. The notion of moduli space of solutions is an important tool in general and in particular for instantons. We have encountered it in the context of magnetic monopoles in (21.10). We will elaborate in this section about the basic properties of the moduli space of instantons such as its dimension, complex structure, metric, symmetries, and singularities.

Consider a small fluctuation $\delta A_\mu(x)$ around an instanton solution $A_\mu(x)$ which is also a self-dual solution of the YM equation, namely, it obeys to linear in $\delta A_\mu(x)$,

$$D_\mu \delta A_\nu - D_\nu \delta A_\mu = \epsilon_{\mu\nu\rho\sigma} D^\rho \delta A^\sigma, \tag{22.44}$$

where D_μ is the covariant derivative in the instanton background. In terms of the quateronic notation this equation reads,

$$\vec{\sigma}^{\dot{\alpha}}_{\dot{\beta}} \not{D}^{\dagger\,\dot{\beta}\alpha} \delta A_{\alpha\dot{\alpha}} = 0, \tag{22.45}$$

where $\not{D} = \sigma^\mu D_\mu$. Next we would like to guarantee that the fluctuation is not a local gauge transformation. This can be done by requiring that the fluctuation be orthogonal to any gauge transformation, namely,

$$\int \mathrm{d}^3x \, \mathrm{Tr}[\delta A_\mu D^\mu \Lambda] = 0 \;\to D^\mu \delta A_\mu = 0, \tag{22.46}$$

where we have made use of an integration by parts to derive the last expression. In the quaternionic notation this condition takes the form of $\not{D}^{\dagger\,\dot{\alpha}\alpha} \delta A_{\alpha\dot{\alpha}} = 0$ which combined with (22.45) is given by,

$$\not{D}^{\dagger\,\dot{\alpha}\alpha} \delta A_{\alpha\dot{\beta}} = 0. \tag{22.47}$$

[4] The properties of the moduli space of instantons were discussed by many authors. In particular [155], [141] and [80]. The review about the moduli space that we are using is [81].

The fluctuation that obeys this equation is referred to as a *zero mode* since it is a physical fluctuation that does not change the action. Note that this is exactly the equation of motion of a Weyl spinor in the background of the instanton $A_\mu(x)$. The zero modes defined by (22.47) are the collective coordinates of the instantons.

Since the YM instantons are characterized by the topological charge defined in (22.2) so is the corresponding moduli space. We thus discuss the moduli space of instantons of charge k which we denote by $\mathcal{M}_k$.

It can be proven that the moduli space of instantons is a manifold. In fact we will see below that $\mathcal{M}_k$ has some conical singular points associated with zero size instantons. The coordinates on the moduli space are the collective coordinates that were just shown to be equivalent to the zero model (22.47). We denote by X_n the collective coordinates where $n = 1, \ldots, \dim \mathcal{M}_k$. A trivial set of coordinates are the space-time coordinates of the center of the instanton X_μ accordingly the moduli space is a product of the form,

$$\mathcal{M}_k = \mathcal{R}^4 \times \hat{\mathcal{M}}_k. \tag{22.48}$$

The collective coordinates X_μ follow from the fact that the instanton solution breaks the symmetry of the action under space-time translations. There are other collective coordinates that associate with symmetries of the theory that the instanton configuration breaks. However, not all symmetries yield non equivalent collective coordinates and not all the coordinates associate with broken symmetries.

From the ADHM construction it follows that the moduli space is identified with the variable $a_{\dot\alpha}$ subject to the ADHM constraints (22.31) quotiented by the residual $U(K)$ symmetry transformation (22.32) with,

$$\mathcal{A} = \begin{pmatrix} 1_{[N]\times[N]} & 0 \\ 0 & \mathcal{C}1_{[2]\times[2]} \end{pmatrix}, \quad \mathcal{B} = \mathcal{C} \quad \mathcal{C} \in U(k), \tag{22.49}$$

which preserve the canonical form of b (22.33) and transform a as follows,

$$\hat{a}_{iI\dot\alpha} \to \hat{a}_{\dot\alpha}\mathcal{C} \quad a'_\mu \to \mathcal{C}^\dagger a'_\mu \mathcal{C}. \tag{22.50}$$

Thus the dimension of the moduli space is,

$$\dim \mathcal{M}_k = 4k(N+K) - 4\dim\ U(k) = 4k(N+k) - 4k^2 = 4NK. \tag{22.51}$$

This result can be derived also by using an index theorem that counts the zero modes at a point in the moduli space. In fact as we will see shortly the space $\mathcal{M}_k$ is a hyper-Kahler quotient of the flat space $\mathcal{R}^{4k(k+N)}$ by the $U(k)$ group. The one instanton solution of $SU(2)$ is indeed characterized by the four coordinates of its center, its size and three global $SU(2)$ gauge transformations.

The moduli space is a complex manifold. A complex manifold is an even-dimensional manifold that admits a complex structure $\mathbf{I}$ a linear map of the

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

tangent space to itself such that $\mathbf{I}^2 = 0$. There are always local holomorphic coordinates $(Z^i, \bar{Z}^i), i = 1, \ldots, n$ (see Section 1.1) for which,

$$\mathbf{I} = \begin{pmatrix} i\delta^{\bar{i}}_{\bar{j}} & 0 \\ 0 & -i\delta^i_j \end{pmatrix} \quad g = g_{i\bar{j}} dZ^i d\bar{Z}^{\bar{j}} \quad w = ig_{i\bar{j}} dZ^i \wedge d\bar{Z}^{\bar{j}}. \tag{22.52}$$

where g is an Hermitian metric and w is referred to as the fundamental 2 form. In the case that the fundamental 2 form is closed namely $\mathrm{d}w = 0$ it is called the *Kähler form* and the associated manifold is a Kähler manifold. The latter is also characterized by the fact that the complex structure is covariantly constant and the Kähler metric can be derived from a *Kähler potential,*

$$\nabla_\mu \mathbf{I} = 0 \quad g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K. \tag{22.53}$$

The moduli space of instanton is not only a Kähler manifold but in fact a *hyper Kähler manifold* which means that it admits three linearly independent complex structures, $\mathbf{I}^{(c)}, c = 1, 2, 3$ that satisfies the algebra

$$\mathbf{I}^{(c)} \mathbf{I}^{(d)} = -\delta^{cd} + \epsilon^{cde} \mathbf{I}^{(e)}. \tag{22.54}$$

The four-dimensional Euclidean space $\mathcal{R}^4$ is hyper Kähler and the three complex structures are,

$$\mathbf{I}^{(c)}_{\mu\nu} = -\eta^c_{\mu\nu} \quad (\vec{\mathbf{I}} \cdot x)_{\alpha\dot{\alpha}} = ix_{\alpha\dot{\beta}} \vec{\sigma}^{\dot{\beta}}_{\dot{\alpha}}, \tag{22.55}$$

where $\eta^c_{\mu\nu}$ is the 't Hooft η symbol defined in (22.14) and the expression in the left-hand side is the quaterionic formulation. Now recall that by the definition of the zero modes (22.47), if $\delta_n A_{\alpha\dot{\alpha}}$ is a zero mode so is also $\delta_n A_{\alpha\dot{\alpha}} C^{\dot{\beta}}_{\dot{\alpha}}$ for any constant matrix $\mathbf{C}$ and in particular also to $\vec{\sigma}$ and hence if $\delta_n A_{\alpha\dot{\alpha}}$ so is also $(\vec{\mathbf{I}} \cdot \delta_n A)_{\alpha\dot{\alpha}} = i\delta_n A_{\alpha\dot{\alpha}} \vec{\sigma}^{\dot{\beta}}_{\dot{\alpha}}$. Since the zero modes form a complete set there must exist $\vec{\mathbf{I}}^n_m$ such that,

$$(\vec{\mathbf{I}} \cdot \delta_m A)_{\alpha\dot{\alpha}} = \delta_n A_{\alpha\dot{\alpha}} \vec{\mathbf{I}}^n_m, \tag{22.56}$$

from which it implies that $\vec{\mathbf{I}}^n_m$ satisfies the algebra (22.54).

The Kähler potential which is common to the three complex structures of the moduli space of instantons takes the form,

$$K = -\frac{g^2}{4} \int \mathrm{d}^4 x x^2 \,\mathrm{Tr} F_{\mu\nu} F^{\mu\nu}. \tag{22.57}$$

Using the form of $\mathbf{I}^{(c)}$ on $\mathcal{R}^4$ given in (22.52) for instance $I^{(3)}$ associated with the complex coordinates on $\mathcal{R}^4$ $ix^3 + x^4$ and $ix^1 - x^2$, we find that,

$$(\mathbf{I}^{(c)} \cdot \partial_{Z^i} A)_{\alpha\dot{\alpha}} = i\partial_{Z^i} A_{\alpha\dot{\alpha}}, \tag{22.58}$$

for instance for $\mathbf{I}^{(3)}$ we get $\partial_{Z^i} A_{\alpha 2} = \partial_{\bar{Z}^{\bar{i}}} A_{\alpha 1} = 0$. Furthermore the derivative with the respect to the holomorphic and anti-holomorphic coordinates of the gauge fields obey the equations of zero modes namely,

$$\delta_i A_\mu \equiv \partial_{Z^i} A_\mu \quad \bar{\delta}_{\bar{i}} A_\mu \equiv \partial_{\bar{Z}^{\bar{i}}} A_\mu. \tag{22.59}$$

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

It further follows that $\partial_{\bar{Z}^{\bar{j}}}\partial_{Z^i} A_\mu = 0$ and hence,

$$\partial_{\bar{Z}^{\bar{j}}}\partial_{Z^i}\,\mathrm{Tr}\,[F_{\mu\nu}^2] = \partial_\mu\partial^\mu\,\mathrm{Tr}\,[\delta_i A_\mu \bar{\delta}_{\bar{j}} A_\mu - 2\partial_\mu\partial_\nu]\,\mathrm{Tr}\,\delta_i A_\mu \bar{\delta}_{\bar{j}} A_\nu. \tag{22.60}$$

Upon integrating by parts twice we find the metric on the moduli space,

$$\partial_{\bar{Z}^{\bar{j}}}\partial_{Z^i} K = -2g^2 \int \mathrm{d}^4x\,\mathrm{Tr}\,\delta_i A_\mu \bar{\delta}_{\bar{j}} A_\mu = g_{i\bar{j}}. \tag{22.61}$$

Next let us now discuss the symmetries of the moduli space, in particular the realization of symmetries of the gauge theory which are broken by the instanton configuration. We start with the four-dimensional conformal group (see Section 17). In the quaternionic formulation the basic variable of the ADHM construction Δ is transformed as follows,

$$x \to x' = (Ax+B)(cX+D)^{-1} \qquad \det\begin{pmatrix} AB \\ CD \end{pmatrix} = 1$$
$$\Delta(x;a,b) \to \Delta(x';a,b) = \Delta(x;aD+bB,aC+bA)(Cx+D)^{-1} \tag{22.62}$$

In fact the term $(Cx+D)^{-1}$ in the right-hand side of the last equation is irrelevant since the gauge field depends on U and $U^\dagger$ defined in (22.26) is redundant.

We can now use transformations of (22.32) that keeps the canonical structure of b (22.33). Upon applying this transformation a goes into,

$$a \to \mathcal{A}(aD+bB)\mathcal{B}^{-1}. \tag{22.63}$$

A particular example of transformations which belong to the conformal group are the translations. For this case,

$$\Delta(x;a,b) \to \Delta(x;a+b\epsilon,b), \tag{22.64}$$

from which it follows that,

$$a'_\mu \to a'_\mu + \epsilon_\mu 1_{[k]\times[k]} \qquad \hat{a}_{\dot{\alpha}} \to \hat{a}_{\dot{\alpha}}. \tag{22.65}$$

It is thus clear that indeed the components a'_μ are proportional to the coordinates of the center of the instanton,

$$\mathrm{tr}_k a'_\mu = kX_\mu. \tag{22.66}$$

Global gauge transformations act non trivially on the ADHM variables if $N \leq 2k$, while if $N \geq 2k$ there are transformations that leave the instantons fixed. This is the stability group of the instanton. One can embed the k instanton solution in an $SU(2k)$ subgroup of $SU(N)$ and show that the stability group is $S(U(N-2k)\times U(1))$.

The moduli space $\hat{\mathcal{M}}_k$ is in fact not a smooth manifold due to certain singularities. However, these singularities do not signal any pathology of the moduli space and integrals over the moduli space are well defined. It can be shown that $\hat{\mathcal{M}}_k$ is a cone. For the moduli space of single instanton $k=1$ the apex of the cone is the point $\rho=0$ where the instanton has a zero size. This structure can be generalized also to the $k\neq 1$ instantons.

Topological characteristics of the moduli space can be described by a topological field theory where the observables of the theory are the topological invariants. This is beyond the scope of this book and we refer the interested reader to the list of references for this chapter.

22.4 Instantons and tunneling between the vacua of the YM theory

The vacua of the YM theory in Minkowsi space-time are defined to be the gauge configurations for which the energy vanishes. Using the temporal gauge $A_0 = 0$, the Hamiltonian of the theory is

$$H = \frac{1}{2g^2}\int \mathrm{d}^3x\, \mathrm{Tr}\,[E^2 + B^2]. \tag{22.67}$$

Thus a classical vacuum has a vanishing field strength,

$$F_{\mu\nu} = 0 \quad \rightarrow A_i(x,t) = iU(x,t)\partial_i U(x,t)^\dagger. \tag{22.68}$$

Thus the vacuum gauge configuration is that of a pure gauge. Prior to a discussion of how to tunnel between two vacuum states, we have to classify and enumerate the vacua namely following (22.68) the group elements $U(x)$. This translates to the equivalence classes of maps from S^3 to the $SU(N)$ group manifold. This is done by the topological charge or winding number or Pontryagin number defined in (22.2). Since this step is very essential in the discussion of the tunneling let us clarify this point. Let us analyze the tunneling between a vacuum state $A_i(x,t_1)$ at $t = t_1$ into another vacuum state $A_i(x,t_2)$ at $t = t_2$. On top of fixing $A_0(x,t) = 0$ we can use the residual gauge symmetry to set $A_i(x,t_1) = 0$. Next we consider a path in the space of gauge configurations that connects the two vacua points and has a finite energy H (in Minkowski space-time). Finite energy implies that for large $|\vec{x}| \to \infty$ it has to be a pure gauge,

$$A_\mu \rightarrow_{|x|\to\infty} U\partial_\mu U^\dagger. \tag{22.69}$$

Since $A_0 = 0$, $U(x,t) = U(x)$ is time independent. Because $U(x,t_1) = 1$ we obtain that for all t and $|x| \to \infty$, $U = 1$ and hence also $A_i(x,t_2) \longrightarrow_{|x|\to\infty} 0$. The fact that asymptotically in $|\vec{x}|$ all $A_i = 0$ allows us to compactify the spacelike hypersurface at fixed t into S^3. The following Fig. 22.1 describes the situation.

On the boundary of hyper-cylinder the gauge fields A_i vanish apart from on the hyper-disk at $t = t_2$ where A_i is a pure gauge. Consider now the topological charge which we have seen (22.2) is in fact a surface term. Since on the boundary $F_{\mu\nu} =$ the contribution to the surface integral takes the form,

$$\begin{aligned} Q &= \frac{1}{24\pi^2}\epsilon^{\mu\nu\rho\sigma}\oint \mathrm{d}\sigma_\mu\, \mathrm{Tr}\,[A_\nu A_\rho A_\sigma] \\ &= \frac{1}{24\pi^2}\epsilon^{0ijk}\int \mathrm{d}^3xr[U\partial_i U^\dagger U\partial_j U^\dagger U\partial_k U^\dagger], \end{aligned} \tag{22.70}$$

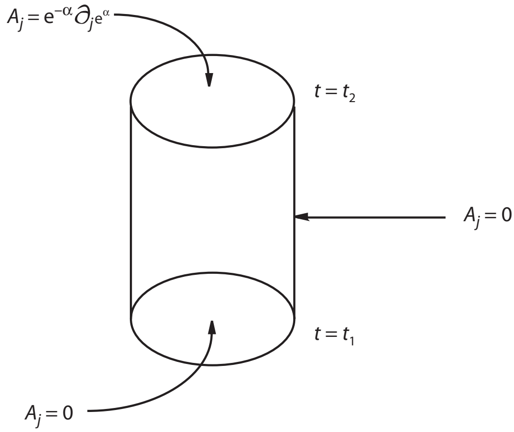

Fig. 22.1. Compactification of the space coordinates on S^3 at fixed t.

where in the last expression the integration is over the three sphere at $t = t_2$ since at all other parts of the surface of the hyper-cylinder $A_i = 0$. Thus the configurations in Minkowski space-time that connect a vacuum state at $t = t_1$ to one at $t = t_2$ are classified by the winding number of the maps of S^3(space) $\to S^3$(group) just as the maps of instanton in Euclidean space-time.[5] In the latter case S^3(space) is the boundary of R^4 whereas in the former it is a compactification of R^3 at $t = t_2$.

It is easy to realize that there is no way to interpolate a vacuum at $t = t_1$ of zero winding number with a one at $t = t_2$ with non vanishing winding number with a configuration of zero energy. The latter corresponds to a pure gauge configuration which has $F_{\mu\nu} = 0$ everywhere and hence also vanishing topological number. Thus the energy of the tunneling configuration as a function of time should look as in Fig. 22.2.

To identify the configuration that has the largest tunneling rate we consider a family of gauge configurations characterized by the collective coordinates associated with a coordinate transformation from t to $\lambda(t)$ such that,

$$A_i^{(\lambda)}(x,t) = A_i(x, \lambda(t)), \tag{22.71}$$

with the requirement that $\lambda(t_1) = t_1$ and $\lambda(t_2) = t_2$. Next we compute the electric and magnetic fields,

$$\begin{aligned} E_i &= F_{i0} = -\partial_0 A^{(\lambda)}(x,t) = \frac{\partial A_i}{\partial \lambda}(x, \lambda(t))\dot{\lambda} \\ B_i &= \frac{1}{2}\epsilon_{ijk}F_{jk} = \frac{1}{2}\epsilon_{ijk}(\partial_j A_k(x,\lambda(t)) + A_j(x,\lambda(t)(A_k(x,\lambda(t)) - (j \leftrightarrow k) \end{aligned} \tag{22.72}$$

[5] The role of instantons in tunneling between different vacua was proposed in [131]. It was also discussed in [26], [40] and [44]. This topic is reviewed in [185] and in [209]. We follow the latter.

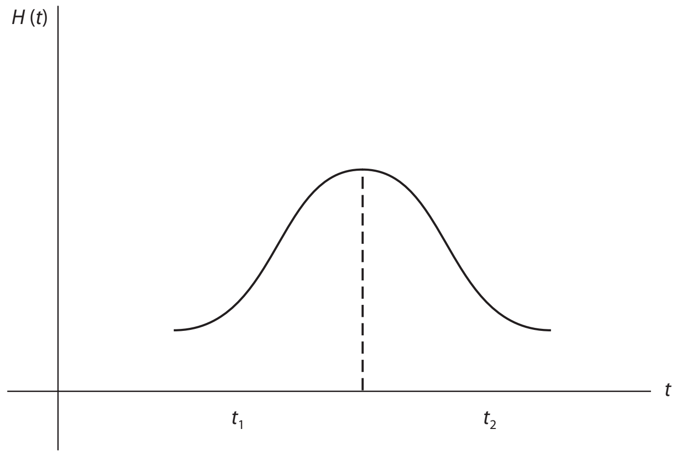

Fig. 22.2. The energy of the tunneling configuration as a function of time.

and substitute them into the Lagrangian $L = \int d^3x\mathcal{L} = \int d^3x[-\frac{1}{g^2}\,\mathrm{Tr}\,[E^2 - B^2]]$ which can be written in the following form,

$$
\begin{aligned}
L &= -\frac{1}{2}m(\lambda)(\dot{\lambda})^2 - V(\lambda), \\
m(\lambda) &= \frac{2}{g^2}\int d^3x\,\mathrm{Tr}\left(\frac{\partial A_i}{\partial\lambda}\right)^2 \geq 0, \\
V(\lambda) &= -\frac{1}{g^2}\int d^3x\,\mathrm{Tr}\,(B^i)^2 \geq 0.
\end{aligned}
\tag{22.73}
$$

The Lagrangian (22.73) is the Lagrangian of a particle that moves from one vacuum at $t = t_1$ where $V(\lambda) = m(\lambda) = 0$ to a vacuum at $t = t_2$, where again $V(\lambda) = m(\lambda) = 0$. The quantum mechanical tunneling rate is proportional to e^{-2R} where R is given by,

$$
\begin{aligned}
R &= \int_{\lambda_1}^{\lambda^2} d\lambda\sqrt{2m(\lambda)(V(\lambda) - E)} \\
&= \int_{\lambda_1}^{\lambda^2} d\lambda\sqrt{\left[\left(\frac{1}{g^2}\int d^3x\,\mathrm{Tr}\left(\frac{\partial A_i}{\partial\lambda}\right)^2\right)\left(\frac{1}{g^2}\int d^3x\,\mathrm{Tr}\,(B_i)^2\right)\right]} \\
&= \frac{2}{g^2}\int_{t_1}^{t_2} dt\sqrt{\left(\int d^3x\,\mathrm{Tr}\,(E_i)^2\right)\left(\frac{1}{g^2}\int d^3x\,\mathrm{Tr}\,(B_i)^2\right)}.
\end{aligned}
\tag{22.74}
$$

Using the triangle inequality we can relate the tunneling rate to the winding or instanton number as follows,

$$
R \geq \frac{2}{g^2}\left|\int_{t_1}^{t_2} d^4x\,\mathrm{Tr}\,[E^iB^i]\right| = \frac{8\pi^2}{g^2}|k|, \tag{22.75}
$$

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press

where the instanton number is $Q = k$. Thus we see that the tunneling rate is bounded by,

$$\mathrm{e}^{-R} \leq \mathrm{e}^{-\frac{8\pi^2}{g^2}|k|}, \tag{22.76}$$

and the bound is saturated for instanton configurations. To be more precise the most probable tunneling paths are given by a Minkowski gauge configuration with $\vec{E} = \pm\vec{B}$ which when viewed as a configuration in Euclidean space are instantons. Conversely given an instanton $A^E_\mu(x,t)$ in Euclidean space one can construct a set of paths in Minkowski space-time $A^M_\mu(x,\lambda(t))$ such that $A^{M,(\lambda)}_i(x,t) = A^E_\mu(x,\lambda(t))$ and $A^{M,(\lambda)}_0(x,t) = A^E_4(x,\lambda(t))$.

22.5 Instantons, theta vacua and the $U_A(1)$ anomaly

It was shown in the previous section that the instantons connect different vacua. This means that the vacuum of the YM theory cannot be described by any of the states of zero energy and a specific topological charge, but instead has to be a superposition of all these states, namely,

$$|\theta\rangle = \sum_k \mathrm{e}^{ik\theta}|k\rangle\ . \tag{22.77}$$

The generator of large gauge transformation that changes the winding number by one unit, namely, $T|k\rangle = |k+1\rangle$ has to be a symmetry generator that commutes with the Hamiltonian so that $T|\mathrm{vac}\rangle = \mathrm{e}^{i\varphi}|\mathrm{vac}\rangle$ for some phase φ. Indeed for the θ vacuum we get $T|\theta\rangle = \sum_k \mathrm{e}^{ik\theta}|k+1\rangle = \mathrm{e}^{-i\theta}|\theta\rangle$.

The energy associated with the θ vacua given by,

$$E(\theta) = -2K\cos(\theta)e^{-S}. \tag{22.78}$$

This follows from the following steps. Consider the amplitude to tunnel from a vacuum $|i\rangle$ to a vacuum $|j\rangle$ is given by,

$$\langle j|e^{-Ht}|i\rangle = \sum N_\pm \frac{\delta_{N_+-N_--(j-i)}}{N_+!N_-!}(Kte^{-S})^{N_++N_-}, \tag{22.79}$$

when the instantons are sufficiently dilute and where K is the pre-exponential factor in the tunneling amplitude, and $N_\pm$ are the number of instantons and anti-instantons. We introduce the parameter θ via a representation of the Kroneker delta function,

$$\delta_{N_+-N_++(i-j)} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{i\theta(N_+-N_++(i-j))}. \tag{22.80}$$

Upon performing the summations over N_+ and N_- we get,

$$\langle j|\mathrm{e}^{-Ht}|i\rangle = \int_0^{2\pi} \mathrm{e}^{i\theta(i-j)}e^{2Kt\cos(\theta)\mathrm{e}^{-S}}, \tag{22.81}$$

which implies that the energy of the θ vacuum is as given in (22.78). Note that this does not imply that the YM theory has a continuous spectrum without mass gap since the θ parameter is fixed for a given theory and it cannot be changed. Fixing the value of θ can be achieved by adding a θ term to the action (22.10),

$$S_{\rm YM} = \frac{1}{2g^2}\int \mathrm{d}^4x \,\mathrm{Tr}\, F_{\mu\nu}F^{\mu\nu} \to \frac{1}{2g^2}\int \mathrm{d}^4x \,\mathrm{Tr}\, F_{\mu\nu}F^{\mu\nu} + \frac{\theta}{16\pi^2}\int \mathrm{d}^4x \,\mathrm{Tr}\,[F_{\mu\nu}{}^*F^{\mu\nu}]. \tag{22.82}$$

The additional θ term is a surface term and hence does not affect the equations of motion, however it is not invariant under CP or T transformations.[6] As will be discussed below, with no massless quarks indeed the θ term implies a strong CP violation. The most severe restriction on CP violation comes from the electric dipole moment of the neutron. This sets the upper bound to theta to be,

$$\theta < 10^{-9}. \tag{22.83}$$

The puzzle of why θ is so tiny is referred to as the strong CP problem. One proposal to handle this problem is the introduction of the axion, χ, a pseudo scalar field with a coupling of the form $\chi\,\mathrm{Tr}\,[F_{\mu\nu}{}^*F^{\mu\nu}]$ so that the effective θ is the sum of $\sqrt{<\chi>}$ and the θ term. As will be discussed below there is in fact an even simpler mechanism to resolve the strong CP problem and that is having a massless u quark. This brings us to the next topic which is the incorporation of light quarks to the game.

In the presence of light quarks there is a simple physical observable that distinguishes between the different topological vacua, the axial current. Recall that for N_f massless quarks the theory is classically invariant under global $U_L(N_f)\times U_R(N_f) \equiv SU_L(N_f)\times SU_R(N_f)\times U_B(1)\times U_A(1)$ symmetry. The $SU(N_f)\times U_V(1)$ symmetry group factors are realized in nature also quantum mechanically. The invariance under the axial $SU(N_f)$ transformations is broken spontaneously and there are N_f^2-1 Goldstone bosons. For $N_f=2$ these are the pions. The $U_A(1)$ axial symmetry is not conserved quantum mechanically. In analogy to the anomaly of the axial symmetry in two dimensions discussed in Section 9.1, in four dimensions as well one can show using various different methods that,

$$\partial^\mu J^5_\mu = \frac{N_f}{8\pi^2}\,\mathrm{Tr}\,[F_{\mu\nu}{}^*F^{\mu\nu}], \tag{22.84}$$

where $J^5_\mu = \sum_i \bar\psi_i\gamma_\mu\gamma_5\psi_i$ is the axial current and ψ_i $i=1,\ldots,N_f$ are the fields of the various flavored quarks. This resolves the so-called $U_A(1)$ puzzle, namely the absence of the fourth Goldstone boson for $N_f=2$ or the ninth one for $N_f=3$. Indeed for the former case one could associate the η pseudo-scalar meson with the fourth Goldstone boson, however it has a mass of $478MeV$ whereas a current

[6] A proposal for resolving the strong CP problem was proposed in [172]. The $U_A(1)$ problem has been resolved by 't Hooft. The mass of the η' was proposed by Witten [221] and by Veneziano [216]. Our discussion of this topic follows the review [185].

algebra theorem states that it has to be lighter than $\sqrt{3}m_\pi$. The right-hand side of (22.84) is proportional to the divergence of the topological current (22.2) $\partial_\mu K_\mu$ so one can define a modified conserved axial current $\tilde{j}^5_\mu = j^5_\mu - N_f K_\mu$. However, unlike the topological charge, the topological current is not gauge invariant. A massless pole in the correlator of K_μ does not necessarily correspond to a massless particle. One may wonder also about the fact that the right-hand side of (22.84) is a surface term and hence cannot have a physical significance. However, as was emphasized above due to instantons the surface term is relevant. Let us see this explicitly. We start by computing the change in the axial charge,

$$\begin{aligned}\Delta Q_5 &= Q_5(t=+\infty) - Q_5(t=-\infty) = \int \mathrm{d}^4x \partial^\mu J^5_\mu \\ &= N_f \int \mathrm{d}^4x \partial^\mu \,\mathrm{Tr}\,[S(x,x)\gamma_\mu\gamma_5],\end{aligned} \tag{22.85}$$

where $S(x,y)$ is the fermion propagator $S(x,y) = <x|(i\,\not{D})^{-1}|y>$ that can be determined from the eigenfunction equation $i\,\not{D}\,\psi_\lambda = \lambda\psi_\lambda$ in the form $S(x,y) = \sum_\lambda \frac{\psi_\lambda(x)\psi^\dagger_\lambda(y)}{\lambda}$. Substituting this expression we get,

$$\Delta Q_5 = N_f \int \mathrm{d}^4x \partial^\mu \,\mathrm{Tr}\left(\sum_\lambda \frac{\psi_\lambda(x)\psi^\dagger_\lambda(y)}{\lambda} 2\lambda\gamma_5\right) = 2N_f(n_\mathrm{L} - n_\mathrm{R}), \tag{22.86}$$

where we have used the fact that ψ_λ and $\gamma_5\psi_\lambda$ are orthogonal so only the $n_\mathrm{L}(n_\mathrm{R})$ left (right) zero modes contribute.

Integrating the left-hand side of (22.84) we get the topological charge Q which is thus related to ΔQ_5. The latter counts the number of left-handed zero modes minus the number of right-handed zero modes. This is obviously associated with instantons. Each instanton contributes one unit to the topological charge and has a left-handed zero mode, whereas an anti-instanton has a right-handed zero mode and $Q = -1$. This is the way the instantons contribute to the axial anomaly and hence to the resolution of the $U_A(1)$ problem. For the case of $N_f = 3$ this implies that this would be the ninth Goldstone boson, the η' is massive even if the quark masses vanish. It was shown that the mass of the η' is related to the topological susceptibility in the following form,

$$\frac{2N_f}{f_\pi^2}\chi_\mathrm{top} = \frac{2N_f}{f_\pi^2}\int \mathrm{d}^4x <Q(x)Q(0)>= m^2_\eta + m^2_{\eta'} - 2m^2_K. \tag{22.87}$$

The combination of masses on the right-hand side corresponds to the part of the η' mass which is not due to the strange quark mass.

https://doi.org/10.1017/9781009401654.023 Published online by Cambridge University Press