path integral with $L_M \to L_M - (\theta/2\pi)\dot{\phi}$; the extra term in the Minkowski Lagrangian is real.

5.2 Large Gauge Transformations in Non-Abelian Theory.

Wave functionals representing the solutions of the equation system (4.15) are annihilated by the constraints $\hat{G}^a$, which means that they stay invariant under the action of the operator

$$\exp\left\{-i\int d^3x\ \chi^a(\boldsymbol{x})\hat{G}^a(\boldsymbol{x})\right\}$$

with small $\chi^a(\boldsymbol{x})$. This is tantamount to saying that the wave functionals are invariant under infinitesimal gauge transformations (1.10) of their arguments, $A_i \to A_i^\chi$, and hence also under finite *topologically trivial* gauge transformations (1.9), i.e. the transformations which can be reduced to a continuous series of infinitesimal gauge transformations (1.10). However, not all gauge transformations are topologically trivial. An example of a topologically nontrivial one is

$$\Omega_*(\boldsymbol{x}) \;=\; \exp\left\{\frac{i\pi x^a\sigma^a}{\sqrt{\boldsymbol{x}^2+\rho^2}}\right\} \tag{5.22}$$

with an arbitrary finite parameter ρ. We have $\Omega_*(\boldsymbol{0}) = 1$. As $|\boldsymbol{x}| \to \infty$, $\Omega_*(\boldsymbol{x}) \to -1$ irrespectively of the direction $\boldsymbol{x}/|\boldsymbol{x}|$ along which infinity is approached. For any other $|\boldsymbol{x}|$, the function $\Omega_*(\boldsymbol{x})$ gives an element of $SU(2)$, with every such element appearing only once. Therefore the function (5.22) represents a topologically nontrivial mapping of S^3 (R^3 compactified at infinity) onto the gauge group $SU(2)$. Indeed, the Chern–Simons number of the pure gauge configuration $A_i = -i\partial_i\Omega_*\Omega_*^\dagger$,

$$\begin{aligned} N_{C.S.} &= \int K_0 d^3x \\ &= \frac{1}{24\pi^2}\int d\boldsymbol{x}\ \epsilon_{ijk}\mathrm{Tr}\left\{(\partial_i\Omega_*)\Omega_*^\dagger(\partial_j\Omega_*)\Omega_*^\dagger(\partial_k\Omega_*)\Omega_*^\dagger\right\}, \end{aligned} \tag{5.23}$$

with K_μ being defined in Eq. (2.7), is equal to one. The gauge transformation (5.22) is topologically nontrivial, which means that one cannot find a family of gauge transformations $\Omega(\alpha, \boldsymbol{x})$ continuous in α such that each $\Omega(\alpha, \boldsymbol{x})$ represents a smooth function on S^3 (the smoothness requirement means in particular that the limiting value of $\Omega(\alpha, \boldsymbol{x})$ at $|\boldsymbol{x}| \to \infty$ exists and is uniquely defined; the necessity of this restriction will be seen a little bit later), $\Omega(0, \boldsymbol{x}) = 1$, and $\Omega(1, \boldsymbol{x}) = \Omega_*(\boldsymbol{x})$. $\Omega_*(\boldsymbol{x})$ may be called a "large" gauge transformation. The Gauss law constraint says nothing about what happens with the wave functional after its arguments are transformed by $\Omega_*(\boldsymbol{x})$.

Along with $\Omega_*(\boldsymbol{x})$, there are further distinct topologically nontrivial transformations $\Omega^{(q)}(\boldsymbol{x})$ belonging to the same homotopy class as (i.e. continuously deformable to) $[\Omega_*(\boldsymbol{x})]^q$ with integer q. Two $\Omega^{(q)}(\boldsymbol{x})$ with different q cannot be smoothly transformed into one another. The classes $\{\Omega^{(q)}(\boldsymbol{x})\}$ form what is called a homotopy group $\pi_3[G] = \mathbb{Z}$.

Note now that the following large gauge transformation operator,

$$\hat{U}\Psi[A_i(\boldsymbol{x})] = \Psi[A_i^{\Omega_*}(\boldsymbol{x})] \;, \tag{5.24}$$

commutes with the Hamiltonian and with all other physical gauge-invariant operators. This means that the operator (5.24) has the same meaning and can be handled in the same way as the operator (5.5) of the rotation by 2π for the pendulum. Thus, we diagonalize the Hamiltonian and the operator of the large gauge transformation (5.24) simultaneously, and divide the large Hilbert space spanned by all the eigenstates of the Hamiltonian satisfying the Gauss law constraints into sectors with a definite eigenvalue of $\hat{U}$. The latter is $e^{i\theta}$, as a large gauge transformation does not change the norm of the wave function. We are allowed and, moreover, are forced to consider only one such distinct sector because no physical operator has nonvanishing matrix elements between the states with different θ. Once finding ourselves in a state characterized by some θ, we can go over to a different state in the same sector as a result of some perturbation, but the value of θ *can*not be changed. It is a fundamental constant of our World, fixed once and forever. The dynamics of quantum Yang–Mills theory is simply not specified until the value of θ is given.

The analog of Eq. (5.11) reads

$$Z(\theta) \sim \sum_{q=-\infty}^{\infty} e^{-iq\theta} \int \prod_{\tau, \boldsymbol{x}} dA_\mu^a(\tau, \boldsymbol{x}) \exp\left\{ -\frac{1}{4g^2} \int_0^\beta d\tau \int d\boldsymbol{x} \left[(F_{\mu\nu}^a)^2 \right] \right\} , \tag{5.25}$$

with the boundary conditions

$$A_\mu(\boldsymbol{x},\beta) = [A_\mu(\boldsymbol{x},0)]^{[\Omega_*^q(\boldsymbol{x})]} \tag{5.26}$$

for the $q^{\underline{th}}$ term in the sum. The expression (5.25) defines the thermal partition function. Tending $\beta \to \infty$, and substituting $\int_0^\beta \equiv \int_{-\beta/2}^{\beta/2} \to \int_{-\infty}^{\infty}$, we obtain the partition function concentrated on the vacuum state.

Let us look at the $q = 1$ term in Eq. (5.25). The important fact is that the Pontryagin index (2.6) of Euclidean field configurations in the corresponding path integral is equal to 1. To see that, consider the configuration $A_i(-\infty, \boldsymbol{x}) = 0$, $A_i(\infty, \boldsymbol{x}) = -i\partial_i\Omega_*\Omega_*^\dagger$. Present our Euclidean space as the cylinder $S^3 \times R$ where R is Euclidean time, and S^3 is 3-dimensional space with the point at infinity added. Recall that the Pontryagin density $\sim F\tilde{F}$ is a total derivative, and that the 4-volume integral of $F\tilde{F}$ can be written as a surface integral according to Eq. (2.8). As the base S^3 of our cylinder has no boundary, the surface of $S^3 \times R$ consists of two spheres S^3 at $\tau = -\infty$ and $\tau = \infty$. The relation (2.7) is reduced to

$$\int d^4x \, \frac{1}{32\pi^2} F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} \; = \; N_{C.S.}(\tau = \infty) - N_{C.S.}(\tau = -\infty) \; . \tag{5.27}$$

Using compact S^3 instead of R^3 as a base is quite appropriate in our case, but may seem at first a bit confusing. Alternatively, one could consider a disk D^3 with large, finite radius, and an integral of the Pontryagin density over $D^3 \times R$. One can then show that the integral of the Chern–Simons current in Eq. (2.7) over the side surface of the cylinder vanishes if the radius of the disk is much larger than the characteristic scale ρ.

When the coupling constant g is small (for the time being, we do not yet understand under what conditions this is true, this will be explained in Lecture 9, devoted to asymptotic freedom), the action of all topologically non-trivial configurations is large. The integral can be done semi-classically and $Z_1 = Z_{-1} \sim \exp\{-8\pi^2/g^2\}$, where $8\pi^2/g^2$ is the action of the instanton solution studied in Lecture 2.

Thus, we understand now the physical meaning of the instanton. It is a semi-classical tunneling trajectory connecting the classical, trivial vacuum $A_i = 0$ and the configuration $-i\partial_i\Omega_*\Omega_*^\dagger$ obtained from the trivial vacuum when applying a large gauge transformation. When the coupling constant g is small, the tunneling amplitude is suppressed as $\exp\{-S_I\}$, where $S_I = 8\pi^2/g^2$ is the instanton action. Everything is the same as for quantum pendulum and the physical interpretation of θ — the quasi-momentum describing the drift of the system between the degenerate vacua characterized by different definite Chern–Simons numbers — is also the same.

We are now able to explain why the requirement that $\Omega(|\boldsymbol{x}| \to \infty)$ be uniquely defined was imposed in the first place. One could, of course, consider gauge transformations which tend to different values when $\boldsymbol{x} \to \infty$ along different directions. For example, so-called Gribov copies with $\tilde{\Omega}(\boldsymbol{x}) \sim \sqrt{\Omega_*(\boldsymbol{x})}$ have this property (see Lecture 7 for more details). It turns out, however, that the action on a tunneling trajectory interpolating between the trivial vacuum $A_i = 0$ and the configuration $-i\partial_i\tilde{\Omega}(\boldsymbol{x})\tilde{\Omega}^\dagger(\boldsymbol{x})$ has an *infinite* action and therefore does not contribute to the path integral. In fact, it is just the presence of this infinite barrier which justifies the topological classification.

This tunneling trajectory (the so-called *meron* solution) may still present some interest, however (see Lecture 7 for further discussion). For reference purposes, let us write here the explicit form of the meron solution in a covariant gauge:

$$A^a_\mu = \eta^a_{\mu\nu} x_\nu / x^2 . \tag{5.28}$$

It differs from the pure gauge field configuration (2.16) by the absense of the overall factor 2 and, in contrast to (2.16), has a nonzero field strength

$$F^a_{\mu\nu} = \frac{x_\alpha(x_\mu \eta^a_{\alpha\nu} - x_\nu \eta^a_{\alpha\mu}) - x^2 \eta^a_{\mu\nu}}{x^4} \tag{5.29}$$

and action density. The latter is proportional to $1/x^4$, and the action integral diverges logarithmically both at small and large $|x|$.

In the same way as we did it for the pendulum, one can rewrite the integral (5.25) in the form

$$Z(\theta) \sim \int \prod_{\tau,\boldsymbol{x}} dA^a_\mu(\tau,\boldsymbol{x}) \exp\left\{-\int_0^\beta d\tau \int d\boldsymbol{x} \left[\frac{1}{4g^2} F^a_{\mu\nu} F^a_{\mu\nu} + i\frac{\theta}{32\pi^2} F^a_{\mu\nu} \tilde{F}^a_{\mu\nu}\right]\right\}, \tag{5.30}$$

where the integral is performed over all field configurations which are periodic in τ up to a large gauge transformation. The second term in the integrand is a total derivative and is relevant when topologically nontrivial configurations come into play. It is the famous θ-term of QCD. Analytical continuation of the path integral to Minkowski spacetime gives

$$\mathcal{L}_M \;\to\; \mathcal{L}_M + \frac{\theta}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{a\,\mu\nu} \;, \tag{5.31}$$

and the θ-term is real.

What is the value of θ in the World we live? It is experiment rather than theory which can answer this question. Note that the θ-term breaks both P and T invariance. (Indeed, $F^a_{\mu\nu}\tilde{F}^{a\,\mu\nu}$ is the same as $\boldsymbol{E}^a \cdot \boldsymbol{B}^a$, where $\boldsymbol{E}^a$ is chromoelectric and $\boldsymbol{B}^a$ is chromomagnetic field. Now, $\boldsymbol{E}^a$ is a vector, which is odd under time reflection, while $\boldsymbol{B}^a$ is a pseudo-vector, even under time reflection.) We know, however, that neither parity nor T invariance are broken in strong interactions. This means that θ is either just zero or very small. The best experimental restriction comes from the measurements on the electric dipole moment d_n of the neutron. We know that $d_n \lesssim 10^{-25} e \cdot cm$ which means that $\theta \lesssim 10^{-9}$.[§]

The fact that θ is zero or very small has no explanation in the framework of QCD. Maybe the problem will be eventually solved when we understand the nature of the sought unified theory of all interactions, incorporating also QCD. The status of this problem is roughly the same as the status of the cosmological term problem: experiment tells us that the cosmological term is zero or very small, but we do not understand why.[¶]

[§] There is a subtlety here. All physical effects noninvariant with respect to P and T transformations are actually proportional to $\sin\theta$. Thus, the symmetry requirements alone restrict θ to be close to *either* 0 *or* π. We simply note here that the second possibility *is* excluded by what we know from experiments on the properties of light pseudoscalar mesons. This can be shown in the framework of the *effective chiral Lagrangian* describing light meson dynamics to be introduced in Lecture 12. We will discuss more of the physics of the imaginary world with $\theta \approx \pi$ in Lecture 16.

[¶] Probably, the analogy between θ and the cosmological term is even deeper. In the Ogievetsky–Sokachev formulation of supergravity, the cosmological term is also written as a total derivative [15]. But this question is far beyond the scope of this book.