DISSERTATION

't Hooft anomaly, global inconsistency, and some of their applications

Yuta Kikuchi
Department of Physics, Graduate School of Science, Kyoto University

January, 2018

Abstract

The 't Hooft anomaly, an obstruction to promoting global symmetries to local gauge symmetries, is one of the most powerful tools to study the infrared properties of the quantum field theories. The UV/IR anomaly matching condition claims that, provided the global symmetry has an 't Hooft anomaly in the ultraviolet theory, the infrared theory reached via renormalization group flow must reproduces the same 't Hooft anomaly, and particularly, the realization of the symmetric gapped state without topological order is strictly forbidden in the infrared theory. Thus, the existence of 't Hooft anomalies, combined with the anomaly matching condition, imposes strong constraints on the low-energy effective theories.

This dissertation presents 't Hooft anomalies and global inconsistencies, the latter of which are other type of obstructions to gauging the global symmetry, involving discrete symmetries and higher-form symmetries. Consequences of the 't Hooft anomaly and global inconsistency are discussed in detail via the UV/IR anomaly matching argument. Several examples are given in increasing difficulty but desirability for the purpose of extracting nonperturbative data of infrared structure of quantum chromodynamics (QCD). To be specific, we deal with quantum mechanical models, the pure Yang-Mills and bifundamental gauge theories with the theta angles, and QCD and QCD-like theories at finite temperatures. The various quantum mechanical models serve as pedagogical illustrations of 't Hooft anomalies, global inconsistencies, and their UV/IR matching arguments. Their consequences manifest themselves as level degeneracies in energy spectra, that are protected from symmetry-preserving perturbations. While the pure Yang-Mills and bifundamental gauge theories with topological terms share a lot of features in common with the quantum mechanical models, additional concepts are introduced with a tricky global symmetry: the center symmetry, which is nowadays understood as a discrete one-form symmetry. To derive the 't Hooft anomaly and global inconsistency involving the center symmetry, we describe the generalized global symmetry along with certain types of topological quantum field theories in a separated chapter. Combining all these ingredients together, we finally attack the 't Hooft anomaly in QCD. Also, massless QCD with twisted boundary conditions at finite temperatures T and chemical potentials μ is discussed, in which, despite the appearance of 't Hooft anomalies in the vacuum, its existence at finite temperatures is totally nontrivial because of the circle compactification in the imaginary-time direction. The outcome is quite remarkable: the symmetric gapped phase is strictly excluded in so called massless $\mathbb{Z}_N$-QCD at any temperature and chemical potential, i.e., the persistent order is realized for this theory at finite (T, μ). Furthermore, it turns out that an 't Hooft anomaly and a global inconsistency for certain global symmetry are responsible for the Roberge-Weiss transition at finite imaginary chemical potential.

Publication List

This dissertation is mainly based on the following three papers that I was involved in.

1. Chapters 2.2 and 3 are primarily based on [1]:
   Y. Kikuchi and Y. Tanizaki, "Global inconsistency, 't Hooft anomaly, and level crossing in quantum mechanics", *PTEP* **2017** (2017) 113B05.

2. The description in Chapter 6 is mostly following [2]:
   Y. Tanizaki and Y. Kikuchi, "Vacuum structure of bifundamental gauge theories at finite topological angles", *JHEP* **06** (2017) 102.

3. A part of results shown in Chapter 7 is presented in an unpublished paper [3]:
   Y. Tanizaki, Y. Kikuchi, T. Misumi, and N. Sakai, "Anomaly matching for phase diagram of massless $\mathbb{Z}_N$-QCD", arXiv:1711.10487 [hep-th], under review.

Acknowledgements

It is cetainly fortunate that during my graduate program I could work with my supervisor, Teiji Kunihiro, whose great works on the rich phase structure of QCD had immediately attracted my mind even before I came to Kyoto. Even though we have not (yet) worked on the exploration of QCD phase diagram, the fantastic thing is that no matter what I am working on, Kunihiro-san always provides insightful comments and patiently keeps encouraging me to work out our projects. I would like to express my gratitude to him for his all the supports. I also deeply indebted to him for his generous attitude and helpful advices on my decision to stay in the United States for a large part of my Doctor program.

It has been wonderful to have graduate students, postdocs, professors, secretaries, administrative staffs around at Kyoto University: special thanks to Yoshiko En'yo, Yoshitaka Hatta, Tetsuo Hyodo, Akira Ohnishi, Hideo Suganuma, Toshitaka Tatsumi, Kouji Kashiwa, Tongyu Lee, Mayumi Hiraoka, Tomoko Ozaki, Kiyoe Yokota and all the graduate students whom I have shared the time with for discussions, dinners, chitchatting, etc.

Due to my long-term stay at Stony Brook University in the U.S. for more than half of my Doctor program, there are many people I should acknowledge as well. I deeply thank to Dima Kharzeev for kindly having accepted my visit and supervised me during my stay at Stony Brook. His physical insight has always amazed and helped me a lot to reach physically significant results in our projects. It has been a fantastic experience for me to be in Nuclear Theory Group at Stony Brook with graduate students, postdocs and professors around: special thanks to Edward Shuryak, Derek Teaney, Jac Verbaarschot, Ismail Zahed, Yukinao Akamatsu, Takumi Iritani, Rene Meyer, Jean-Francois Paquet, Yiyang Jia, Rasmus Larsen, and the members of the Chiral Matter Friday —Yuji Hirono, Sahal Kaushik, Dima Kharzeev, Mark Mace, Evan Philips, and Yuya Tanizaki— for stimulating discussions. Especially, I spent a great amount of time with Rene in my first half year in the U.S. and he helped me adopt to unfamiliar American life including language barrier, for which I feel particular appreciation.

During my graduate program there have been many collaborations that I was involved in and I feel deep appreciation for having such brilliant collaborators: Kyosuke Tsumura, Shin'ya Gongyo, Tetsuo Hyodo, Kouji Kashiwa, Yasuhiro Yamaguchi, Shigeki Sugimoto, Chris Pak Hang Lau, Dima Kharzeev, Rene Meyer, Tomoya Hayata, Tatsuhiro Misumi, Norisuke Sakai and Yuya Tanizaki. I am especially grateful to Tanizaki-san for patiently helping me to understand from basics to advanced topics of the quantum field theory including recent development of quantum anomalies, which turns into the main subject of my dissertation.

Besides the research projects it has been fantastic to spend a certain amount of time in playing soccer with terrific friends both at Kyoto and Stony Brook. I thank all the members of the YITP soccer group at Kyoto and the Friday soccer group at Stony Brook for providing me such lovely moments.

I would like to acknowledge financial support that I have received over the past three years from the Japan Society for the Promotion of Science (JSPS), without which my stay in the U.S. as well as the work reported in this dissertation would not have been possible. I should acknowledge the RIKEN for the RIKEN Brain Circulation Program as well, that allowed me to visit the Brookhaven National Laboratory and the Stony Brook University when I was in the Master program. The stimulating experience back then urged me to stay in the U.S. My Acknowledgement also goes to the Yukawa Institute of Theoretical Physics for their financial support for my short-term stays in Kyoto every time I came back from the U.S.

I am absolutely sure that there are many people that I have failed to list here but to whom I am equally grateful. I wish to thank them all together anonymously.

At last but not least, I would like to thank my parents, Atsushi Kikuchi and Chieko Kikuchi, for their tremendous support.

Contents

Chapter 1

Introduction to quantum anomaly

Quantum field theory (QFT) provides a powerful tool unifying the relativity and quantum mechanics in high energy physics as well as the long-range description of many-body physics. Universality of QFT sometimes exhibits common aspects in seemingly quite different systems and allows us to treat them in an analogous fashion [4]. While perturbative aspects of QFTs are very well understood, and it gives an approximate result successfully when QFTs are weakly coupled, many examples of our interest are described by strongly-coupled QFTs, and a tremendous amount of effort has been made with various nonperturbative techniques to extract clues to clarify its properties. However, unraveling nonperturbative nature of the quantum field theory (QFT) is extremely challenging in general. A typical example is the quantum chromodynamic (QCD), whose infrared properties as well as phase structure are still far from complete understanding. Solution of strongly-coupled QFT is generically unknown unless it is in a special situation, such as in low dimensions, with strong-weak dualities, with certain supersymmetries, etc. It is therefore of great importance to give a rigorous statement on QFTs that applies even when they are strongly coupled.

A key clue is global symmetry of QFT. One cornerstone of traditional many-body physics is Landau's characterization of phases [5, 6]: different phases realize different symmetries. At generic values of coupling constants, the free energy is an analytic function, but some singularities must appear when symmetry breaking pattern changes. Another interesting consequence of spontaneous symmetry breaking is the existence of massless bosons called Nambu-Goldstone bosons when a continuous symmetry is spontaneously broken [7, 8]. As we see in these two famous examples, we can give rigorous statements about strongly-coupled field theories by assuming patterns of spontaneous breaking of global symmetries. This is already surprising, but it requires other nonperturbative data to answer the question about whether the symmetry is spontaneously broken or not.

Throughout this dissertation, we attempt to understand the properties of quantum anomaly to answer these questions to some extent. Among several types of anomalies, the 't Hooft anomaly [9] plays a key role to extract the nonperturbative data as we will discuss in detail. The 't Hooft anomaly, defined as an obstruction to promoting the global symmetry to local gauge symmetry, imposes strong constraint on the infrared theory via UV/IR anomaly matching argument [9–11]. The constraints involving various symmetries such as discrete symmetries [12–14] and higher-form symmetries [15–17] in addition to conventional continuous symmetries shed new light on the nonperturbative aspects of QFTs from condensed matter physics to high energy physics (see e.g. [2, 3, 18–28]).

We here give an organization of the dissertation. After a brief exposition to several types of quantum anomalies in this chapter, we give a rather abstract account of the 't Hooft anomaly and global inconsistency in Chapter 2. The global inconsistency, originally proposed in [21], is described in general context in [1, 2] with a certain refinement on the UV/IR matching condition. In Chapter 3, we elucidate 't Hooft anomalies and global inconsistencies in simple quantum mechanical models following [1] for the purpose of capturing the central idea of their UV/IR matching argument. We then switch gears from Chapter 4 to start exploring 't Hooft anomalies and global inconsistencies in QFTs. As preliminaries, the generalized global symmetry, some aspects of topological quantum field theories, and some related concepts are covered, which will be indispensable tools to investigate 't Hooft anomalies and global inconsistencies involving the "center symmetry" in the following three chapters. In Chapter 5, we start exploration of QFTs with reviewing an 't Hooft anomaly and a global inconsistency in the $SU(N)$ pure Yang-Mills theory with the θ term mostly following [21]. Although the constraints due to the 't Hooft anomaly and global inconsistency look quite similar in the pure Yang-Mills theory, we find in Chapter 6 that some difference appears in $SU(N) \times SU(N)$ gauge theory with Dirac fermions in bifundamental representation [2]. Finally, we try to unravel certain aspects of infrared structure of quantum chromodynamics (QCD) by means of 't Hooft anomalies and global inconsistencies [3]. An obvious difficulty compared with the previous two examples is lack of the center symmetry because of dynamical quarks. Nevertheless, we can detect an 't Hooft anomaly involving an "emergent" center symmetry. Furthermore, 't Hooft anomalies are used to constrain the possible phase diagram of QCD-like theory at finite temperatures and (real and imaginary) chemical potential. We briefly summarize the dissertation in Chapter 8.

Before jumping into the discussion on the 't Hooft anomaly, we first look at the other types of anomalies carefully, that would help us to set up necessary terminologies and avoid possible confusions one may encounter in the following chapters.

The anomaly was originally discovered as a solution to a puzzle: The dominant decay process of the neutral pion, $\pi^0 \to 2\gamma$, could not be explained by low energy effective theory describing the pion as a Nambu-Goldstone boson associated with spontaneous broken flavor $SU(2) \times SU(2)$ symmetry. It turned out that the chiral symmetry of QCD is explicitly broken by the quantization process [29, 30]. The absence of chiral symmetry in massless quantum electrodynamics (QED) or in massless QCD due to this mechanism is called the Adler-Bell-Jackiw (ABJ) anomaly. Generally, we often encounter the infinities in QFTs as a result of quantum corrections and they require certain regularizations, which does not preserve the symmetry of classical action of the QFT. Then, this regularizing procedure might result in the symmetry violation at the end of the computation.[1] This violation is called quantum anomaly and the anomaly sometimes breaks global symmetries and sometimes breaks gauge symmetries. In the latter case, the theory itself is inconsistent and could be used as a nonperturbative data to impose constraints on consistent gauge theory. Here, we see classic examples of different types: perturbative anomaly and global anomaly.

1.1 Perturbative anomaly

The ABJ anomaly is the most famous example of so called perturbative anomaly, which yields symmetry violation due to an weak gauge fields. The computation of the anomaly can be done by either diagramatic computation or the Fujikawa's method [31, 32] dealing with the path-integral measure under the gauge transformation. "Perturbative" sounds somewhat misleading because the anomaly is one-loop exact, and hence, provides the nonperturbative data of QFTs. The "perturbative" is actually put to emphasize the anomaly resulted under weak gauge/gravitational background fields as opposed to the global anomaly, which is absent in those circumstances as discussed in the next section. We take a brief look at the computation of the ABJ anomaly for the four dimensional

[1] The symmetry-violating regularization is not the only source of anomalies as we will see in the quantum mechanical examples in Chapter 3, where divergences do not appear but still anomalies show up.

massless Dirac fermions under $U(1)$ background fields by using the Fujikawa's method, which makes the nonperturbative nature becomes clear. The partition function is give by

$$\mathcal{Z}[A] = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, \mathrm{e}^{-S}, \tag{1.1}$$

with the Euclidean action

$$S = \int \mathrm{d}^4x \bar{\psi}\not{D}\psi, \quad \not{D} \equiv \gamma^\mu(\partial_\mu + \mathrm{i}A_\mu). \tag{1.2}$$

We consider a local chiral transformation:

$$\psi(x) \to \mathrm{e}^{\mathrm{i}\gamma_5\theta(x)}\psi(x), \quad \bar{\psi}(x) \to \bar{\psi}(x)\mathrm{e}^{\mathrm{i}\gamma_5\theta(x)}, \tag{1.3}$$

which is just a redefinition of integration variables. The key observation is that the path-integral measure transforms in nontrivial way:

$$\begin{aligned}\mathcal{D}\psi\mathcal{D}\bar{\psi} &\to \exp\left\{-2\mathrm{i}\int \mathrm{d}^4x\, \theta(x)\mathrm{tr}[\gamma_5]\delta^4(x-x)\right\}\mathcal{D}\psi\mathcal{D}\bar{\psi} \\ &\equiv \exp\left\{\int \mathrm{d}^4x\, \theta(x)\mathcal{A}(x)\right\}\mathcal{D}\psi\mathcal{D}\bar{\psi}. \end{aligned}\tag{1.4}$$

As a result of an infinitesimal chiral transformation $\delta\psi = \mathrm{i}\gamma_5\theta(x)\psi$, we obtain

$$0 = \delta\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, \mathrm{e}^{-S} = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\int \mathrm{d}^4x\, (\theta(x)\mathcal{A}(x) + J_5^\mu\partial_\mu\theta(x))\mathrm{e}^{-S} \tag{1.5}$$

which leads to a violation of local conservation law, i.e., the ABJ anomaly,

$$\langle\partial_\mu J_5^\mu\rangle = \mathcal{A}(x). \tag{1.6}$$

At first sight, the prefactor in (1.4) seems to be trivial due to $\mathrm{tr}[\gamma_5]$, which is however divergent and requires a certain regularization. The standard procedure is to introduce a gauge invariant convergence factor $\lim_{\Lambda\to\infty}\mathrm{e}^{-\not{D}^2/\Lambda^2}$ inside the trace and calculated to be,

$$\begin{aligned}&\lim_{\Lambda\to\infty}\lim_{x\to y}\mathrm{tr}[\gamma_5\mathrm{e}^{-\not{D}^2/\Lambda^2}]\delta^4(x-y) \\ &= \lim_{\Lambda\to\infty}\lim_{x\to y}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\mathrm{e}^{-ik_\mu y^\mu}\mathrm{tr}[\gamma_5\mathrm{e}^{-\not{D}^2/\Lambda^2}]\mathrm{e}^{ik_\mu x^\mu} \\ &= \lim_{\Lambda\to\infty}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\mathrm{tr}\left[\gamma_5\exp\left(\frac{(D+\mathrm{i}k)^2+[\gamma^\mu,\gamma^\nu]F_{\mu\nu}/4}{\Lambda^2}\right)\right] \\ &= \lim_{\Lambda\to\infty}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\mathrm{e}^{-k^2/\Lambda^2}\mathrm{tr}\left[\frac{1}{32\Lambda^4}\gamma_5[\gamma^\mu,\gamma^\nu][\gamma^\rho,\gamma^\sigma]F_{\mu\nu}F_{\rho\sigma}\right] \\ &= -\frac{1}{32\pi^2}\mathrm{i}\epsilon^{\mu\nu\rho\sigma}\mathrm{tr}[F_{\mu\nu}F_{\rho\sigma}], \end{aligned}\tag{1.7}$$

where we have used $\mathrm{tr}(\gamma_5[\gamma^\mu,\gamma^\nu][\gamma^\rho,\gamma^\sigma]) = -16\mathrm{i}\epsilon^{\mu\nu\rho\sigma}$ in the last step.

1.2 Global anomaly

The global anomaly involves the gauge transformations[2] which is not continuously connected to identity. In this section, we shall discuss the Witten's $SU(2)$ global anomaly by the traditional

[2] diffeomorphisms if we discuss the gravitational anomaly but we consider only in flat spacetime

spectral flow argument [33, 34]. Let us consider single Weyl fermion coupled with background $SU(2)$ gauge field in four dimensional Euclidean space. A gauge transformation g which vanishes at spatial infinity is classified by the homotopy group $\pi^4(SU(2)) = \mathbb{Z}_2$, meaning that the gauge transformations of nontrivial element of the homotopy group is not continuously connected to identity. The partition function is given by

$$\mathcal{Z}[A] = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \mathrm{e}^{-\int \mathrm{d}^4x \bar{\psi}\mathrm{i}\not{D}\psi} = \mathrm{Pf}[\mathrm{i}\not{D}] \tag{1.8}$$

where Pf denotes the Pfaffian. This expression is invariant under the infinitesimal gauge transformation, corresponding to a trivial element of the homotopy group. The $\mathrm{Pf}[\mathrm{i}\not{D}]$ is formally expressed by a product of $\mathrm{i}\not{D}$,

$$\mathrm{Pf}[\mathrm{i}\not{D}] = \prod_i \lambda_i. \tag{1.9}$$

The problem is that the sign of the partition function cannot be determined in gauge invariant way due to the infinite product of negative eigenvalues.

Let us take a closer look at the indefinite sign from the viewpoint of spectral flow of the eigenvalues. More specifically, we see the spectral flow along a smooth variation of background gauge field $A \to A^g$. g is a topologically nontrivial gauge transformation under which A transforms into A^g. We parametrize the smooth variation of A by

$$A_t = (1-t)A + tA^g, \quad 0 \le t \le 1. \tag{1.10}$$

It would be problematic if the partition function $\mathcal{Z}[A]$ changes the sign odd times along the variation of t from 0 to 1 because the gauge invariance requires

$$\mathcal{Z}[A] = \mathcal{Z}[A^g] = |\mathcal{Z}[A]|(-1)^{\mathfrak{I}}, \tag{1.11}$$

for some integer $\mathfrak{I}$. In other words, such a spectral flow contradicts to the gauge invariance. This is a typical spectral flow argument for the global anomaly, that is, the global aspect of gauge fields encoded in the spectral flow results in the ambiguity of the partition function and gauge theory itself is inconsistent.

Now, what we need to know is the integer $\mathfrak{I}$, which is given by the index of the Dirac operator on the five dimensional manifold $X = M^4 \times \mathbb{R}$. The manifold M^4 is the original four dimensional one parametrized by x, and the fifth direction $\mathbb{R}$ is parametrized by $-\infty < \tau < \infty$. The gauge configuration is given by A at $\tau \to -\infty$ and by A^g at $\tau \to \infty$, between which A and A^g are adiabatically connected by $A(x,\tau)$.[3] The five dimensional Dirac operator is concretely given by

$$\not{D}_X = \gamma_\tau \partial_t + \sum_{\mu=1}^{4} \gamma_\mu D_\mu. \tag{1.12}$$

Under the adiabatic limit, $\partial_t D_\mu \approx 0$, the solution of Dirac equation $\not{D}_X \Psi{=}0$ can be expanded in terms of the eigenfunction of $\gamma_\tau \not{D}$,

$$\Psi(x,\tau) = \sum_\lambda f_\lambda(\tau)\phi_\lambda(x). \tag{1.13}$$

$f_\lambda(\tau)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}\tau} f_\lambda(\tau) = -\lambda(\tau) f_\lambda(\tau), \tag{1.14}$$

[3] The index $\mathfrak{I}$ may be equally calculated by the index of the Dirac operator on the mapping torus. It is constructed by identifying the fifth direction as t in (1.10) and glueing the ends of the cylinder, $t = 0$ and $t = 1$.

which is solved by

$$f_\lambda(\tau) = f_\lambda(0)\exp\left(\int_0^\tau \mathrm{d}\tau\lambda(\tau)\right) \tag{1.15}$$

Therefore, the number of normalizable zero modes of five dimensional Dirac operator is equal to the number of zero crossings of eigenvalues from negative to positive. In other words, the five dimensional Dirac index is equivalent to the sign change of the partition function. We will repeat almost the same argument in Section 1.4 to introduce the Atiyah-Patodi-Singer's (APS) index theorem. If we take M^4 as S^4 for instance, the index can be readily calculated to be odd, i.e., $\mathfrak{I} = 1 \pmod 2$ by the APS index theorem. Therefore, the partition function is not invariant under the topologically nontrivial gauge transformation. Hence, the theory of single four dimensional Weyl fermion coupled to $SU(2)$ background gauge field is inconsistent.

1.3 Parity anomaly

We shall discuss the parity anomaly [12–14] by closely following [18]. Although the spectral flow of the Dirac operator exhibits similar behavior to that in the last section, further analysis is required to draw a precise and even more interesting conclusion.

Let ψ be a three dimensional massless Dirac fermion with action

$$S = \int \mathrm{d}^3x\bar{\psi}\mathrm{i}\not{D}\psi, \tag{1.16}$$

which is time reversal (T) invariant and finite mass of Dirac fermion would violate T invariance. The partition function in Minkowski space is formally calculated to be

$$\mathcal{Z}[A] = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \mathrm{e}^{\mathrm{i}S} = \det \mathrm{i}\not{D} \tag{1.17}$$

Since the Dirac operator is hermitian, the eigenvalues λ_i is real. Thus, the partition function is also real and given by

$$\det \mathrm{i}\not{D} = \prod_i \lambda_i. \tag{1.18}$$

It is noted that $\{\lambda_i\}$ again contains infinitely many negative eigenvalues, which leads to indefinite sign of the partition function. This is a symptom of the global anomaly that we saw in the last section.

The spectral flow argument developed around (1.10) can be equally applied to result in ambiguity in the sign of the partition function. The sign is indeed calculated by introducing a Dirac operator on a four dimensional cylinder or a mapping torus. But it is too early to conclude that this theory is inconsistent even if the index is odd or equivalently the sign is minus. The question is whether it is still possible to draw a physically sensible answer out of the seemingly pathological partition function.

One of the crucial conditions which was implicitly assumed in the above discussion is the reality of the partition function, i.e., T invariance of the theory. Indeed we can obtain a consistent theory by discarding the T symmetry, leading to the parity anomaly [12–14]. We introduce Pauli-Villars regulator with a large mass M, which explicitly breaks T invariance,

$$\mathcal{Z} = \lim_{M\to\infty}\prod_i \frac{\lambda_i}{\lambda_i + \mathrm{i}M} = |\mathcal{Z}|\exp\left(-\frac{\mathrm{i}\pi}{2}\sum_i \mathrm{sign}(\lambda_i)\right). \tag{1.19}$$

We note that $|\lambda_i| \gg M$ does not contribute to the sign of $\mathcal{Z}$ thanks to the regularization. With a proper regularization we obtain

$$\mathcal{Z} = |\mathcal{Z}| \exp\left(-\mathrm{i}\pi \frac{\eta}{2}\right), \tag{1.20}$$

where η is defined by [35–37],

$$\eta \equiv \lim_{s \to 0} \sum_i \frac{\mathrm{sign}(\lambda_i)}{|\lambda_i|^s}. \tag{1.21}$$

The resultant partition function is gauge invariant due to the regularization at the cost of T invariance, which is broken by an imaginary part controlled by η. As we noted, the classical action for the massless Dirac fermion has T symmetry, which is however broken on the quantum level. This is the parity anomaly.

In this model, the relation between the global anomaly and the parity anomaly is summarized as follows. If we try to preserve the T invariance, or equivalently the reality of the partition function, the theory becomes inconsistent because the gauge invariance is lost according to the spectral flow argument. Although this appears to be the global gauge anomaly, it is not quite correct in this case because we do not have to keep the reality of the partition function at the cost of gauge invariance. Indeed the gauge invariance can be maintained by giving up the T invariance, which makes more sense because the violation of gauge invariance invalidate the theory itself while the theory with broken T symmetry is still acceptable. The regularization which preserves gauge invariance results in the T broken partition function, i.e., the parity anomaly manifest itself via its imaginary part.

1.4 Index theorem, boundary, and anomaly inflow

In the last section, we had to give up T invariance in order to obtain gauge invariant theory for massless Dirac fermion in three dimension. But it is well known that three dimensional massless Dirac fermion exists with T symmetry without violating $U(1)$ gauge invariance, that is, the boundary state of the T invariant topological insulator. We see what is going on in this system from the viewpoint of the anomaly inflow.

1.4.1 Index theorem

Before getting into the physical setup realized in the topological insulator, we recall the index theorem.

The Atiyah-Singer's (AS) index theorem [38–44] states that, given a Dirac operator on a closed four-manifold X, the Dirac index $\mathfrak{I}$, which is defined to be a difference between the number of the Dirac zero modes with positive and negative chirality $\mathfrak{I} = n_+ - n_-$, and the instanton number are tied together via the following relation:

$$\mathfrak{I} = \int_X \frac{\mathrm{tr} F \wedge F}{8\pi^2}, \tag{1.22}$$

where F is a field strength of gauge group G and the right-hand side yields an integer on a closed four-manifold X.

The situation relevant to the topological insulator is described on a four-manifold X with a boundary three-manifold Y. A subtle issue arises on the boundary, where the topological nature of the right hand-side in eq. (1.22) is invalidated, and so is the AS index. This invalidation of the index can also be attributed to the absence of the local boundary condition which preserves chirality. Indeed, the most general local boundary condition maintaining the rotational symmetry on the boundary Y is given by

$$n \cdot \gamma\psi = \pm\psi, \tag{1.23}$$

where n is a unit vector normal to Y. But it does not preserve the chirality because γ^μ anticommutes with $\gamma_5 = \frac{1}{4!}\epsilon^{ijkl}\gamma_i\gamma_j\gamma_k\gamma_l$. The Atiyah-Patodi-Singer's (APS) index theorem [35–37] provides a sensible formula for $\mathfrak{I}$ even on a manifold with boundaries by introducing nonlocal and chirality-conserving boundary condition:

$$\mathfrak{I} = \int_X \frac{\mathrm{tr}F \wedge F}{8\pi^2} - \frac{\eta(D_Y)}{2}, \tag{1.24}$$

where D_Y is the Dirac operator on Y and $\eta(D_Y)$ is an associated η-invariant, which measures the spectral asymmetry of D_Y.

Instead of following the original derivation [35], we present the derivation in a simplified setup considered in [45] (see also [12, 46]), which allows us to understand the connection between four-dimensional Dirac index and three dimensional spectral flow. We consider the Dirac operator,

$$D_X \equiv \mathrm{i}\not{D} = \mathrm{i}(\gamma_t\partial_t + \sum_{i=1}^{3}\gamma_i D_i) = \mathrm{i}\tau_1 \otimes 1_{2\times 2}\partial_t + \tau_2 \otimes D_Y \tag{1.25}$$

on a Euclidean four-manifold $X = Y \times \mathbb{R}$ with $-\infty < t < \infty$ parametrizing $\mathbb{R}$, and boundaries exist at the ends $t = \{-\infty, \infty\}$. $D_Y = \mathrm{i}\sum_{i=1,2,3}\sigma_i D_i$ is the Dirac operator on Y and we took $A_t = 0$ gauge. The relations between γ matrices and Pauli matrices are given in chiral representation by

$$\begin{aligned}
&\gamma_t = \begin{pmatrix} 0 & 1_{2\times 2} \\ 1_{2\times 2} & 0 \end{pmatrix} = \tau_1 \otimes 1_{2\times 2}, \quad \gamma_{i=1,2,3} = \begin{pmatrix} 0 & -\mathrm{i}\sigma_i \\ \mathrm{i}\sigma_i & 0 \end{pmatrix} = \tau_2 \otimes \sigma_i, \\
&\gamma_5 = \gamma_t\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} -1_{2\times 2} & 0 \\ 0 & 1_{2\times 2} \end{pmatrix} = -\tau_3 \otimes 1_{2\times 2},
\end{aligned} \tag{1.26}$$

We compute the index of the Dirac operator (1.25) in the adiabatic limit, where D_Y changes slowly with t. The idea is the following: The index of D_X is given by the spectral flow of D_Y. The spectral flow of D_Y is the signed sum of the zero crossing of the eigenvalues $\{\lambda(t)\}$ along t running from $-\infty$ to ∞, which is related to the η invariant. The function $\eta(D_Y)$ smoothly evolves except at those points where eigenvalues of D_Y pass through zero. At those points $\eta(D_Y)$ jumps discontinuously by ± 2. Therefore, the spectral flow of D_Y is half the discontinuous change in $\eta(D_Y)$ as t runs from $-\infty$ to ∞.

In the adiabatic limit, one can construct the zero modes of D by using an eigenspinor ψ_Y of D_Y with an eigenvalue $\lambda(y)$ and a two-component spinor $f(y)$ such that

$$(\mathrm{i}\tau_1\partial_t + \tau_2 \otimes D_Y)\psi_Y \otimes f(t) \approx 0, \tag{1.27}$$

which leads to $(\partial_t + \tau_3\lambda)f(t) = 0$. Hence,

$$f(t) = \mathrm{e}^{-\int^t \tau_3\lambda}\chi, \tag{1.28}$$

where χ is a constant two-component spinor. To obtain normalizable solutions we require boundary conditions at the ends of cylinder, that is, $\tau_3\lambda > 0$ (< 0) at $t = \infty$ $(-\infty)$. Therefore, there exists one normalizable zero mode for each eigenspinor χ of τ_3 whose eigenvalue is $+(-)$ with the positive (negative) zero crossing of $\lambda(t)$. Noting that the chirality is given by $-\tau_3 \otimes 1_{2\times 2}$ we see that the index of D corresponds to the spectral flow of D_Y.

Assuming that zero-crossings occur at points $\{t_1, \cdots, t_k\}$, the index $\mathfrak{I}$ and $\eta(D_Y)$ are related as

$$
\begin{aligned}
-\mathfrak{I} &= \sum_{\{t_i\}} \mathrm{sgn}(\text{zero crossing}) = \frac{1}{2}\sum_{i=1}^{k}\left[\eta(D_Y)|_{t_i^+} - \eta(D_Y)|_{t_i^-}\right] \\
&= \frac{1}{2}\left[\eta(D_Y)|_{t=+\infty} - \eta(D_Y)|_{t=-\infty}\right] - \frac{1}{2}\sum_{i=0}^{k}\left[\eta(D_Y)|_{t_{i+1}^-} - \eta(D_Y)|_{t_i^+}\right] \\
&= \frac{1}{2}\left[\eta(D_Y)|_{t=+\infty} - \eta(D_Y)|_{t=-\infty}\right] - \frac{1}{2}\sum_{i=0}^{k}\int_{t_i^+}^{t_{i+1}^-} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t}\eta(D_Y).
\end{aligned}
\tag{1.29}
$$

where we have defined $t_0 = -\infty$ and $t_{k+1} = +\infty$.

We shall show that the second term corresponds to the four-dimensional index density. It is useful to introduce the following integral:

$$
\eta(s) = -\frac{2}{\Gamma(\frac{s+1}{2})}\int_0^\infty \mathrm{d}u\, u^s \mathrm{Tr}[D_Y \mathrm{e}^{-u^2 D_Y^2}],
\tag{1.30}
$$

where $\eta(D_Y) = \eta(0)$. Then $\mathrm{d}\eta(D_Y)/\mathrm{d}t$ is calculated to be

$$
\begin{aligned}
\frac{\mathrm{d}\eta(D_Y)}{\mathrm{d}t} &= -\lim_{s\to 0}\frac{2}{\Gamma(\frac{s+1}{2})}\int_0^\infty \mathrm{d}u\, u^s \mathrm{Tr}\left[\left(\frac{\partial D_Y}{\partial t} - 2u^2 D_Y^2 \frac{\partial D_Y}{\partial t}\right)\mathrm{e}^{-u^2 D_Y^2}\right] \\
&= -\lim_{s\to 0}\frac{2}{\Gamma(\frac{s+1}{2})}\int_0^\infty \mathrm{d}u\, u^s \frac{\partial}{\partial u}\mathrm{Tr}\left[u\frac{\partial D_Y}{\partial t}\mathrm{e}^{-u^2 D_Y^2}\right] \\
&= \frac{2}{\sqrt{\pi}}\lim_{u\to 0}\mathrm{Tr}\left[u\frac{\partial D_Y}{\partial t}\mathrm{e}^{-u^2 D_Y^2}\right].
\end{aligned}
\tag{1.31}
$$

This expression is indeed shown to be identical to the index density of D as follows. From (1.25), its square is given by

$$
D_X^2 = -\partial_t^2 + D_Y^2 - \tau_3 \partial_t D_Y.
\tag{1.32}
$$

The index density is

$$
\begin{aligned}
&\lim_{\Lambda\to\infty}\int_Y \mathrm{tr}[\tau_3 \mathrm{e}^{-D^2/\Lambda^2}] \\
&= \lim_{\Lambda\to\infty}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\mathrm{Tr}\,\tau_3 \mathrm{e}^{[(\mathrm{i}\omega+\partial_t)^2 - D_Y^2 + \tau_3\partial_t D_Y]/\Lambda^2} \\
&= \lim_{\Lambda\to\infty}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{-\omega^2/\Lambda^2}\mathrm{Tr}\,\tau_3\left(\tau_3\frac{\partial_t D_Y}{\Lambda^2}\right)\mathrm{e}^{-D_Y^2/\Lambda^2} \\
&= \frac{1}{\sqrt{\pi}}\lim_{\Lambda\to\infty}\mathrm{Tr}\,\frac{1}{\Lambda^2}\frac{\partial D_Y}{\partial t}\mathrm{e}^{-D_Y^2/\Lambda^2} \\
&= \frac{1}{2}\frac{\mathrm{d}\eta(D_Y)}{\mathrm{d}t}.
\end{aligned}
\tag{1.33}
$$

Since we already showed in Section 1.1,

$$
\lim_{\Lambda\to\infty}\int_{\mathbb{R}}\mathrm{d}t\int_Y \mathrm{d}^3y\, \mathrm{tr}[\sigma_3 \mathrm{e}^{-D^2/\Lambda^2}] = \int_{Y\times\mathbb{R}}\frac{\mathrm{tr}F\wedge F}{8\pi^2},
\tag{1.34}
$$

we finally obtain,

$$\mathfrak{I} = \int_{Y\times\mathbb{R}} \frac{\mathrm{tr} F\wedge F}{8\pi^2} - \frac{1}{2}\big[\eta(D_Y)|_{y=+\infty} - \eta(D_Y)|_{y=-\infty}\big]. \tag{1.35}$$

The APS index theorem can be generalized to more complicated manifold X with a boundary Y as discussed in [35–37]. Then, the η in (1.35) is simply replaced by that on the corresponding boundary Y[4]:

$$\mathfrak{I} = \int_{X} \frac{\mathrm{tr} F\wedge F}{8\pi^2} - \frac{1}{2}\eta(D_Y). \tag{1.37}$$

1.4.2 Anomaly inflow

After getting a necessary tool, the APS index formula (1.37), let us go back to our original problem. Single massless Dirac fermion coupled to $U(1)$ background gauge field suffers from the Parity anomaly as we saw in Section 1.3, and the corresponding partition function takes the form,

$$\mathcal{Z} = |\mathcal{Z}| \exp\left(-\mathrm{i}\pi\frac{\eta}{2}\right). \tag{1.38}$$

Now, the bulk of the topological insulator plays an pivotal role. Letting the bulk manifold X and its boundary manifold Y, we suppose the three dimensional theory is realized on a boundary state of the bulk theory which provides

$$\mathcal{Z}_{\mathrm{bulk}} = \exp\left(\mathrm{i}\pi\int_X \frac{\mathrm{tr} F\wedge F}{8\pi^2}\right). \tag{1.39}$$

Then, by combining (1.38) and (1.39) together and employing the APS formula (1.37),

$$\mathcal{Z}\mathcal{Z}_{\mathrm{bulk}} = |\mathcal{Z}| \exp\left(\mathrm{i}\pi\int_X \frac{\mathrm{tr} F\wedge F}{8\pi^2} - \mathrm{i}\pi\frac{\eta}{2}\right) = |\mathcal{Z}|(-1)^{\mathfrak{I}}. \tag{1.40}$$

Now, we have got a satisfactory answer in the sense that both gauge and T invariances are maintained. The anomaly on the boundary Y is compensated by that in the bulk X, which leads to the anomaly-free state as a whole system. This anomaly cancellation is called anomaly inflow mechanism [47]. Due to the parity anomaly, the massless Dirac fermion cannot exist on three manifold by itself, but it can be realized as a boundary state of certain bulk theory. Hence, it is a manifestation of the bulk-boundary correspondence.

[4]More generally, the APS index theorem is given by

$$\mathfrak{I} = \int_X \widehat{A}(R)\mathrm{ch}(F) - \frac{1}{2}\big[\eta(D_Y)|_{y=+\infty} - \eta(D_Y)|_{y=-\infty}\big]. \tag{1.36}$$

on a curved even-dimensional space X with a boundary Y.

Chapter 2

’t Hooft anomaly, global inconsistency, and UV/IR matching

An ’t Hooft anomaly is defined as an obstruction to promoting the global symmetry to local gauge symmetry [9–11, 48, 49]. The anomalies that we discussed in the last Chapter are not of this type. What is the difference? Let us start with clarifying their difference.

- Gauge anomaly :
  Gauge invariance is violated on the quantum level. Hence, the theory is inconsistent. For instance, four dimensional massless Dirac fermion with $U_V(1) \times U_A(1)$ gauge symmetry has gauge anomaly.

- ABJ-type anomaly :
  Global symmetry is violated on the quantum level. Four dimensional massless Dirac fermion with $U_V(1)$ gauge symmetry and classical $U_A(1)$ global symmetry is an example. The $U_A(1)$ global symmetry of the classical action is explicitly broken due to the quantization. This is the ABJ anomaly discussed in Section 1.1.

- ’t Hooft anomaly[1] :
  There exists an obstruction when one try to promote the global symmetry of the theory to local gauge symmetry. It is noted that the theory has nothing wrong on the quantum level in contrast to above two types of anomalies in the sense that the theory is perfectly consistent and no global symmetry is broken in the quantum theory. An example is four dimensional massless Dirac fermion with $U_V(1) \times U_A(1)$ global symmetry. The theory is consistent and its global symmetry is exact. An inconsistency shows up only if you try to gauge the $U_V(1) \times U_A(1)$ symmetry or couple to its background gauge field.

We should emphasize that, although the existence of an ’t Hooft anomaly itself does not mean breaking of symmetries without coupling to background gauge fields, it provides important non-perturbative data of QFTs as it imposes nontrivial constraints on low energy dynamics of theories because of the anomaly matching condition. More specifically, the anomaly matching condition states that the low-energy effective field theory of the QFT $\mathcal{T}$ must also follow the same transformation law (2.1) under the background G-gauge field A and the G-gauge transformation $A \mapsto A + \mathrm{d}\alpha$. We have more to say on this argument later in this chapter.

[1] Although the concept of the ’t Hooft anomaly and anomaly matching condition were introduced by ’t Hooft in 1979 [9], the anomaly was named ’t Hooft anomaly only recently by Kapustin and Thorngren in 2014 [10] to distinguish it from anomalies of the other two types mentioned above.

One of the questions we would like to address in this and following chapters is whether we can derive a nontrivial result even when the 't Hooft anomaly is absent. In Ref. [21], a new condition, so called the global inconsistency of gauging symmetries, is proposed in order to claim the nontrivial consequence similar to the 't Hooft anomaly. They considered about the four dimensional $SU(N)$ Yang Mills theory at $\theta = \pi$, and the mixed 't Hooft anomaly is found for the center symmetry and time reversal symmetry when N is even. This derives the spontaneous breaking of time reversal symmetry at $\theta = \pi$ under a certain assumption (see [50–63] for early related discussions). But it turns out that the 't Hooft anomaly does not exists when N is odd. Nevertheless they found a similar obstruction to gauging the symmetry by pointing out that the local counter terms for gauging the center symmetry at $\theta = 0$ and $\theta = \pi$ must be different in order to be compatible with time reversal symmetry at those points, and this global inconsistency is claimed to lead to the same consequence as the 't Hooft anomaly either at $\theta = 0$ and $\theta = \pi$: If the phase of one side (say, $\theta = 0$) is trivial, then the phase of the other side ($\theta = \pi$) must be nontrivial. In Refs. [1, 2], we suggested a new possibility that is compatible with the global inconsistency: The global inconsistency can be saturated in IR by the phase transition separating those time reversal symmetric points when the vacua at those points are trivially gapped. The phase structure of the $SU(N) \times SU(N)$ bifundamental gauge theory with finite topological angles is determined under some assumptions with this proposal. In this situation, it would be nice to discuss various solvable models with the global inconsistency to check what kinds of possibility can be realized.

All these discussions will be described in a lot more detail in the following sections from various perspectives. The purpose of this chapter is to give a review of 't Hooft anomalies followed by an introduction of the global inconsistency for gauging symmetries, which will be fully utilized to unravel the infrared behavior of quantum theories briefly mentioned above in the forthcoming chapters.

2.1 't Hooft anomaly and anomaly matching argument

An 't Hooft anomaly is defined as an obstruction to promoting the global symmetry to local gauge symmetry [9, 10]. We consider a QFT $\mathcal{T}$ with a global symmetry G, and let $\mathcal{Z}[A]$ be the partition function of $\mathcal{T}$ under the background G-gauge field A. We say that G has an 't Hooft anomaly if the partition function $\mathcal{Z}$ follows the nontrivial transformation law[2],

$$\mathcal{Z}[A + \mathrm{d}\alpha] = \mathcal{Z}[A] \exp\left(\mathrm{i}\mathcal{A}[\alpha, A]\right), \tag{2.1}$$

under the G-gauge transformation $A \mapsto A + \mathrm{d}\alpha$ and $\mathcal{A}[\alpha, A]$ cannot be canceled by local counter terms. Especially when $G = G_1 \times G_2$, G_1 and G_2 is said to have a mixed 't Hooft anomaly if G_1 and G_2 themselves have no 't Hooft anomaly but $G_1 \times G_2$ has an 't Hooft anomaly.

Let us have a look at an example raised above, a four dimensional massless Dirac fermion with the global symmetry $G = U_V(1) \times U_A(1)$. To confirm the existence of the 't Hooft anomaly we consider the theory under the background G-gauge field with a partition function $\mathcal{Z}[A_V, A_A]$, where A_V and A_A are background gauge fields for symmetry group $U_V(1)$ and $U_A(1)$, respectively. $\mathcal{Z}[A_V, A_A]$ transforms

$$\mathcal{Z}[A_A + \mathrm{d}\alpha_A] = \mathcal{Z}[A] \exp\left[\frac{\mathrm{i}}{8\pi^2}\int \alpha_A \left(F_V \wedge F_V + \frac{1}{3}F_A \wedge F_A\right)\right], \tag{2.2}$$

[2]Other obstructions to gauging the symmetry exist as shown in Ref. [11] when the symmetry is discrete, but we do not consider such subtle obstructions in this paper. The anomaly inflow for 't Hooft anomaly not of Dijkgraaf–Witten type is discussed in Ref. [64].

under $U_V(1)$ local gauge transformations.[3] Indeed, the gauge invariance is violated by a nontrivial phase. Hence, it is forbidden to promote the global symmetry group G to gauge group, meaning that the global G symmetry has an 't Hooft anomaly.

2.1.1 UV/IR anomaly matching

The 't Hooft UV/IR anomaly matching states that the low-energy effective field theory of the QFT $\mathcal{T}$ must also follow the same transformation law (2.1) under the background G-gauge field A and the G-gauge transformation $A \mapsto A + \mathrm{d}\alpha$. Original proof of this statement is given, when G is the continuous chiral symmetry, by introducing the spectator chiral fermions canceling the 't Hooft anomaly and by making the G-gauge field A dynamical. Since the coupling of $\mathcal{T}$ to the G-gauge field A can be made arbitrarily small, the low-energy effective theory of $\mathcal{T}$ is unaffected by the presence of A and should produce the same phase $\mathcal{A}[\alpha, A]$ under the G-gauge transformation in order to cancel the G-gauge anomaly from the spectator fermions [9] (See also [65, 66] for review).

Another proof is given by the important observation that the phase functional $\mathcal{A}[\alpha, A]$ can be written as the boundary term of the gauge transformation of a topological G-gauge theory in one-higher dimension. This is proven when G is the continuous chiral symmetry in even dimension [67, 68], and it is true in many examples with discrete global symmetries [10, 11, 69]. When this is true, we can put the theory $\mathcal{T}$ on the boundary manifold of the topological G-gauge theory, and then the low-energy effective theory must be able to lie on the same boundary manifold. As a result, the anomaly inflow [47] derives the anomaly matching. Latest developments on the understanding of topological materials lead to discoveries of new 't Hooft anomalies that include discrete syemmetries [12–14] or higher-form symmetries [15–17] in the context of high energy and condensed matter physics, and they derive nontrivial consequences of low-energy effective theories [2, 18–28].

What kind of constraints can be obtained for infrared theories provided G has an 't Hooft anomaly? There are three possible candidates which satisfy the anomaly matching condition:

- G is unbroken and the theory contains massless excitation,
- G is unbroken, the vacuum (or vacua) is gapped, and the theory possesses topological degrees of freedom,
- G is spontaneously broken.

These realizations of vacua are referred to as "nontrivial". In other words, the existence of 't Hooft anomaly rules out the realization of the "trivial" vacuum.

Here, we give an example to illustrate in more detail how the anomaly matching argument works. The classic example of 't Hooft anomaly is the flavor symmetry $G = SU(N_\mathrm{f})_\mathrm{L} \times SU(N_\mathrm{f})_\mathrm{R} \times U(1)_\mathrm{V}$ of massless QCD [9]. The anomalous phase factor of the 't Hooft G-anomaly is characterized by the anomaly index $d_{\alpha\beta\gamma} = \mathrm{tr}[\{T_\alpha, T_\beta\}, T_\gamma]$, with the flavor symmetry generators $T_\alpha, T_\beta, T_\gamma$, which is the obstruction to gauging the G-symmetry. As outlined above we introduce the spectator fermions which generate the anomaly index $-d_{\alpha\beta\gamma}$ to cancel the obstruction to gauging. "Spectator" indicates a property that it is singlet under the color gauge group and hence does not participate in the strong dynamics. The existence of spectator allows us to safely promote the global G-symmetry group to the local gauge group. Now, we look at the corresponding infrared theory, which is described by a certain low energy effective theory. Since the strong dynamics does not affect the spectator, its anomaly should not change in the ultraviolet and infrared theories. Therefore, the 't Hooft anomaly for the

[3] The transformation given in (2.2) corresponds to the consistent anomaly, which satisfies the Wess-Zumino consistency condition, as opposed to the covariant anomaly, whose local current is covariant by adding a local counter term at the cost of the consistency condition.

G-symmetry of the original theory $d_{\alpha\beta\gamma}$ should also be unchanged because the gauge invariance of the whole theory including spectator is always maintained. The consequence is as follows: If there is no massless fermionic degrees of freedom to reproduce the required anomaly in infrared theory, the global symmetry has to be spontaneously broken in infrared, where the anomaly is indeed saturated by the Wess-Zumino term.

To see this argument more explicitly, let us derive the anomaly constraint for the massless QCD with the gauge group $SU(N_c)$ and the flavor group $G = SU(N_f)_L \times SU(N_f)_R \times U(1)_V$ [9, 65]. The constraints come from the 't Hooft anomaly associated to $SU(N_f)_L$-$SU(N_f)_L$-$SU(N_f)_L$ (and $SU(N_f)_R$-$SU(N_f)_R$-$SU(N_f)_R$) current triplet and $SU(N_f)_L$-$SU(N_f)_L$-$U(1)$ (and $SU(N_f)_R$-$SU(N_f)_R$-$U(1)$) current triplet. The former yields the anomaly index,

$$d_{\alpha\beta\gamma} = N_c \mathrm{tr}[\{t_\alpha, t_\beta\}, t_\gamma], \tag{2.3}$$

with the $SU(N_f)$ generators $t_\alpha, t_\beta, t_\gamma$ satisfying $\mathrm{tr}[t_\alpha, t_\beta] = \delta_{\alpha\beta}/2$. The anomaly for the latter is

$$d_{\alpha\beta} \equiv N_c \mathrm{tr}[\{t_\alpha, t_\beta\}] = N_c \delta_{\alpha\beta}, \tag{2.4}$$

where the $U(1)$ charge is taken to be unity.

Next, we look at the low-energy effective theory assuming that the flavor symmetry is unbroken. The assumption requires that the anomaly indices should be matched. In other words, the flavor symmetry must be spontaneously broken if there is no way to match the anomaly. We suppose that a color-singlet fermionic bound state emerging in the infrared theory is composed of N_c massless elementary quarks, i.e.,

$$m_L + m_R = N_c, \tag{2.5}$$

where m_L and m_R are numbers of left- and right-handed quarks in the bound state. Letting (r, s) be the irreducible representation of $SU(N_f)_L \times SU(N_f)_R$ and $\ell(r, s)$ be the number of times the irreducible representation (r, s) appears in the bound state, the anomaly matching conditions read

$$d_{\alpha\beta\gamma} = N_c \mathrm{tr}[\{t_\alpha, t_\beta\}, t_\gamma] = \sum_{r,s} \ell(r,s) d_s \mathrm{tr}[\{t^r_\alpha, t^r_\beta\}, t^r_\gamma], \tag{2.6}$$

$$d_{\alpha\beta} = N_c \delta_{\alpha\beta} = \sum_{r,s} \ell(r,s) d_s N_c \mathrm{tr}[\{t^r_\alpha, t^r_\beta\}], \tag{2.7}$$

where tr_r and $t^r_\alpha, t^r_\beta, t^r_\gamma$ are the trace and generators in the irreducible representation r of $SU(N_f)_L$. d_s is the dimension of representation s of $SU(N_f)_R$. Useful relations to reduce the above constraints are

$$\mathrm{tr}[\{t^r_\alpha, t^r_\beta\}, t^r_\gamma] = C_r \mathrm{tr}[\{t_\alpha, t_\beta\}, t_\gamma], \tag{2.8}$$

$$\mathrm{tr}[\{t^r_\alpha, t^r_\beta\}] = K_r \mathrm{tr}[\{t^r_\alpha, t^r_\beta\}]. \tag{2.9}$$

The constraint (2.5) with $N_\mathrm{c}=3$ allows the following representations

<table>
<tr><th>representation</th><th>ℓ</th><th>d_s</th><th>K_r</th><th>C_r</th></tr>
<tr><td>$\begin{array}{|c|c|c|}\hline \mathrm{L} & \mathrm{L} & \mathrm{L} \\ \hline \end{array}$</td><td>$\ell_a$</td><td>$1$</td><td>$\frac{(N_\mathrm{f}+3)(N_\mathrm{f}+6)}{2}$</td><td>$\frac{(N_\mathrm{f}+2)(N_\mathrm{f}+3)}{2}$</td></tr>
<tr><td>$\begin{array}{|c|}\hline \mathrm{L} \\ \hline \mathrm{L} \\ \hline \mathrm{L} \\ \hline \end{array}$</td><td>$\ell_b$</td><td>$1$</td><td>$\frac{(N_\mathrm{f}-3)(N_\mathrm{f}-6)}{2}$</td><td>$\frac{(N_\mathrm{f}-2)(N_\mathrm{f}-3)}{2}$</td></tr>
<tr><td>$\begin{array}{|c|c|}\hline \mathrm{L} & \mathrm{L} \\ \hline \mathrm{L} \\ \cline{1-1} \end{array}$</td><td>$\ell_c$</td><td>$1$</td><td>$N_\mathrm{f}^2-9$</td><td>$N_\mathrm{f}^2-3$</td></tr>
<tr><td>$\begin{array}{|c|c|}\hline \mathrm{L} & \mathrm{L} \\ \hline \end{array} \otimes \begin{array}{|c|}\hline \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_d$</td><td>$N_\mathrm{f}$</td><td>$N_\mathrm{f}+4$</td><td>$N_\mathrm{f}+2$</td></tr>
<tr><td>$\begin{array}{|c|}\hline \mathrm{L} \\ \hline \mathrm{L} \\ \hline \end{array} \otimes \begin{array}{|c|}\hline \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_e$</td><td>$N_\mathrm{f}$</td><td>$N_\mathrm{f}-4$</td><td>$N_\mathrm{f}-2$</td></tr>
<tr><td>$\begin{array}{|c|}\hline \mathrm{L} \\ \hline \end{array} \otimes \begin{array}{|c|c|}\hline \mathrm{R} & \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_f$</td><td>$\frac{N_\mathrm{f}(N_\mathrm{f}+1)}{2}$</td><td>$1$</td><td>$1$</td></tr>
<tr><td>$\begin{array}{|c|}\hline \mathrm{L} \\ \hline \end{array} \otimes \begin{array}{|c|}\hline \mathrm{R} \\ \hline \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_g$</td><td>$\frac{N_\mathrm{f}(N_\mathrm{f}-1)}{2}$</td><td>$1$</td><td>$1$</td></tr>
<tr><td>$\begin{array}{|c|c|c|}\hline \mathrm{R} & \mathrm{R} & \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_h$</td><td>$\frac{N_\mathrm{f}(N_\mathrm{f}+1)(N_\mathrm{f}+2)}{6}$</td><td>$0$</td><td>$0$</td></tr>
<tr><td>$\begin{array}{|c|}\hline \mathrm{R} \\ \hline \mathrm{R} \\ \hline \mathrm{R} \\ \hline \end{array}$</td><td>$\ell_i$</td><td>$\frac{N_\mathrm{f}(N_\mathrm{f}-1)(N_\mathrm{f}-2)}{6}$</td><td>$0$</td><td>$0$</td></tr>
<tr><td>$\begin{array}{|c|c|}\hline \mathrm{R} & \mathrm{R} \\ \hline \mathrm{R} \\ \cline{1-1} \end{array}$</td><td>$\ell_j$</td><td>$\frac{N_\mathrm{f}(N_\mathrm{f}^2-1)}{3}$</td><td>$0$</td><td>$0$</td></tr>
</table>

(2.10)

with parameters needed in evaluation of (2.8) and (2.9). $SU(N_\mathrm{f})_\mathrm{L}$-$SU(N_\mathrm{f})_\mathrm{L}$-$SU(N_\mathrm{f})_\mathrm{L}$ anomaly (2.8) leads to the constraint,

$$\begin{aligned} &\ell_a\frac{(N_\mathrm{f}+3)(N_\mathrm{f}+6)}{2}+\ell_b\frac{(N_\mathrm{f}-3)(N_\mathrm{f}-6)}{2}+\ell_c(N_\mathrm{f}^2-9) \\ &+\ell_d N_\mathrm{f}(N_\mathrm{f}+4)+\ell_e N_\mathrm{f}(N_\mathrm{f}-4)+\ell_f\frac{N_\mathrm{f}(N_\mathrm{f}+1)}{2}+\ell_g\frac{N_\mathrm{f}(N_\mathrm{f}-1)}{2}=3. \end{aligned} \tag{2.11}$$

$SU(N_\mathrm{f})_\mathrm{L}$-$SU(N_\mathrm{f})_\mathrm{L}$-$U(1)$ anomaly (2.9) leads to the constraint,

$$\begin{aligned} &\ell_a\frac{(N_\mathrm{f}+2)(N_\mathrm{f}+3)}{2}+\ell_b\frac{(N_\mathrm{f}-2)(N_\mathrm{f}-3)}{2}+\ell_c(N_\mathrm{f}^2-3) \\ &+\ell_d N_\mathrm{f}(N_\mathrm{f}+2)+\ell_e N_\mathrm{f}(N_\mathrm{f}-2)+\ell_f\frac{N_\mathrm{f}(N_\mathrm{f}+1)}{2}+\ell_g\frac{N_\mathrm{f}(N_\mathrm{f}-1)}{2}=1. \end{aligned} \tag{2.12}$$

The latter cannot have a solution if N_f is an integer multiple of three as the left-hand side is always an integer multiple of three but the right-hand side is clearly not. Therefore, the flavor symmetry $SU(N_\mathrm{f})_\mathrm{L}\times SU(N_\mathrm{f})_\mathrm{R}\times U(1)_\mathrm{V}$ must be spontaneously broken for $N_\mathrm{c}=N_\mathrm{f}=3$, for instance. Furthermore, since vector-like symmetries are protected in the vacuum [70], we conclude that the chiral symmetry is always broken. These constraints are also applied to the finite density except that the vector-like symmetry are not protected anymore. This anomaly matching argument is consistent with the prediction of the color-superconductivity phase at high quark number density region of the QCD phase diagram [71, 72]. We will have a lot to say on anomaly constraints at finite temperature in Section 7.2.

2.2 Global inconsistency

In this section, we define the global inconsistency, that is proposed in [21] and refined in our work [1] in general context of QFT.

Let $\mathcal{T}$ be a QFT parametrized by continuous parameters $\vec{g} = (g^1, g^2, \cdots)$ such as mass parameters, coupling constants, theta angles, and so on, which is described by a partition function $\mathcal{Z}_{\vec{g}}$. At generic values of $\vec{g}$, the QFT $\mathcal{T}(\vec{g})$ has the global symmetry G, and we assume that G has no 't Hooft anomaly. By this assumption, we can couple the theory $\mathcal{T}(\vec{g})$ to the background G-gauge field without breaking the invariance under the G-gauge transformation. In this process, the topological G-gauge theory on the same dimension is introduced, and the parameter space is extended by new couplings $\vec{k}$ of the topological G-gauge theory (see Chapter 4 for a detailed discussion on topological quantum field theories). Some of them might be continuous but the other of them will be quantized to ensure the G-gauge invariance, and we assume, for simplicity, that all the new couplings $\vec{k}$ is quantized to discrete values[4]. We denote the partition function under the background G-gauge field A as $\mathcal{Z}_{\vec{g},\vec{k}}[A]$, and it satisfies

$$\mathcal{Z}_{\vec{g},\vec{k}}[A + \mathrm{d}\alpha] = \mathcal{Z}_{\vec{g},\vec{k}}[A] \tag{2.13}$$

under the G-gauge transformation $A \mapsto A + \mathrm{d}\alpha$. When making the G-gauge field A dynamical, we call the obtained theory as $(\mathcal{T}(\vec{g})/G)_{\vec{k}}$, and the global symmetry disappears at generic point of $\vec{g}$.

Although the symmetry of the theory $\mathcal{T}(\vec{g})$ is G for generic $\vec{g}$, it may be enhanced to other group at special points. Let $\vec{g}_1$ and $\vec{g}_2$ be such special points, where the symmetry is enhanced to $G \times H$ by the group H, and we shall refer to these points $\vec{g}_1$ and $\vec{g}_2$ as high symmetry points. We restrict our attention to the case where $G \times H$ has no 't Hooft anomaly both at $\vec{g}_1$ and $\vec{g}_2$. In this setting, the global inconsistency is defined as follows: There exists no $\vec{k}$ such that $\mathcal{Z}_{\vec{g},\vec{k}}[A]$ is compatible with the H-gauge invariance both at $\vec{g}_1$ and $\vec{g}_2$.

Let us take a closer look at the global inconsistency. Since there is no 't Hooft anomaly for $G \times H$ at $\vec{g}_i$ $(i = 1, 2)$, there exists $\vec{k}_i$ such that

$$\mathcal{Z}_{\vec{g}_i,\vec{k}_i}[h \cdot A] = \mathcal{Z}_{\vec{g}_i,\vec{k}_i}[A], \tag{2.14}$$

where $h \cdot A$ is the transformation of G-gauge field A by $h \in H$. The condition for the global inconsistency states that $\vec{k}_1 \neq \vec{k}_2$. When $\vec{k} = \vec{k}_1$ is chosen, the symmetry H at $\vec{g}_2$ is explicitly broken as

$$\mathcal{Z}_{\vec{g}_2,\vec{k}_1}[h \cdot A] = \mathcal{Z}_{\vec{g}_2,\vec{k}_1}[A] \exp\left(\mathrm{i}\mathcal{A}_{\vec{g}_2,\vec{k}_1}[h, A]\right) \tag{2.15}$$

for some phase functional $\mathcal{A}_{\vec{g}_2,\vec{k}_1}$. Therefore, $(\mathcal{T}(\vec{g}_1)/G)_{\vec{k}_1}$ has the symmetry H, but $(\mathcal{T}(\vec{g})/G)_{\vec{k}_1}$ has no symmetry including $\vec{g} = \vec{g}_2$. The similar equation

$$\mathcal{Z}_{\vec{g}_1,\vec{k}_2}[h \cdot A] = \mathcal{Z}_{\vec{g}_1,\vec{k}_2}[A] \exp\left(\mathrm{i}\mathcal{A}_{\vec{g}_1,\vec{k}_2}[h, A]\right), \tag{2.16}$$

is true at $\vec{g}_1$ when $\vec{k} = \vec{k}_2$ is chosen: $(\mathcal{T}(\vec{g}_2)/G)_{\vec{k}_2}$ has the symmetry H, but $(\mathcal{T}(\vec{g})/G)_{\vec{k}_2}$ has no symmetry including $\vec{g} = \vec{g}_1$. It should be noted that $\vec{k}$ cannot be chosen individually at each point

[4] An example of the discrete parameter $\vec{k}$ is the level of the Chern-Simons theory.

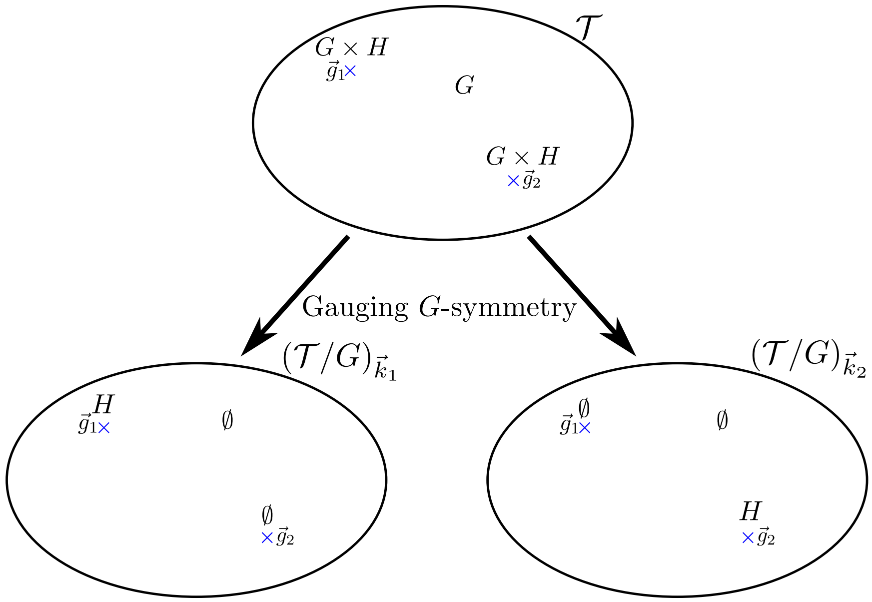


Figure 2.1: The schematic figure illustrating the global inconsistency in the space of coupling constants $\vec{g}$. In the original theory $\mathcal{T}$, symmetry G exists at generic couplings $\vec{g}$ and is enhanced by H at $\vec{g}_1$ and $\vec{g}_2$. To gauge the symmetry G, $\mathcal{T}$ is coupled to the topological G-gauge theory with the discrete parameter $\vec{k}$. In $(\mathcal{T}/G)_{\vec{k}_1}$, the symmetry is absent except at $\vec{g} = \vec{g}_1$. In $(\mathcal{T}/G)_{\vec{k}_2}$, the symmetry is absent except at $\vec{g} = \vec{g}_2$. This figure is taken from Ref. [1] with a slight modification.

because the two points are continuously connected in the parameter space and $\vec{k}$, being a discrete parameter, does not change discontinuously on the path connecting the points[5]. This situation is schematically shown in Fig. 2.1.

It should be emphasized that the inconsistent points have to be connected continuously in parameter space. When there is the global inconsistency between $\vec{g}_1$ and $\vec{g}_2$, we claim that

- The vacuum either of $\mathcal{T}(\vec{g}_1)$ or of $\mathcal{T}(\vec{g}_2)$ is nontrivial, or
- $\vec{g}_1$ and $\vec{g}_2$ are separated by the phase transition.

This is the consequence of UV/IR matching condition for the global inconsistency. When the first statement is realized, the global inconsistency shows the existence of the nontrivial phase at one of the high symmetry points. Meanwhile, the second statement, whose possibility was pointed out and carefully examined in some models in our work [1, 2], suggests that the global inconsistency is automatically satisfied if there is a phase transition separating the high symmetry points where the discrete parameter k may jump. This aspect makes the global inconsistency a milder obstruction than the 't Hooft anomaly and an important corollary is that the existence of global inconsistency does not necessarily lead to nontrivial infrared theory at high symmetry points.

[5]If $\vec{k}$ contains continuous parameters, the corresponding condition is replaced as follows: The global inconsistency exists if there is no connected component of the $\vec{k}$ space that respects full symmetries at both $\vec{g}_1$ and $\vec{g}_2$.

Chapter 3

Quantum mechanics with topological terms

The purpose of this chapter is to elucidate significance of global inconsistency as well as mixed 't Hooft anomaly in rather simple quantum mechanical models mostly following our work [1]. One of the model is reminiscent of $SU(N)$ Yang Mills theory, which possesses 't Hooft anomaly for even N and global inconsistency for odd N [21] as we will carefully study in Chapter 5. The similarity was emphasized in Ref. [23] in terms of two- and three-dimensional Abelian-Higgs models. The other is similar to $SU(N) \times SU(N)$ bifundamental gauge theory, and they also share several properties in common in view of symmetries and anomalies, which we will cover in full detail in Chapter 6. We analyze these models in two ways: the operator formalism and the path integral formalism. In the former method, we find the central extension of representations of symmetry groups. In the latter method, we see inconsistency in the local counter term when promoting global symmetries to local gauge redundancies. Although these two methods do not necessarily give the same information about anomalies, we shall see their connection by explicit computation in our models. Since energy spectra and corresponding states are calculable, we can clarify consequences of global inconsistency and 't Hooft anomaly explicitly.

Since this chapter is slightly longer than the others, we give an organization of this Chapter: In Section 3.1, we discuss a particle moving around a circle with a periodic potential. We see how to detect 't Hooft anomaly and global inconsistency in the system and discuss consequences on the energy spectrum. In Section 3.2, we add another variable to the model discussed in Section 3.1 to mimic the $SU(N) \times SU(N)$ bifundamental gauge theory at finite θ angles. The global inconsistency plays an even more important role in this model and we present the resultant energy spectrum and its interpretations. We give a short summary of this chapter in Section 3.4.

3.1 Quantum mechanics of a particle on S^1

We consider the quantum mechanics on a circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ with the topological θ term, describing a particle with unit mass moving on a ring of unit radius. θ term arises due to the flux threading the ring. The Euclidean classical action is

$$S[q] = \int \mathrm{d}\tau \left[\frac{1}{2}\dot{q}^2 + V(Nq)\right] - \frac{\mathrm{i}\theta}{2\pi}\int \mathrm{d}q. \tag{3.1}$$

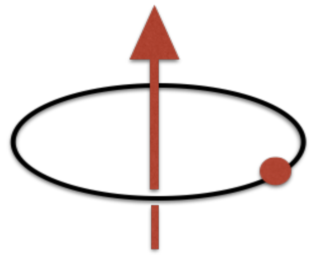


Figure 3.1: A particle on a ring with fluxes.

The potential $V(x)$ is an arbitrary 2π periodic smooth function, $V(x+2\pi)=V(x)$, and it can be represented as the Fourier series,

$$V(x)=\sum_{\ell\geq 1}\lambda_\ell \cos(\ell x+\alpha_\ell). \tag{3.2}$$

Each q is the map $q: S^1_\beta \to \mathbb{R}/2\pi\mathbb{Z}$, where S^1_β is the circle with the circumference β, $\dot{q}_i=\mathrm{d}q_i/\mathrm{d}\tau$, and $N\geq 2$ is an integer. The set of parameters is $\vec{g}=(\theta,\lambda_1,\ldots,\alpha_1,\ldots)$, and we often denote only θ instead of $\vec{g}$ since the most important parameter in our discussion is θ. The parameter θ is identified with $\theta+2\pi$ because $\int \mathrm{d}q\in 2\pi\mathbb{Z}$. The partition function $\mathcal{Z}_\theta$ is defined by the path integral,

$$\mathcal{Z}_\theta=\int \mathcal{D}q \exp\left(-S[q]\right). \tag{3.3}$$

In the operator formalism, the Hamiltonian of this system is given by

$$\hat{H}(\hat{p},\hat{q})=\frac{1}{2}\left(\hat{p}-\frac{\theta}{2\pi}\right)^2+V(N\hat{q}), \tag{3.4}$$

where $[\hat{q},\hat{p}]=\mathrm{i}$ and the Hilbert space $\mathcal{H}$ is the set of 2π-periodic L^2-functions; the partition function is $\mathcal{Z}_\theta=\mathrm{tr}_{\mathcal{H}}[\exp(-\beta\hat{H})]$.

The goal of this section is to figure out the consequences of 't Hooft anomaly and global inconsistency in this model. The aspect of the 't Hooft anomaly for this model is already discussed in detail when $N=2$ and $\alpha_\ell=0$ in Appendix of Ref. [21]. We would like to start with this model since it is the simplest case where the global inconsistency shows up when N is odd. We will indeed see that (non-accidental) level crossings appearing in the energy spectrum can be explained in terms not only of 't Hooft anomalies but also of global inconsistency.

3.1.1 Symmetries, central extension, and global inconsistency

The system (3.1) has the $\mathbb{Z}_N$ symmetry, generated by

$$\mathsf{U}: q(\tau)\mapsto q(\tau)+\frac{2\pi}{N}. \tag{3.5}$$

Since $q(\tau)$ and $q(\tau)+2\pi$ is identified on the circle, $\mathsf{U}^N=1$. Quantum mechanically, the symmetry operator U can be realized as

$$\mathsf{U}=\exp\left(\mathrm{i}\frac{2\pi}{N}\hat{p}\right), \tag{3.6}$$

and it is easy to check that $\mathsf{U}\hat{H}\mathsf{U}^{-1}=\hat{H}$ for any θ. We take this convention for U in the following.

The symmetry of the system is $\mathbb{Z}_N$ for generic θ, but there are additional symmetry at $\theta = 0, \pi$. These two points are the high symmetry points of (3.1), where we have the time reversal symmetry T,

$$\mathsf{T} : q(\tau) \mapsto q(-\tau),\ \dot{q}(\tau) \mapsto -\dot{q}(-\tau). \tag{3.7}$$

At $\theta = 0$, the action S is quadratic in $\dot{q}$, and this symmetry exists trivially. At generic θ, the topological term is linear in $\dot{q}$ and the time reversal symmetry is absent. At $\theta = \pi$, if we perform this transformation, the topological term changes as

$$-\frac{\mathrm{i}}{2}\int \mathrm{d}q \mapsto \frac{\mathrm{i}}{2}\int \mathrm{d}q = -\frac{\mathrm{i}}{2}\int \mathrm{d}q + \mathrm{i}\int \mathrm{d}q. \tag{3.8}$$

Since $\int \mathrm{d}q \in 2\pi\mathbb{Z}$, the path-integral weight $\exp(-S)$ does not change under T. Therefore, the time reversal is the symmetry also at $\theta = \pi$.

Let us study the commutation relation of U and T [21]. Two requirements, $\mathsf{T}\hat{H}\mathsf{T}^{-1} = \hat{H}$ and $\mathsf{T}\mathrm{i}\mathsf{T}^{-1} = -\mathrm{i}$, are satisfied by

$$\mathsf{T}\hat{q}\mathsf{T}^{-1} = \hat{q}, \quad \mathsf{T}\hat{p}\mathsf{T}^{-1} = \begin{cases} -\hat{p} & (\theta = 0), \\ -\hat{p} + 1 & (\theta = \pi). \end{cases} \tag{3.9}$$

If we choose the coordinate basis (i.e. $\hat{q} = q$ and $\hat{p} = -\mathrm{i}\partial_q$), we can realize T as $\mathsf{T} = \mathcal{K}$ at $\theta = 0$, and $\mathsf{T} = \exp(\mathrm{i}q)\mathcal{K}$ at $\theta = \pi$, where $\mathcal{K}$ is the complex conjugation. Using the expression (3.6) and the above commutation relation for T, we find that

$$\mathsf{T}\mathsf{U}\mathsf{T}^{-1} = \begin{cases} \mathsf{U}, & (\theta = 0), \\ \mathrm{e}^{-2\pi\mathrm{i}/N}\mathsf{U}, & (\theta = \pi). \end{cases} \tag{3.10}$$

We have several remarks on the central extension of symmetry group based on the commutation relations (3.10). At $\theta = 0$, the $\mathbb{Z}_N$ transformation and time reversal ($\mathbb{Z}_2$) transformation commute as we expected from the enhanced symmetry $\mathbb{Z}_N \times \mathbb{Z}_2$. However, at $\theta = \pi$ we have an additional phase factor, which may or may not be absorbed by properly redefining the operator. The symmetry group $\mathbb{Z}_N \times \mathbb{Z}_2$ is said to be centrally extended when there is no proper redefinition to absorb the phase factor, which is the central element. Let us redefine the operator by $\mathsf{U}' \equiv \mathrm{e}^{-\frac{2\pi\mathrm{i}k}{N}}\mathsf{U}$ for some integer k. Substituting U' back into the second commutation relation (3.10), we obtain

$$\mathsf{T}\mathsf{U}'\mathsf{T}^{-1} = \exp\left(\frac{2\pi\mathrm{i}}{N}(2k-1)\right)\mathsf{U}'. \tag{3.11}$$

Hence the phase factor can be absorbed when the following condition is satisfied:

$$2k - 1 = 0 \pmod{N}. \tag{3.12}$$

Since there is no solution for k to be an integer when $N \in 2\mathbb{Z}$, the symmetry group is centrally extended. If we try to redefine the operator with a solution of (3.12), which is a half integer for even N, the redefined operator U' satisfies $(\mathsf{U}')^N = -1$ unlike $\mathsf{U}^N = 1$. This means that we get a double cover of the original symmetry group $\mathbb{Z}_N \times \mathbb{Z}_2$. We shall see in the next section that this is the consequence of the 't Hooft anomaly between $\mathbb{Z}_N$ and the time-reversal symmetry [21]. When $N \in 2\mathbb{Z}+1$, we can redefine the operator U' by choosing $k = (N+1)/2$, which is an integer. Since we succeeded in defining U' with maintaining $(\mathsf{U}')^N = 1$, there is no central extension for odd N. This is not the end of story. Although there is no central extension at $\theta = 0$ and π separately for odd integer N, we cannot avoid the central extension at $\theta = 0$ and π simultaneously by choosing a common operator U (or U'). This fact implies the global inconsistency.

Let us discuss how the above argument constraints the energy spectrum. First let us consider the case when $N \in 2\mathbb{Z}$. Let us ask whether there exists a simultaneous eigenstate of U and T at $\theta = \pi$. We assume for contradiction that such a state exists and denote it by $|\psi\rangle$. By assumption, we can set

$$\mathsf{U}|\psi\rangle = \mathrm{e}^{2\pi \mathrm{i} k/N}|\psi\rangle, \quad \mathsf{T}|\psi\rangle = \eta|\psi\rangle. \tag{3.13}$$

Here $k \in \mathbb{Z}$, because $\mathsf{U}^N = 1$ on $\mathcal{H}$. Using the commutation relation (3.10), we obtain

$$\exp\left(\frac{2\pi \mathrm{i}}{N} k\right) = \exp\left(\frac{2\pi \mathrm{i}}{N}(1-k)\right). \tag{3.14}$$

This can be rewritten as (3.12). When N is even, this does not have any integer solutions: The simultaneous eigenstate of U and T cannot exist at $\theta = \pi$, and all the energy eigenvalues is two-fold degenerate.

Next, let us consider the case when $N \in 2\mathbb{Z}+1$. In this case, we shall find no 't Hooft anomaly, and thus the simultaneous eigenstate can exist at $\theta = \pi$. Indeed, we obtain the same condition (3.14) for the simultaneous eigenstates of U and T at $\theta = \pi$, and the possible $\mathbb{Z}_N$ charge is determined as $k = (N+1)/2$ modulo N when N is an odd integer. Even in this situation, the global inconsistency between $\theta = 0$ and π can derive a nontrivial result: No states can be singlet both at $\theta = 0$ and $\theta = \pi$. Let $|\psi_0\rangle$ be a simultaneous eigenstate of U and T at $\theta = 0$, then the similar computation shows that

$$\mathsf{U}|\psi_0\rangle = |\psi_0\rangle. \tag{3.15}$$

Let $|\psi_\pi\rangle$ be a simultaneous eigenstate of U and T at $\theta = \pi$, then the above argument has shown that

$$\mathsf{U}|\psi_\pi\rangle = \exp\left(\frac{2\pi \mathrm{i}}{N}\frac{N+1}{2}\right)|\psi_\pi\rangle. \tag{3.16}$$

Since $|\psi_0\rangle$ and $|\psi_\pi\rangle$ have different $\mathbb{Z}_N$ charge, those states cannot be continuously connected by changing the parameter θ of the theory. In other words, the T-invariant states at $\theta = 0$ break T at $\theta = \pi$, and vice versa.

To make the above arguments more convincing, let us compute the energy spectrum explicitly for a potential,

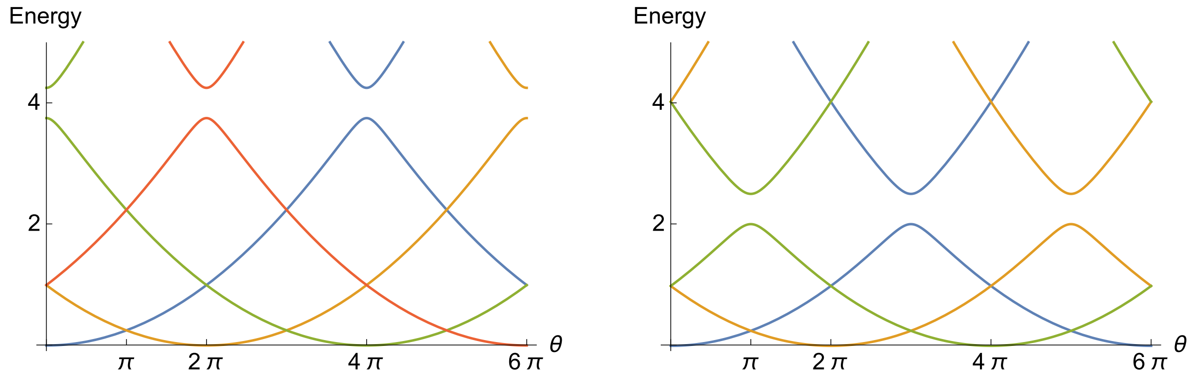
$$V(Nq) = \lambda \cos(Nq). \tag{3.17}$$

Figure 3.2 shows the energy spectra for the cases $N = 4$ and $N = 3$ that are computed numerically by diagonalizing the Hamiltonian.

As we can see in Fig. 3.2a, no state can be singlet at $\theta = \pi$ when $N = 4$ and this is expected because of the nontrivial commutation relation between U and T. When $N = 3$, there are singlet states at $\theta = \pi$ as shown in Fig. 3.2b, and this is allowed from the commutation relation. The point is that a singlet state at $\theta = 0$ and a singlet state at $\theta = \pi$ are not connected continuously by changing θ from 0 to π. Since there is no level crossing between 0 and π in this example, this condition suggests that the ground state at $\theta = \pi$ is two-fold degenerate and the time reversal symmetry is spontaneously broken, and this is realized in Fig. 3.2b.

3.1.2 Gauging $\mathbb{Z}_N$ symmetry, 't Hooft anomaly, and global inconsistency

In order to make the connection between the general discussion in Sec. 2.2 and the computation in Sec. 3.1.1, we rewrite everything using the path integral formalism of this model. We discuss the 't Hooft anomaly and global inconsistency of the quantum mechanics (3.1) in this subsection, and the connection between them will be established in the next subsection.



(a) $n = 4$

(b) $n = 3$

Figure 3.2: Energy levels as functions of θ with $\lambda = 0.5$ in (3.17) for $\mathbb{Z}_4$ and $\mathbb{Z}_3$ symmetric cases, respectively. Each color corresponds to different $\mathbb{Z}_N$ charge. (a) Every state forms a pair at $\theta = \pi, 3\pi, 5\pi$, which is a consequence of the t' Hooft anomaly. (b) Not every state forms a pair at $\theta = \pi, 3\pi, 5\pi$. But, a singlet state at $\theta = 0$ are not continuously connected to a singlet state at $\theta = \pi$, which is a consequence of the global inconsistency. This figure is taken from Ref. [1].

To analyze the 't Hooft anomaly or global inconsistency, we promote the global $\mathbb{Z}_N$ symmetry of (3.1) to the local gauge symmetry, and it can be done by coupling the theory (3.1) to a $\mathbb{Z}_N$ topological gauge theory [16]. First, let us write down the continuum description of the $\mathbb{Z}_N$ topological gauge theory,

$$S_{\mathrm{top},k} = \mathrm{i}\int F \wedge (\mathrm{d}B - NA) + \mathrm{i}k\int A. \tag{3.18}$$

Here, $A = A_0 \mathrm{d}\tau$ is the $U(1)$ one-form gauge field, B is the $U(1)$ scalar gauge field, and F is the scalar auxiliary field introduced as the Lagrange multiplier. The second term is the one-dimensional Chern-Simons term, and the level k must be an integer for invariance under the $U(1)$ gauge transformation,

$$A \mapsto A + \mathrm{d}\lambda,\ B \mapsto B + N\lambda,\ F \mapsto F. \tag{3.19}$$

The level k is identified with $k + N$ because the equation of motion of F gives

$$NA = \mathrm{d}B, \tag{3.20}$$

and thus $N\int A = \int \mathrm{d}B \in 2\pi\mathbb{Z}$. We can regard this pair (A, B) as the $\mathbb{Z}_N$ gauge field. Let us briefly describe why they can be seen as $\mathbb{Z}_N$ gauge field. We notice that, since A is flat $\mathrm{d}A = 0$, it is solely characterized by the holonomy $\exp\left[\mathrm{i}\int_{S^1} A\right]$, which is a gauge-invariant observable of the theory. Because the constraint implies

$$\left(\exp\left[\mathrm{i}\int_{S^1} A\right]\right)^N = \exp\left(\mathrm{i}\int_{S^1} \mathrm{d}B\right) = 1, \tag{3.21}$$

the holonomy takes values $\mathrm{e}^{2\pi\mathrm{i}n/N}$ for $n \in \mathbb{Z}_N$, which indeed agrees with those in $\mathbb{Z}_N$ gauge theory. Therefore, we claim that A becomes a $\mathbb{Z}_N$ gauge field under the constraint (3.20). We will have a lot more to say on gauging discrete symmetry and the topological quantum field theory in Chapter 4.

In order to make the following discussion simpler, we integrate out F: The topological action becomes

$$S_{\mathrm{top},k}[A, B] = \mathrm{i}k\int A, \tag{3.22}$$

and B dependence appears implicitly through the constraints (3.20). Next, we couple (3.1) to the topological $\mathbb{Z}_N$-gauge theory (3.22) by postulating the following transformation of q under the $U(1)$ guage transformation (3.19),

$$q \mapsto q - \lambda. \tag{3.23}$$

The gauge-invariant combinations are $\mathrm{d}q + A$ and $Nq + B$, and thus the gauge-invariant action becomes

$$S[q, A, B] = \int \mathrm{d}\tau \left[\frac{1}{2}(\dot{q} + A_0)^2 + V(Nq + B)\right] - \frac{\mathrm{i}\theta}{2\pi}\int (\mathrm{d}q + A) + S_{\mathrm{top},k}. \tag{3.24}$$

We can readily get the partition function $\mathcal{Z}_{\theta,k}[(A, B)]$ under the background $\mathbb{Z}_N$ gauge field (A, B) as

$$\mathcal{Z}_{\theta,k}[(A, B)] = \int \mathcal{D}q \exp(-S[q, A, B]), \tag{3.25}$$

and the set of couplings is extended by the Chern-Simons level $k \in \mathbb{Z}_N$.

The time reversal operation T of the Euclidean path integral is chosen as follows:

$$q(\tau) \mapsto q(-\tau),\ A_0(\tau) \mapsto -A_0(-\tau),\ B(\tau) \mapsto B(-\tau). \tag{3.26}$$

The transformation of the dynamical variable q is same as the original one (3.7), and the transformation of background fields are chosen in such a way that the equation of motion is unchanged. That is, the covariant derivative $(\dot{q} + A_0)$ is changed to $-(\dot{q} + A_0)$, and $nA = \mathrm{d}B$ is unchanged under this time reversal transformation. Under this transformation, let us check the property of the partition function under the background gauge field at $\theta = 0, \pi$.

The original theory is time reversal invariant at $\theta = 0$ and π. At $\theta = 0$, the topological θ term is absent, and thus the T transformation only flips the sign of the Chern-Simons term:

$$\mathrm{i}k \int A \mapsto -\mathrm{i}k \int A = \mathrm{i}k \int A - 2\mathrm{i}k \int A. \tag{3.27}$$

Therefore, the transformation law of the partition function at $\theta = 0$ is

$$\mathcal{Z}_{0,k}[\mathsf{T} \cdot (A, B)] = \mathcal{Z}_{0,k}[(A, B)] \exp\left(2\mathrm{i}k \int A\right). \tag{3.28}$$

We can eliminate the additional phase of (3.28) by choosing appropriate k, i.e.,

$$2k = 0 \pmod N. \tag{3.29}$$

When N is even, we have two solutions, $k = 0, N/2 \pmod N$, and when $N(\geq 3)$ is odd, we have the unique solution, $k = 0 \pmod N$. It should be noted that these values of k are identical with the $\mathbb{Z}_N$ charges for singlet states at $\theta = 0$ that are calculated in Sec. 3.1.1.

At $\theta = \pi$, a nontrivial thing happens because the topological θ term also flips its sign under time reversal T. To see it, let us apply the T transformation to the θ term at $\theta = \pi$:

$$\begin{aligned} -\frac{\mathrm{i}\pi}{2\pi}\int (\mathrm{d}q + A) \quad &\mapsto \quad \frac{\mathrm{i}\pi}{2\pi}\int (\mathrm{d}q + A) \\ &= \quad -\frac{\mathrm{i}\pi}{2\pi}\int (\mathrm{d}q + A) + \mathrm{i}\int \mathrm{d}q + \mathrm{i}\int A. \end{aligned} \tag{3.30}$$

Two additional terms appear after the T transformation of the θ term; $\int \mathrm{d}q$ and $\int A$. $\int \mathrm{d}q$ does not play any role in the path integral, because $\mathrm{i}\int \mathrm{d}q \in 2\pi\mathrm{i}\mathbb{Z}$. Additional $\int A$ shifts the Chern-Simons level by 1. Combined with the flip of the Chern-Simons term, the T transformation of the partition function at $\theta = \pi$ is obtained as

$$\mathcal{Z}_{\pi,k}[\mathsf{T} \cdot (A, B)] = \mathcal{Z}_{\pi,k}[(A, B)] \exp\left(\mathrm{i}(2k-1)\int A\right). \tag{3.31}$$

In order to preserve the time reversal symmetry under background the $\mathbb{Z}_N$-gauge field, we must choose k, such that

$$2k - 1 = 0, \quad (\mathrm{mod}\ N). \tag{3.32}$$

For even N, the condition has no solution. The phase factor of (3.31) cannot be eliminated by local counter terms, and thus there is the mixed 't Hooft anomaly between $\mathbb{Z}_N$ and the time reversal symmetry. The anomaly matching claims that the ground state must be degenerate at $\theta = \pi$ when N is even. For odd $N(\geq 3)$, this has the solution $k = (N+1)/2$ modulo N, and no 't Hooft anomaly exists. It should again be noticed that this is same with the $\mathbb{Z}_N$ charge of the singlet state at $\theta = \pi$ as computed in Sec. 3.1.1.

For odd $N \geq 3$, there is a global inconsistency between $\theta = 0$ and $\theta = \pi$. To eliminate phases at $\theta = 0$ and $\theta = \pi$, the Chern-Simons level k should be chosen as

$$k_0 = 0, \quad k_\pi = \frac{N+1}{2}, \tag{3.33}$$

respectively. We cannot choose simultaneous k eliminating phases because $k_0 \neq k_\pi$ and k is the discrete parameter. To circumvent it, we need the 2-dimensional bulk Σ with $\partial\Sigma = S^1_\beta$ as in the case of the anomaly inflow, and then the bulk topological field theory,

$$S_{\mathrm{2d},\Sigma}[A] = \mathrm{i}\theta\frac{N+1}{2\pi}\int_\Sigma \mathrm{d}A, \tag{3.34}$$

can simultaneously eliminate the phases at $\theta = 0, \pi$ [23]. At $\theta = 0, \pi$, this topological action is independent of the choice of Σ unlike the case of 't Hooft anomaly, but it is not true for generic $0 < \theta < \pi$ and the information of the bulk Σ is necessary in order to connect $\theta = 0, \pi$.

3.1.3 Relation between two formalisms

The central extension in operator formalism and local counter terms resulted from gauging global symmetry in path integral formalism seemingly give same information about mixed anomaly and global inconsistency. Here, we show the connection by an explicit computation. We start with the path integral formalism (one can go the other way around) with the action (3.24). We fix a gauge by requiring $B = 0$ (mod 2π), and the equation of motion $\mathrm{d}B = NA$ is solved by

$$B = \sum_i 2\pi\ell_i\Theta(\tau - \tau_i), \quad A = \sum_i \frac{2\pi\ell_i}{N}\delta(\tau - \tau_i)\mathrm{d}\tau, \tag{3.35}$$

where $\Theta(\tau)$ is the step function and $\delta(\tau)$ is the delta function, for $\tau_i \in \mathbb{R}$ and $\ell_i \in \mathbb{Z}$. Let us calculate the partition function under this background $\mathbb{Z}_N$ gauge field,

$$\begin{aligned}
&\mathcal{Z}_{\theta,k}[(A,B)]\\
=&\int \mathcal{D}q\mathcal{D}p\exp\left[\int \mathrm{d}\tau\left(\mathrm{i}p(\dot{q}+A_0)-\frac{1}{2}\left(p-\frac{\theta}{2}\right)^2-V(Nq+B)-\mathrm{i}kA_0\right)\right]\\
=&\int \mathcal{D}q\mathcal{D}p\exp\left[\int \mathrm{d}\tau\Big(\mathrm{i}p\dot{q}-H(p,q)\Big)\right]\exp\left[\sum_i \frac{2\pi\mathrm{i}\ell_i}{N}(p(\tau_i)-k)\right]\\
=&\left\langle \prod_i \left(\mathrm{e}^{-2\pi\mathrm{i}k/N}\mathsf{U}(\tau_i)\right)^{\ell_i}\right\rangle. \qquad (3.36)
\end{aligned}$$

It can now be explicitly shown the relation between the commutation relation (3.10) and the phases in (3.28) and (3.31). Using the commutation relation, we get

$$\mathsf{T}\left(\mathrm{e}^{-2\pi\mathrm{i}k/N}\mathsf{U}\right)\mathsf{T}^{-1}=\begin{cases}\mathrm{e}^{2\pi\mathrm{i}(2k)/N}\left(\mathrm{e}^{-2\pi\mathrm{i}k/N}\mathsf{U}\right), & (\theta=0),\\ \mathrm{e}^{2\pi\mathrm{i}(2k-1)/N}\left(\mathrm{e}^{-2\pi\mathrm{i}k/N}\mathsf{U}\right), & (\theta=\pi).\end{cases} \qquad (3.37)$$

The T transformation acting on the right hand side of (3.36) gives the correct additional phases: At $\theta=0$, we get

$$\prod_{\mathrm{i}}\left(\mathrm{e}^{2\pi\mathrm{i}(2k)/N}\right)^{\ell_i}=\exp\left(2\mathrm{i}k\int A\right), \qquad (3.38)$$

and, at $\theta=\pi$, we get

$$\prod_{\mathrm{i}}\left(\mathrm{e}^{2\pi\mathrm{i}(2k-1)/N}\right)^{\ell_i}=\exp\left(\mathrm{i}(2k-1)\int A\right). \qquad (3.39)$$

We should emphasize that the phase factors which come from the local counter term are precisely same as those appear as a central extension.

3.2 Quantum mechanics of two particles on S^1

We consider the quantum mechanics with the target space $U(1)\times U(1)$ corresponding to two distinguishable particles moving on a ring with flux threading. We shall go through the parallel argument as we have done in the last section, but this model exhibits new ingredients and the global inconsistency plays a particularly important role. The Euclidean classical action is

$$S[q_1,q_2]=\int \mathrm{d}\tau\left[\frac{1}{2}(m_1\dot{q}_1^2+m_2\dot{q}_2^2)+V(q_1-q_2)\right]-\frac{\mathrm{i}\theta_1}{2\pi}\int \mathrm{d}q_1-\frac{\mathrm{i}\theta_2}{2\pi}\int \mathrm{d}q_2, \qquad (3.40)$$

where m_1 and m_2 are distinct mass parameters for each particle and the potential $V(x)$ is represented as the Fourier series (3.2), which is a smooth 2π periodic function. Each q_i $(i=1,2)$ is the map $q_i : S^1_\beta \to \mathbb{R}/2\pi\mathbb{Z}$. The theta parameters θ_i are 2π periodic variables.

With use of the path integral the partition function is expressed as

$$\mathcal{Z}_{(\theta_1,\theta_2)}=\int \mathcal{D}q_1\mathcal{D}q_2\exp(-S[q_1,q_2]). \qquad (3.41)$$

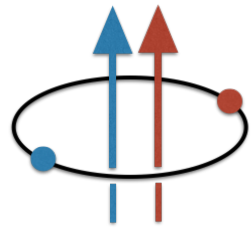


Figure 3.3: Two particles on a ring with fluxes.

In the operator formalism, the partition function is expressed as $\mathcal{Z}_{(\theta_1,\theta_2)} = \mathrm{tr}_{\mathcal{H}}[\exp(-\beta \hat{H})]$ with the hamiltonian given by

$$\hat{H}(\hat{p}_1, \hat{q}_1, \hat{p}_2, \hat{q}_2) = \frac{1}{2m_1}\left(\hat{p}_1 - \frac{\theta_1}{2\pi}\right)^2 + \frac{1}{2m_2}\left(\hat{p}_2 - \frac{\theta_2}{2\pi}\right)^2 + V(\hat{q}_1 - \hat{q}_2). \tag{3.42}$$

where $[\hat{q}_i, \hat{p}_j] = \mathrm{i}\delta_{ij}$.

3.2.1 Symmetries, central extension, and global inconsistency

The action (3.40) possesses $U(1)$ symmetry generated by

$$\mathsf{U}_\alpha : q_i(\tau) \mapsto q_i(\tau) + \alpha, \tag{3.43}$$

where $i = 1, 2$ and α is a 2π periodic constant, i.e., $\mathsf{U}_{2\pi} = 1$. The corresponding generator is given by $\mathsf{U}_\alpha = \mathrm{e}^{\mathrm{i}\alpha(\hat{p}_1+\hat{p}_2)}$ and satisfies a commutation relation $\mathsf{U}_\alpha \hat{H} \mathsf{U}_{-\alpha} = \hat{H}$. The time reversal transformation

$$\mathsf{T} : q_i(\tau) \mapsto q_i(-\tau),\ \dot{q}_i(\tau) \mapsto -\dot{q}_i(-\tau), \tag{3.44}$$

becomes an additional symmetry at $(\theta_1, \theta_2) = (0, 0), (0, \pi), (\pi, 0), (\pi, \pi)$, which are the high symmetry points of the model.

We analyze commutation relations of U_α and T to study the 't Hooft anomaly and global inconsistency. Let the high symmetry points be denoted by $(\theta_1, \theta_2) = (j_1\pi, j_2\pi)$ with $j_1, j_2 \in \mathbb{Z}$. The condition for the time-reversal symmetry, $\mathsf{T}\hat{H}\mathsf{T}^{-1} = \hat{H}$, combined with anti-unitarity $\mathsf{T}\mathrm{i}\mathsf{T} = -\mathrm{i}$ requires $\mathsf{T}\hat{q}_i\mathsf{T}^{-1} = \hat{q}_i$ $(i = 1, 2)$ and

$$\mathsf{T}\hat{p}_1\mathsf{T}^{-1} = -\hat{p}_1 + j_1, \quad \mathsf{T}\hat{p}_2\mathsf{T}^{-1} = -\hat{p}_2 + j_2, \quad (\theta_1, \theta_2) = (j_1\pi, j_2\pi). \tag{3.45}$$

Therefore, the commutation relations between U and T are,

$$\mathsf{T}\mathsf{U}_\alpha\mathsf{T}^{-1} = \mathrm{e}^{\mathrm{i}(j_1+j_2)\alpha}\mathsf{U}_\alpha, \quad (\theta_1, \theta_2) = (j_1\pi, j_2\pi). \tag{3.46}$$

At $(\theta_1, \theta_2) = (0, 0)$, we obtained expected relation from $U(1) \times \mathbb{Z}_2$ symmetry. At $(\theta_1, \theta_2) = (0, \pi)$ and $(\pi, 0)$, we have additional phase factor $\mathrm{e}^{\mathrm{i}\alpha}$. We again try to absorb it by redefining the operator $\mathsf{U}'_\alpha \equiv \mathrm{e}^{-\mathrm{i}\alpha/2}\mathsf{U}_\alpha$. But U'_α forces the periodicity of α to be extended to 4π. Thus, avoiding the central extension necessarily yields the double cover of $U(1) \times \mathbb{Z}_2$ and this is a symptom of a mixed 't Hooft anomaly. Although similar issue seems to appear at $(\theta_1, \theta_2) = (\pi, \pi)$ this is not true because the phase factor $\mathrm{e}^{2\mathrm{i}\alpha}$ can be absorbed by a redefinition $\mathsf{U}''_\alpha \equiv \mathrm{e}^{-\mathrm{i}\alpha}\hat{\mathcal{O}}_\alpha$ without extending periodicity of α. It is again noted that, although there is no central extension at $(\theta_1, \theta_2) = (0, 0)$ and (π, π) respectively, we cannot choose common operator U_α (or U''_α). This is the global inconsistency. So far, we found similar observations as those we saw in the last section.

An interesting thing happens at $(\theta_1,\theta_2)=(\pi,-\pi)$; $\mathsf{T}\hat{p}_1\mathsf{T}^{-1}=-\hat{p}_1+1$ and $\mathsf{T}\hat{p}_2\mathsf{T}^{-1}=-\hat{p}_2-1$ lead to a commutation relation

$$\mathsf{T}\mathsf{U}_\alpha\mathsf{T}^{-1}=\mathsf{U}_\alpha. \tag{3.47}$$

Hence, there is not a mixed 't Hooft anomaly and also a global inconsistency does not exist between $(0,0)$ and $(\pi,-\pi)$. A global inconsistency however exists between $(0,0)$ and (π,π). It is noted that the theory at $(\pi,-\pi)$ must show the same property as one at $(\theta_1,\theta_2)=(\pi,\pi)$ because θ_2 is 2π periodic parameter.[1] This observation is not a contradiction and yields an important constraint on the energy spectrum as we will see momentarily.

We explore the implication of the above argument to energy spectrum and phase diagram. The same argument as we gave in the last section results in the existence of degenerate state at $(\theta_1,\theta_2)=(0,\pi),(\pi,0)$ and we do not repeat here. Instead, we restrict our attention to $(\theta_1,\theta_2)=(0,0),(\pi,\pi),(\pi,-\pi)$. The simultaneous eigenstate of U_α and T would satisfy

$$\mathsf{U}_\alpha|\psi\rangle=\mathrm{e}^{\mathrm{i}\alpha k}|\psi\rangle,\ \mathsf{T}|\psi\rangle=\eta|\psi\rangle, \tag{3.48}$$

where $k\in\mathbb{Z}$ because $\mathsf{U}_{2\pi}=1$. Then, by using the commutation relations (3.46) and (3.47), the parallel discussion given in Sec. 3.1.1 leads to the following $U(1)$ transformation law of states,

$$\mathsf{U}_\alpha|\psi_{(0,0)}\rangle=|\psi_{(0,0)}\rangle,\ \mathsf{U}_\alpha|\psi_{(\pi,\pi)}\rangle=\mathrm{e}^{i\alpha}|\psi_{(\pi,\pi)}\rangle,\ \mathsf{U}_\alpha|\psi_{(\pi,-\pi)}\rangle=|\psi_{(\pi,-\pi)}\rangle, \tag{3.49}$$

at $(\theta_1,\theta_2)=(0,0),(\pi,\pi),(\pi,-\pi)$, respectively. Since $|\psi_{(\pi,\pi)}\rangle$ has different $U(1)$ charge from $|\psi_{(0,0)}\rangle$ and $|\psi_{(\pi,-\pi)}\rangle$, $|\psi_{(\pi,\pi)}\rangle$ cannot be continuously connected to the other two states at high symmetry points. In addition, (π,π) and $(\pi,-\pi)$ must be identified because θ_2 is a 2π periodic parameter as we mentioned before. The compatible consequence is that (π,π) and $(\pi,-\pi)$ are separated by a phase transition as shown in Fig. 3.5. Otherwise, T-invariant state at (π,π) would be connected to T-broken state at $(\pi,-\pi)$ without a level crossing, which contradicts to the fact that (π,π) and $(\pi,-\pi)$ must have identical energy spectra.

The above arguments are indeed checked by a explicit computation of the energy spectra with a specific potential

$$V(q_1-q_2)=\lambda\cos(q_1-q_2). \tag{3.50}$$

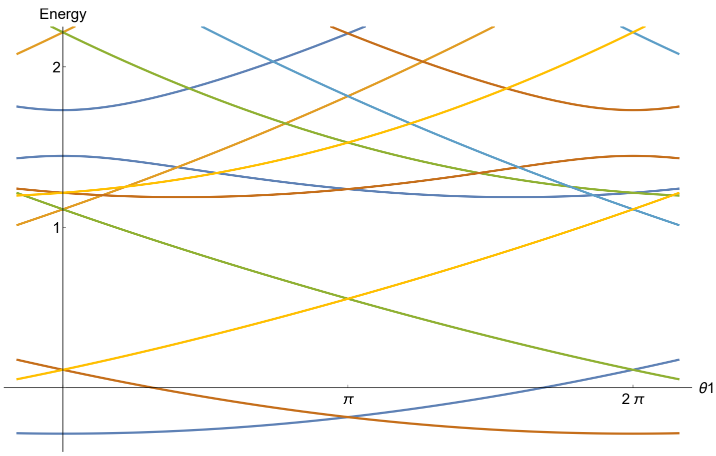
As shown in Fig. 3.4a, all the states at $(\theta_1,\theta_2)=(\pi,0)$ form pairs and the time reversal symmetry is spontaneously broken.

Fig. 3.4b shows the energy spectra as function of $\theta_1=\theta_2=\theta$. If a nondegenerate state exists at $\theta=0$, it is continuously connected to a degenerate state at $\theta=0$ (see the lowest blue curve in Fig. 3.4b, for instance) and vice versa (the lowest brown curve). Interestingly, the vacuum (lowest-energy) states are singlet both at $\theta=0$ and $\theta=\pi$, which is allowed because the level crossing (phase transition) separates these high symmetry points. The $U(1)$ charge of the lowest-energy state can jump at the crossing point since the points are not continuously connected by changing θ. This is the new ingredient which we did not see in the last section. Namely, the global inconsistency does not necessarily lead to the existence of degenerate vacuum at high symmetry points. Therefore, This result does not contradict to the fact that there is a global inconsistency between $(\theta_1,\theta_2)=(0,0)$ and (π,π).

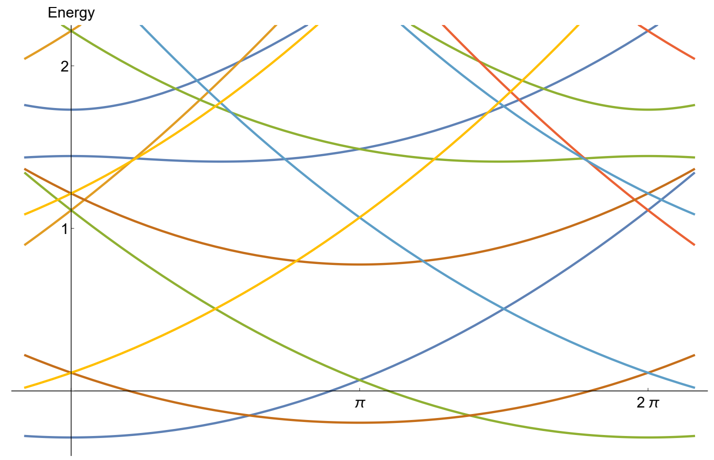
Finally, we see that the vacuum is nondegenerate for $\theta_1=-\theta_2=\theta'$ (Fig. 3.4c), which is consistent with the discussion in the last section that there is neither an 't Hooft anomaly nor global inconsistency at $\theta_1=-\theta_2=\pi$.

A phase diagram on (θ_1,θ_2)-plane (Fig. 3.5) follows the energy spectrum and level crossing computed above. As expected from the 't Hooft anomaly, Level crossing lines pass at $(0,\pm\pi)$ and

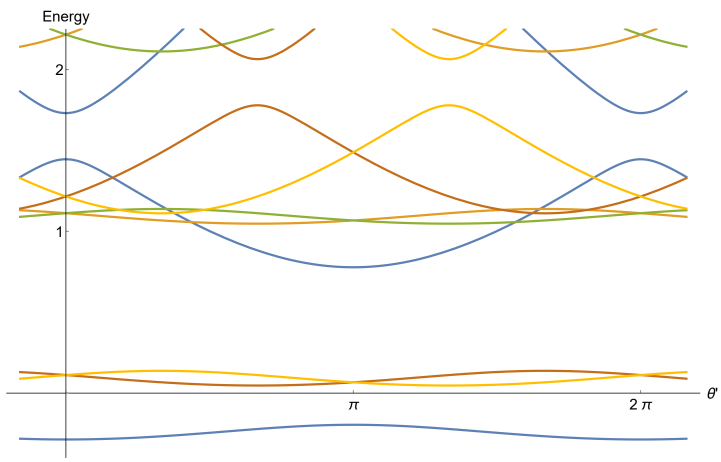
[1] Of course the same is true at $(\theta_1,\theta_2)=(-\pi,\pi)$ by using 2π periodicity of θ_1.



(a) $\theta_2 = 0$.



(b) $\theta = \theta_1 = \theta_2$.



(c) $\theta' = \theta_1 = -\theta_2$.

Figure 3.4: Energy spectra as functions of θ with $m_1 = 1$, $m_2 = 1/2$ and $\lambda = 1$. Color of lines indicates the $U(1)$ charge of states. (a) All the levels are degenerate at $(\theta_1, \theta_2) = (\pi, 0)$ due to the 't Hooft anomaly. (b) A singlet state at $(\theta_1, \theta_2) = (0, 0)$ must be connected to a degenerate state at $(\theta_1, \theta_2) = (\pi, \pi)$ and vice versa due to the global inconsistency. (c) Singlet states at $(\theta_1, \theta_2) = (0, 0)$ are connected to singlet states at $(\theta_1, \theta_2) = (\pi, -\pi)$. These figures are taken from Ref. [1].

$(\pm\pi, 0)$. High symmetry points $(\theta_1, \theta_2) = (0, 0)$ and $(\pi, -\pi)$ are connected without level crossing while $(0, 0)$ and (π, π) are separated by a level crossing line, which agrees with our consideration based on the global inconsistency.

Based on the constraints from the global inconsistency, we could come up with a little more exotic phase diagram which we did not find here. The other possibility we could draw from the global inconsistency between $(0, 0)$ and (π, π) is that the nondegenerate vacuum at $(0, 0)$ is connected to the degenerate vacua at (π, π) without level crossing. Then, there exists the degenerate vacua at $(\pi, -\pi)$ as well due to the 2π periodicity of θ_2. Therefore, the points $(0, 0)$ and $(\pi, -\pi)$ must be separated by another level crossing line because the singlet state at $(0, 0)$ cannot be connected to the T-broken state at $(\pi, -\pi)$ due to the absence of global inconsistency between these two points. See Chapter 6 for detailed discussion on $SU(N) \times SU(N)$ bifundamental gauge theory with two θ parameters corresponding to two gauge groups. In the theory, the almost same conditions are obtained by using global inconsistency and 't Hooft anomaly and two possible diagrams are proposed, and our phase

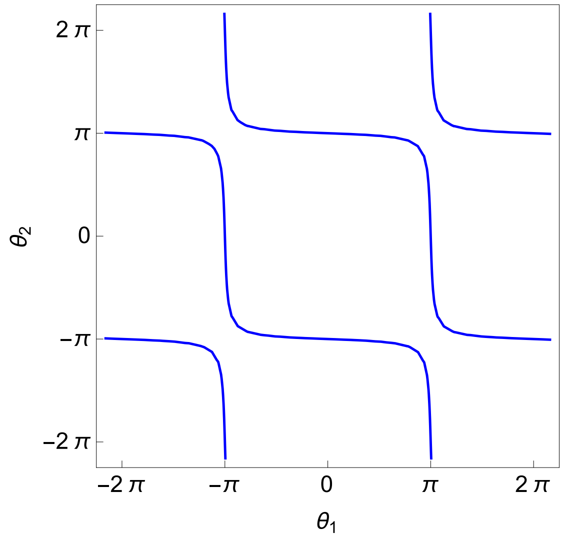


Figure 3.5: Phase diagram on (θ_1, θ_2)-plane with $\lambda = 0.2$. Each line represents a level crossing (phase transition). The phase structure is 2π periodic along θ_1 and θ_2 axises. This figure is taken from Ref. [1]

diagram Fig. 3.5 actually fits to one of the proposal made in [2] as we will discuss in Chapter 6.

3.2.2 Gauging $U(1)$ symmetry, 't Hooft anomaly, and global inconsistency

We promote the global $U(1)$ symmetry to the local gauge symmetry by coupling to the background $U(1)$ gauge field A in order to study the 't Hooft anomaly and global inconsistency for $U(1) \times \mathbb{Z}_2$ symmetry. To this end, we study the model (3.40) in the path integral formalism (3.41) as we have done in Sec. 3.1. The topological $U(1)$ gauge theory we need to couple here is

$$S_{\mathrm{top},k}[A] = \mathrm{i}k \int A, \tag{3.51}$$

which is $U(1)$ level-k Chern-Simons term in one dimension. The invariance under $U(1)$ gauge transformation

$$A \mapsto A + \mathrm{d}\lambda \tag{3.52}$$

requires the level to be an integer, $k \in \mathbb{Z}$. By postulating $U(1)$ gauge transformation,

$$q_1 \mapsto q_1 + \lambda, \;\; q_2 \mapsto q_2 + \lambda, \tag{3.53}$$

we obtain the gauge invariant action coupled to the topological gauge theory,

$$\begin{aligned} S[q_1, q_2, A] \;\; &= \;\; \int \mathrm{d}\tau \left[\frac{m_1}{2}(\dot{q}_1 + A_0)^2 + \frac{m_2}{2}(\dot{q}_2 + A_0)^2 + V(q_1 - q_2)\right] \\ &\quad - \frac{\mathrm{i}\theta_1}{2\pi}\int(\mathrm{d}q_1 + A) - \frac{\mathrm{i}\theta_2}{2\pi}\int(\mathrm{d}q_2 + A) + S_{\mathrm{top},k}[A]. \end{aligned} \tag{3.54}$$

Therefore, the partition function coupled to the background $U(1)$ gauge field is given by

$$\mathcal{Z}_{(\theta_1,\theta_2),k}[A] = \int \mathcal{D}q_1 \mathcal{D}q_2 \exp\left(-S[q_1, q_2, A]\right) \tag{3.55}$$

We will see how the partition function at high symmetry points transforms under time reversal operation,

$$q_1(\tau) \mapsto q_1(-\tau),\ q_2(\tau) \mapsto q_2(-\tau),\ A_0(\tau) \mapsto -A_0(-\tau). \tag{3.56}$$

At $(\theta_1, \theta_2) = (0, 0)$, the partition function transforms as

$$\mathcal{Z}_{(0,0),k}[\mathsf{T} \cdot A] = \mathcal{Z}_{(0,0),k}[A] \exp\left(2\mathrm{i}k \int A\right). \tag{3.57}$$

The time reversal invariance requires $k = 0$. Notice that the same transformation law holds at $(\theta_1, \theta_2) = (\pi, -\pi)$, which also results in $k = 0$.

As we saw in the last section, the transformation of the partition function at $(\theta_1, \theta_2) = (0, \pi)$

$$\mathcal{Z}_{(\pi,0),k}[\mathsf{T} \cdot A] = \mathcal{Z}_{(\pi,0),k}[A] \exp\left(\mathrm{i}(2k-1) \int A\right), \tag{3.58}$$

leads to a nontrivial consequence. The time reversal invariance requires $2k - 1 = 0$. Since this condition cannot be satisfied with integer k, the time reversal invariance cannot be preserved after gauging the $U(1)$ symmetry. Hence, an 't Hooft anomaly exists at $(\theta_1, \theta_2) = (0, \pi)$. Clearly, the same is true at $(\theta_1, \theta_2) = (\pi, 0)$.

Finally, at $(\theta_1, \theta_2) = (\pi, \pi)$, the partition function transforms as

$$\mathcal{Z}_{(\pi,\pi),k}[\mathsf{T} \cdot A] = \mathcal{Z}_{(\pi,\pi),k}[A] \exp\left(\mathrm{i}(2k-2) \int A\right). \tag{3.59}$$

In this case, the time reversal invariance is unbroken by choosing $k = 1$, meaning that there exists no mixed 't Hooft anomaly. By observing the resulting Chern-Simons levels at $(0, 0), (\pi, -\pi), (\pi, \pi)$

$$k_{(0,0)} = 0 = k_{(\pi,-\pi)},\ k_{(\pi,\pi)} = 1, \tag{3.60}$$

we conclude that there are global inconsistencies between $(0, 0)$ and (π, π), and between $(\pi, -\pi)$ and (π, π), respectively.

It is impossible to eliminate the phases coming out of 't Hooft anomalies and global inconsistencies by the local counter term, and we need the 2-dimensional bulk Σ with $\partial\Sigma = S^1_\beta$ to do it keeping the gauge invariance. The 2-dimensional topological action,

$$S_{\mathrm{2d},\Sigma}[A] = \mathrm{i}\frac{(\theta_1 + \theta_2)}{2\pi} \int_\Sigma \mathrm{d}A, \tag{3.61}$$

cancel additional phases of the partition function. At $(\theta_1, \theta_2) = (\pi, 0), (0, \pi)$, this topological action depends on the topology of Σ, and this detects the mixed 't Hooft anomaly. At $(\theta_1, \theta_2) = (\pi, \pi)$, this does not depend on the choice of Σ, but the information of the bulk is necessary to connect it with $(\theta_1, \theta_2) = (0, 0)$, and this is the signal for the global inconsistency.

3.3 More on $\mathbb{Z}_N \times \mathbb{Z}_2$ mixed anomaly

Finally, we briefly look at the model (3.40) with one-particle potentials $V(Nq_1)$ and $V(Nq_2)$ in addition to the inter-particle potential $V(q_1 - q_2)$, which are represented as Fourier series (3.2) with different sets of parameters. $V(Nq_1)$ and $V(Nq_2)$ explicitly break $U(1)$ symmetry down to $\mathbb{Z}_N$ symmetry, which is generated by

$$\mathsf{U} : q_1 \mapsto q_1 + \frac{2\pi}{N},\ q_2 \mapsto q_2(\tau) + \frac{2\pi}{N}. \tag{3.62}$$

The potentials $V(Nq_1)$, $V(Nq_2)$ change the conditions for the 't Hooft anomaly and global inconsistency at high symmetry points. Here, we do not repeat the operator formalism but present only the path integral formalism. To this end, we promote the global $\mathbb{Z}_N$ symmetry by following the procedure employed in Sec. 3.1. The topological gauge theory we need to couple is (3.22) by introducing $\mathbb{Z}_N$ one-form A and $U(1)$ zero-form gauge fields B with the constraint $NA = \mathrm{d}B$. The $\mathbb{Z}_N$ gauge transformation is given by (3.19) and

$$q_1 \mapsto q_1 - \lambda,\ q_2 \mapsto q_2 - \lambda. \tag{3.63}$$

The action invariant under the gauge transformation takes the following form,

$$\begin{aligned} &S_{(\theta_1,\theta_2),k}[q_1, q_2, A, B] \\ &= \int \mathrm{d}\tau \left[\frac{m_1}{2}(\dot{q}_1 + A_0)^2 + \frac{m_2}{2}(\dot{q}_2 + A_0)^2 + V(q_1 - q_2) + V(Nq_1 + B) + V(Nq_2 + B)\right] \\ &- \frac{\mathrm{i}\theta_1}{2\pi}\int(\mathrm{d}q_1 + A) - \frac{\mathrm{i}\theta_2}{2\pi}\int(\mathrm{d}q_2 + A) + S_{\mathrm{top},k}[A], \end{aligned} \tag{3.64}$$

Here, we list the condition for the discrete parameter k at each high symmetry points required by invariance under the time reversal symmetry:

$$\begin{cases} k = -k, & (\theta_1, \theta_2) = (0, 0), (\pi, -\pi) \\ k = -k + 1, & (\theta_1, \theta_2) = (\pi, 0), (0, \pi), \qquad (\mathrm{mod}\ N) \\ k = -k + 2, & (\theta_1, \theta_2) = (\pi, \pi). \end{cases} \tag{3.65}$$

These restrictions result in the following consequences: For odd $N \geq 3$, an 't Hooft anomaly does not exist at any high symmetry point. In this case, global inconsistencies exists among $(0,0)$, $(\pi,0)$, $(0,\pi)$ and (π,π) because

$$k_{(0,0)} = 0,\ k_{(\pi,0)} = \frac{N+1}{2} = k_{(0,\pi)},\ k_{(\pi,\pi)} = 1, \tag{3.66}$$

which take different values.

For even $N \geq 4$, 't Hooft anomalies appear at $(\theta_1, \theta_2) = (\pi, 0), (0, \pi)$ because there is no integer solution for k, i.e., the gauge invariance cannot be maintained. Although there is no mixed anomaly at $(\theta_1, \theta_2) = (0, 0), (\pi, \pi), (\pi, -\pi)$, a global inconsistency exists between $(0, 0)$ and (π, π) and between $(\pi, -\pi)$ and (π, π) because

$$k_{(0,0)} = 0 = k_{(\pi,-\pi)},\ k_{(\pi,\pi)} = 1. \tag{3.67}$$

In $N = 2$ case, the first and third conditions in (3.65) are equivalent mod N. Hence, there is no global inconsistency although we still have 't Hooft anomalies at $(\theta_1, \theta_2) = (\pi, 0), (0, \pi)$.

3.4 Summary

We have illustrated the nature of the global inconsistency as well as the 't Hooft anomaly and their implication on energy spectra by looking at quantum mechanical models. Let us recall the 't Hooft anomaly and global inconsistency once again. The 't Hooft anomaly shows up as an obstruction to gauging a global G symmetry of the system and inevitably leads to nontrivial infrared theories. The global inconsistency has similar nature in that it also appears as an obstruction to gauging symmetry and imposes constraints on the low-energy theory. The global inconsistency, however, plays a role in more restricted situations, where there exist high symmetry points connected each other by continuous parameters of the theory. The constraints obtained from the global inconsistency is milder than those from the 't Hooft anomaly due to the fact that it does not necessarily rule out the realization of trivial vacuum at high symmetry points. When there is a global inconsistency between two high symmetry points, one can draw a constraint that the vacuum is nontrivial at either of the points, or that those two points are separated by a phase transition.

We carefully analyzed quantum mechanical models which exhibits 't Hooft anomalies and global inconsistencies at high symmetry points of the parameter space spanned by theta parameters. We studied them by the operator formalism and path-integral formalism. In the operator formalism, by studying central extensions of the symmetry group, one can tell how (non-accidental) level crossings appears in energy spectrum. In the path-integral formalism, 't Hooft anomalies and global inconsistencies are detected by gauging a global symmetry as we discussed in Sec. 2.2. We then established a precise connection between these two formalisms in the quantum mechanical models, which allows us to predict the level crossing in energy spectra by studying 't Hooft anomalies and global inconsistency. It is noted that the 't Hooft anomaly matching argument constraints only vacuum property of the QFT because of the assumption on locality of low-energy effective theories. However, they becomes more restrictive in quantum mechanics and one can extract the information on excited states as well by combining the observations drawn from the central extension of symmetry groups.

More specifically, we analyzed the following quantum mechanical models in detail: In the model describing a particle on a ring, the symmetry group is $\mathbb{Z}_N \times \mathbb{Z}_2$ at high symmetry points $\theta = 0$ and π. There is a mixed anomaly at $\theta = \pi$ for even N, and a global inconsistency between $\theta = 0$ and π for odd N. This model is a reminiscent of $SU(N)$ pure Yang-Mills model at $\theta = \pi$ with $\mathbb{Z}_N$ one-form center symmetry and time reversal symmetry. The second model with two particles is a reminiscent of $SU(N) \times SU(N)$ gauge theory with bifundamental matters. There are mixed anomalies at $(\theta_1, \theta_2) = (0, \pi), (\pi, 0)$. The global inconsistency appears between $(\theta_1, \theta_2) = (0, 0)$ and (π, π) but not between $(\theta_1, \theta_2) = (0, 0)$ and $(\pi, -\pi)$ which indeed agrees with the phase diagram for this model in (θ_1, θ_2) plane. The interesting observation which was absent in the first model is that the global inconsistency does not imply the existence of degenerate vacua at $(\theta_1, \theta_2) = (\pi, \pi)$. Instead the high symmetry points, $(\theta_1, \theta_2) = (0, 0)$ and (π, π), are separated by a level crossing line in the (θ_1, θ_2) space.

Chapter 4

Generalized global symmetry

As we have already illustrated in simple quantum mechanical models in the last chapter, for the purpose of detecting an 't Hooft anomaly and global inconsistency we need to specify the global symmetry and then gauge it. The procedure is more or less the same in the QFTs that we will encounter in the rest of this dissertation except that involved symmetries and its gauging are gradually getting complicated. This chapter serves as preliminaries to deal with those symmetries. We pay a primarily attention to so called the generalized global symmetry or higher form symmetry, that acts on gauge invariant extended objects as opposed to the conventional symmetry acting on point-like objects [15–17]. As we will see in the following chapters, the higher form symmetry that we are interested in is the center symmetry and happens to be a discrete symmetry as well. Gauging such symmetries requires some technicalities involving the topological quantum field theory. After reviewing necessary ingredients in nonabelian gauge theories, we introduce the generalized global symmetry followed by demonstrating how to gauge discrete symmetries by making use of the topological quantum field theory.

4.1 Electric and magnetic charge in nonabelian theories

Spectra of particles with electric or/and magnetic charges are constrained by the locality of QFT. In case of nonabelian gauge theories, the spectrum of charged objects are closely related to its gauge group, and sometimes it is crucial to specify the global symmetry group. Some basic facts about the electric and magnetic charges in gauge theories are reviewed in this section.

4.1.1 Dirac quantization condition

Let us start with the $U(1)$ Maxwell theory, in which electrons carry conserved electric charge. More generally, one may think of a particle carrying magnetic charge or both electric and magnetic charge, namely, monopole and dyon, respectively. The Dirac quantization condition states that, given two dyons with electric and magnetic charges given by (e_1, g_1) and (e_2, g_2), they satisfy the "mutual locality" in the quantum theory if

$$e_1 g_2 - e_2 g_1 \in 2\pi\mathbb{Z}, \tag{4.1}$$

which is called the Dirac quantization condition [73–76]. The mutual locality means that there is no correlation between two particles which are infinitely separated. In the topological quantum field theory, since there is no local excitation, two separated objects do not correlate if they are mutually local.

Let us take an electron of charge $(e, 0)$ and a magnetic monopole of charge $(0, g)$. Then, the Dirac quantization condition requires $eg \in 2\pi\mathbb{Z}$, which also holds for an electron and a dyon. Now, suppose the minimum of magnetic charge is $2\pi/e$ due to the existence of electron of charge e, and think about two dyons of charge $(e_1, 2\pi/e)$ and $(e_2, 2\pi/e)$. According to the Dirac quantization condition, $(e_1 - e_2)/e \in \mathbb{Z}$ is required. If we additionally require the time reversal symmetry (or equivalently CP), their electric charges are restricted to $e_1 = -e_2$, leading to $2e_1/e \in \mathbb{Z}$. Hence, the dyon charge has to be quantized by integer or half-integer.

On the other hand, if the time reversal symmetry is broken by finite θ angle described by the action

$$S_\theta = \frac{\mathrm{i}\theta}{8\pi^2} \int F \wedge F, \tag{4.2}$$

in four dimensional spacetime, the electric charge of a dyon of charge (e, g) takes the value $q = e - \theta e/2\pi$, which is not necessarily quantized. This is called Witten effect [77].

4.1.2 $\mathfrak{su}(N)$ gauge theory

We discuss all the possible electric and magnetic charges especially for the gauge group $G = SU(N)$ and $G = SU(N)/\mathbb{Z}_N$ following [78]. To this end, we need to elaborate the representations of each group and algebra. For slightly more general settings, let $\widehat{G} = SU(N)$ be the universal cover of the gauge group, and the gauge group is given by $G = \widehat{G}/H$ with a center subgroup $H \subset \mathbb{Z}_N$. We will give a physical interpretation after a rather abstract description.

The group G determines what kinds of matter fields are allowed to exists, namely, they may exist if they are in the representations of G. While not all dynamical matters in these representations necessarily exist, all the Wilson line operators are required to exist corresponding to the representations of the group. They are labeled by the weight lattice of G modulo the Weyl group, Λ_{w}/W. Therefore, the line operator spectra specify the theory in terms of gauge groups.[1] Now, we loosely identify these points on Λ_{w}/W with "electric charges" of matter fields or the Wilson line operators. We may define the "magnetic charge" in the same way by means of the GNO dual gauge group $G^\vee$ [79], whose universal cover $\widehat{G^\vee}$ has the same center group as the original group $\widehat{G}$. The magnetic charge is labeled by the magnetic weight lattice of $\mathfrak{g}^\vee$, that is GNO dual Lie algebra of $\mathfrak{g}$, modulo the Weyl group Λ^m_{w}/W. Generically, one can consider dyonic matters or line operators carrying both electric and magnetic charge, which are labeled by $(\Lambda_{\mathrm{w}} \times \Lambda^m_{\mathrm{w}})/W$. Finally, as we will physically explain momentarily it is useful to organize these matter fields or line operators in terms of the pair of lattices $(\Lambda_{\mathrm{w}} \times \Lambda^m_{\mathrm{w}})/W$ modulo the root lattices of $\mathfrak{g} \times \mathfrak{g}^\vee$, which is the center of the universal cover group. They are labeled by (z_e, z_m) and we actually call them electric and magnetic charge in nonabelian G gauge theory in the following.

How does this charge assignment physically make sense? Let us consider in $G = SU(N)$ gauge theory. Classically, all the electric charges belong to representations of the Lie algebra $\mathfrak{su}(N)$, i.e., elements of the weight lattice with the identification under the Weyl group. After quantization, each electric charge can emit and absorb gluons that are in the adjoint representation and thus the electric charge z_e is labeled only by classical electric charge modulo N, i.e., $z_e \in \mathbb{Z}_N$, which is the center group of $SU(N)$. That is why we identified the weight lattice under the root lattices which represent gluons. The same is true for the magnetic charge. Hence, the dyonic charge is labeled by

$$(z_e, z_m) \in \mathbb{Z}_N \times \mathbb{Z}_N. \tag{4.3}$$

[1] It turns out that theories of same gauge group can have different line operator spectra, which further distinguish those theories [78].

Now, the difference between abelian and nonabelian theories are rather clear. In the Maxwell theory, the photons do not carry $U(1)$ electric charge and radiative corrections do not affect the charge of electrons. On the other hand, gluons carry color charge, which may be attributed to the nonlinearity of the nonaberian theory, and hence the quarks of charge 0 and N cannot be distinguished.

What is the physical meaning of the gauge group $G = SU(N)/H$? When the gauge group is $G = SU(N)/H$, the dynamical matters as well as the Wilson line operators must be invariant under $H \subset \mathbb{Z}_N$ and only such $z_e \in \mathbb{Z}_N$ are allowed. When $G = SU(N)$, the allowed electric charges are $z_e = 0, 1, \ldots, N-1$. For $G = SU(N)/\mathbb{Z}_N$, the only allowed electric charge is $z_e = 0$ because the particle of charge $k \in \mathbb{Z}_N$ acquires a phase $\mathrm{e}^{\mathrm{i}k/N}$ under the $\mathbb{Z}_N$ transformation,[2] meaning that such particle is not invariant under $H = \mathbb{Z}_N$ unless $k = 0$.

Finally, as we discussed earlier, for particles being genuine point-like objects, the set of allowed charges must satisfy Dirac quantization condition: For (z_e, z_m) and (z'_e, z'_m), they must satisfy

$$(z_e z'_m - z'_e z_m) = 0 \mod N. \tag{4.4}$$

When $G = SU(N)$, the only allowed magnetic charge is $z_m = 0$. For $G = SU(N)/\mathbb{Z}_N$, the allowed magnetic charges are $z_m = 0, \ldots, N-1$.

All these discussions combined together specify allowed electric and magnetic charges. In particular, all the line operators allowed by the discussions based on the group representations and the Dirac quantization conditions must exist, whose spectra in turn characterize distinct gauge theories.

4.1.3 Genuine line operators

So far in this section we have seen the spectra of line operators are highly constrained by the Dirac quantization condition, which requires the mutual locality among line operators. We call the operators that satisfy the condition the genuine line operators. On the other hand, the violation of the mutual locality can be attributed to the existence of surfaces attached to those lines and they still provide important information as we will see later. Those surfaces pick up a nontrivial phase when two line operators which are not mutually local link each other [16]. To distinguish those line operators it is useful to introduce the following three classes [16, 78]:

1. A line operator which bounds a surface operator. We need to specify the location of both line and surface.

2. A line operator which bounds a surface operator as well as the previous case, but correlation functions only depends on the topological class of the surface.

3. A line operator with no surface need to be attatched, which are called genuine line operators.

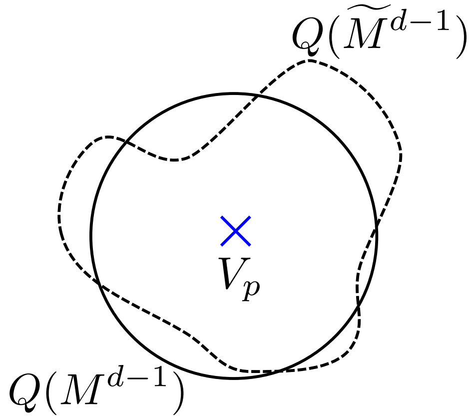
The generalization to higher dimensional operators is similarly done.

4.2 Higher form global symmetry

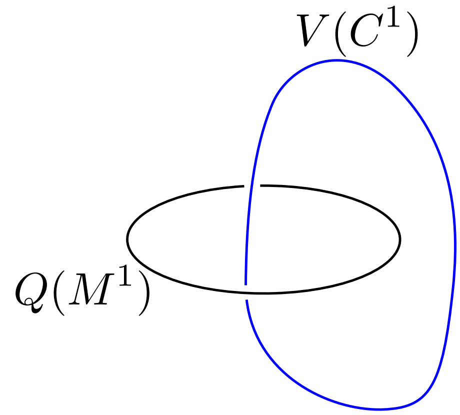
Let us start with recalling the basic properties of ordinary global symmetries in d dimensional spacetime, which faithfully act on the Hilbert space [17] (see also section 2 of [80]).[3]

[2] We will elaborate the precise meaning of the $\mathbb{Z}_N$ transformation after introducing the one-form symmetry

[3] We do not see the gauge symmetry as symmetry because it "trivially" acts on Hilbert space, which is, roughly specking, spanned by gauge invariant operators.



(a) Zero form symmetry in d-dimensional spacetime.



(b) One form symmetry in 3-dimensional spacetime.

Figure 4.1: Schematic figures describing charge operators and charged objects associated to zero-form and one-form symmetry. (a) V_p on the blue point is a point-like object charged under the ordinary (zero-form) symmetry. The charge operator $Q(M^{d-1})$ represented by the black solid line, that is $(d-1)$-dimensional surface in d-dimensional spacetime, measures the charge inside the surface M^{d-1}. The dashed line also represents the charge operator $Q(\widetilde{M}^{d-1})$. Since the measured charge does not depend on the choice of surface unless it crosses the charged objects, the surface is topological. (b) The blue line $V(C^1)$ is a line-object charged under the one-form symmetry. The charge operator $Q(M^1)$ represented by the black solid line, that is 1-dimensional cycle in 3-dimensional spacetime, measures the charge of the line operator $V(C^1)$ linking with M^1.

For a continuous symmetry group G the Noether's theorem provides a conserved charge by the spatial integral of a conserved current,

$$Q(M^{d-1}) = \int_{M^{d-1}} j \tag{4.5}$$

where the current j is given by $(d-1)$-form. Then, the generator of the symmetry transformation is obtained by exponentiating the charge,

$$U_g(M^{d-1}) = \mathrm{e}^{\mathrm{i}\alpha Q(M^{d-1})}, \tag{4.6}$$

where $g \in G$ and α is a transformation parameter. The generator acts on operators $V(p)$ charged under the symmetry in the following way,

$$U_g(M^{d-1})V_i(p)(U_g(M^{d-1}))^{-1} = R_{ij}V_j(p), \tag{4.7}$$

with $p \in M^{d-1}$ and the right-hand side is proportional to the intersection number between p and M^{d-1}. More generally, this action can be expressed as

$$U_g(S^{d-1})V_i(p) = R_{ij}V_j(p), \tag{4.8}$$

where S^{d-1} is a sphere surrounding the point p (see Figure 4.1a). This is understood as follows: The product of operators should be regarded as the time-ordered product with some choice of a time-slice as a result of canonical quantization. Thus, the product in (4.7) can be understood by taking the time slices of $U_g(M^{d-1})$ and $(U_g(M^{d-1}))^{-1}$ at $t = t_p + \epsilon$ and $t = t_p - \epsilon$ respectively where t_p is the time of the point p and ϵ is a infinitesimal constant. Then, we obtain the sphere surrounding p by deforming two surfaces M^{d-1}, which leads to (4.8). The description of global symmetry based on the action of generator on charged objects (4.8) holds for discrete symmetry groups, too.

Next, we generalize the discussion to the symmetry associated with the generator acting on q-dimensional charged object, that is so called q-form global symmetry [15, 17]. The necessary ingredients to discuss q-form symmetry are the following: The transformation parameter α is given by a closed q-form, i.e., $\mathrm{d}\alpha = 0$. The conserved charge is defined by an integration of $(d-q-1)$-form current, $Q(M^{d-q-1}) = \int_{M^{d-q-1}} j$ for continuous symmetries. The generator is given by $U_g(M^{d-q-1})$ supported on a manifold M^{d-q-1} and it acts on q-dimensional charged objects $V(C^q)$ (see Figure 4.1b). With a manifold S^{d-q-1} linking C^q the action of generator is expressed as

$$U_g(S^{d-q-1})V(C^q) = RV(C^q). \tag{4.9}$$

One-form symmetries in Maxwell theory

Let us take a look at the four-dimensional Maxwell theory as an example possessing one-form symmetries, whose action is $S = \int F \wedge *F$. The equation of motion and the Bianchi identity are respectively

$$\mathrm{d} * F = 0, \quad \mathrm{d}F = 0, \tag{4.10}$$

without matter fields. Thus, $*F$ and F may be regarded as two-form conserved currents, to which we associate symmetry generators,[4]

$$U_\mathrm{e}(M^2) = \exp\left(\mathrm{i}\frac{\alpha_\mathrm{e}}{2\pi}\int_{M^2} *F\right), \quad U_\mathrm{m}(M^2) = \exp\left(\mathrm{i}\frac{\alpha_\mathrm{m}}{2\pi}\int_{M^2} F\right). \tag{4.13}$$

The charged objects under the symmetry generated by $U_\mathrm{e}(M^2)$ are the Wilson loops

$$W(C^1) = \exp\left(\mathrm{i}\frac{n}{2\pi}\int_{C^1} A\right), \tag{4.14}$$

with charge n, satisfying the equal-time commutator

$$U_\mathrm{e}(S^2)W(C^1) = \mathrm{e}^{\mathrm{i}\alpha_\mathrm{e} n\ell}W(C^1), \tag{4.15}$$

where ℓ is a linking number between S^2 and C^1. Physically, the generator measures the electric flux carried by the Wilson loop. Hence, the symmetry is called electric one-form symmetry. It is explicitly broken if charged matter fields exist. For the same reason, the symmetry generated by $U_\mathrm{m}(M^2)$ is called magnetic one-form symmetry and its charged objects are the 't Hooft loops

$$T(C^1) = \exp\left(\mathrm{i}\frac{m}{2\pi}\int_{C^1} \widehat{A}\right), \tag{4.16}$$

with a magnetic charge m and magnetic gauge potential $\widehat{A}$ locally satisfying $\mathrm{d}\widehat{A} = *F$.

[4]Provided an open surface M^2 bounded by ∂M^2, does the generator $U_\mathrm{m}(M^2)$ become a genuine Wilson line operator? If so it would be expressed as

$$U_\mathrm{m}(M^2) = U_\mathrm{e}(\partial M^2) = \exp\left(\mathrm{i}\frac{\alpha_\mathrm{m}}{2\pi}\oint_{\partial M^2} *A\right). \tag{4.11}$$

However, the answer is NO because it is not invariant under the gauge transformation $A \to A + \mathrm{d}\lambda$,

$$U_\mathrm{e}(\partial M^2) \to U_\mathrm{e}(\partial M^2)\exp\left(\mathrm{i}\alpha_\mathrm{m} k\right), \tag{4.12}$$

where we used $\oint \mathrm{d}\lambda \in 2\pi k$ with $k \in \mathbb{Z}$. Thus, the surface bounding the loop is needed. The same is true for $U_\mathrm{e}(M^2)$, which cannot be a genuine 't Hooft line operator.

4.3 Topological quantum theory

When we discuss gapped topological states, there is no local excitation in the infrared spectrum. Such systems are described by means of topological quantum field theories (TQFT), which is roughly defined by the QFT independent of geometry, i.e., it does not depend on metrics and hence the energy-momentum tensor vanishes. Even though there is no local degrees of freedom, such states are distinguished from trivially gapped states by the existence of nonlocal operators such as line or higher dimensional surface operators and their nontrivial statistics. In this section, we will briefly review the abelian Chern-Simons theory and the BF theory from the viewpoint of global symmetries. The BF theory plays an important role for gauging discrete symmetries as we will discuss in the next section.

4.3.1 Abelian Chern-Simons theory

The 3d Chern-Simons theory might be the most familiar TQFT for physicists. We restrict ourselves to the $U(1)$ Chern-Simons theory for the sake of simplicity. The action is

$$S_{\mathrm{CS}} = \frac{\mathrm{i}k}{4\pi}\int A\wedge \mathrm{d}A. \tag{4.17}$$

It is noted that there is no metric in this action.[5] Under the gauge transformation $A\to A+\mathrm{d}\alpha$[6], the action changes as

$$\Delta S_{\mathrm{CS}} = \frac{\mathrm{i}k}{2\pi}\int \mathrm{d}\alpha\wedge \mathrm{d}A \in 2\pi\mathrm{i}k\mathbb{Z}. \tag{4.18}$$

Therefore, the level k must be an integer, leading to the invariance of the partition function. This can be explicitly checked on a spin manifold $S^1\times S^2$. Another way to understand the level quantization is to define the Chern-Simons action

$$S_{\mathrm{CS}} = \frac{k}{4\pi}\int_{M^4} F\wedge F, \tag{4.19}$$

by choosing an extension to a bounding four-manifold M^4. Suppose we define the Chern-Simons action on M^3 by extending to M^4 and $\tilde{M}^4$ whose boundary is M^3, the level quantization is derived by requiring that the definition does not depend on the extension,

$$\frac{k}{4\pi}\int_{M^4} F\wedge F - \frac{k}{4\pi}\int_{\tilde{M}^4} F\wedge F \in 2\pi\mathbb{Z}. \tag{4.20}$$

The difference can be rewritten as

$$\frac{k}{4\pi}\int_{M^4-\tilde{M}^4} F\wedge F \in 2\pi k\mathbb{Z}, \tag{4.21}$$

where $M^4-\tilde{M}^4$ is a closed four manifold and we used the index theorem $\frac{1}{8\pi^2}\int F\wedge F\in 2\pi\mathbb{Z}$, which is valid on spin manifolds. Hence, $k\in\mathbb{Z}$ allow us to define the Chern-Simons action without referring to four manifolds.

The theory is invariant under the transformation

$$A\to A+\frac{\lambda}{k}, \tag{4.22}$$

[5] There is a subtle geometric dependence through the framing anomaly, but we do not get into detail.

[6] This is a large gauge transformation with the gauge parameter α satisfying $\oint \mathrm{d}\alpha\in 2\pi\mathbb{Z}$.

with a $U(1)$ gauge field λ satisfying $\mathrm{d}\lambda = 0$, i.e., λ is flat, and this is a one-form $\mathbb{Z}_k$ transformation. By repeating it k times we get the 0-form gauge transformation (4.17). The $\mathbb{Z}_k$ one-form global symmetry is generated by the Wilson loops

$$W_m(M^1) = \exp\left(\mathrm{i}m \oint_{M^1} A\right), \tag{4.23}$$

which are charged objects, too.

$$W_m(M^1)W_n(N^1) = \exp\left(2\pi\mathrm{i}\frac{mn}{k}\right) W_n(N^1), \tag{4.24}$$

for the linked loops M^1 and N^1.

4.3.2 BF theory

The 4d BF theory is described by the action

$$S_{BF} = \frac{\mathrm{i}N}{2\pi} \int B \wedge \mathrm{d}A, \tag{4.25}$$

where B is a two-form gauge field and A is a one-form gauge field. With k being an integer, the action is invariant mod 2π under the 0-form gauge transformation $A \to A + \mathrm{d}\alpha_0$ and the one-form gauge transformation $B \to B + \mathrm{d}\alpha_1$. Accordingly, the theory possesses one-form and two-form global $\mathbb{Z}_N$ symmetries generated by

$$T_{m_1}(M^2) = \exp\left(\mathrm{i}m_1 \oint_{M^2} B\right), \tag{4.26}$$

$$W_{m_2}(M^1) = \exp\left(\mathrm{i}m_2 \oint_{M^1} A\right), \tag{4.27}$$

by which the fields[7] are transformed as

$$A \to A + \frac{1}{N}\lambda_1, \tag{4.28}$$

$$B \to B + \frac{1}{N}\lambda_2, \tag{4.29}$$

respectively, with properly normalized flat one-form gauge field λ_1 and two-form gauge field λ_2.

One may add extra topological term to the BF action (4.25),

$$S^p_{BF} = \frac{\mathrm{i}N}{2\pi} \int B \wedge \mathrm{d}A + \frac{\mathrm{i}Np}{4\pi} \int B \wedge B, \tag{4.30}$$

with an additional parameter p. Since the added term plays an important role in the discussion of the 't Hooft anomaly in the following chapters, we here see the property of p and gauge symmetry of the theory. If we set $p \in \mathbb{Z}$, the action is invariant under one-form gauge transformation,

$$B \to B + \mathrm{d}\alpha_1, \quad A \to A - p\alpha_1. \tag{4.31}$$

[7] Corresponding gauge invariant charged objects are $W_{n_1}(N^1)$ and $T_{n_2}(N^2)$, resepectively.

We can dualize the field A by introducing a Lagrange multiplier $\widehat{A}$ in the action (4.30),

$$\begin{aligned}
S^p_{BF} &\sim \frac{\mathrm{i}}{2\pi}\int (NB + \mathrm{d}\widehat{A})\wedge F + \frac{\mathrm{i}Np}{4\pi}\int B\wedge B \\
&= \frac{\mathrm{i}}{2\pi}\int (NB + \mathrm{d}\widehat{A})\wedge F + \frac{\mathrm{i}p}{4\pi N}\int \mathrm{d}\widehat{A}\wedge \mathrm{d}\widehat{A}.
\end{aligned} \tag{4.32}$$

Here " $\sim$" means that, we get $\mathrm{d}F = 0$ by integrating out $\widehat{A}$, leading to $F = \mathrm{d}A$, and end up with the original form. Now, the dualized action is invariant under the one-form gauge transformation

$$B \to B + \mathrm{d}\alpha_1, \quad \widehat{A} \to \widehat{A} - N\alpha_1 \quad F \to F - p\mathrm{d}\alpha_1. \tag{4.33}$$

From the last term of the action it is clear that $p \sim p + N$ because $\frac{1}{4\pi}\int \mathrm{d}\widehat{A}\wedge\mathrm{d}\widehat{A} \in 2\pi\mathbb{Z}$. Therefore, the added parameter is set to be $p \in \mathbb{Z}_N$.

The symmetry and line operator spectrum of this theory are carefully analyzed in the next chapter.

4.4 Gauging $\mathbb{Z}_N$ symmetry

In the last chapter, we encountered the situation where we need to gauge the $\mathbb{Z}_N$ symmetry to study the 't Hooft anomaly. In the next chapter, we will again face to the similar problem in $SU(N)$ Yang-Mills theory, i.e., gauging the center symmetry, which is $\mathbb{Z}_N$ one-form symmetry. We develop the technique to gauge those discrete symmetries in this section. We consider the discrete symmetry which can be embedded in a continuous symmetry.

One way to gauge the $\mathbb{Z}_N$ symmetry of the theory with an action S is to enforce a $U(1)$ gauge symmetry and then break it to the $\mathbb{Z}_p$ subgroup by the Higgs mechanism [81, 82]. More specifically, we first enlarge the $\mathbb{Z}_N$ symmetry to the $U(1)$ symmetry, which is possible in the cases we are interested in, and gauge the symmetry by introducing $U(1)$ gauge field A. Then, we couple to the theory that contains the Higgs field H enabling us to break $U(1)$ to $\mathbb{Z}_N$ gauge group. The total $\mathbb{Z}_N$-gauged action is

$$S_{\text{total}} = S[A] + S_{\text{Higgs}}[A, H]. \tag{4.34}$$

As we will see below, $S_{\text{Higgs}}[A, H]$ is described by the BF theory if we take low energy limit.

Let us consider a $U(1)$ gauge theory with a gauge field A and a Higgs field H carrying charge N under $U(1)$ gauge group. Then, the $U(1)$ gauge transformation of the charge N Higgs field is realized by $\phi \to \mathrm{e}^{\mathrm{i}N\alpha}\phi$ with a gauge parameter α. After condensing the Higgs field, the nonvanishing vacuum expectation value $\langle H\rangle \neq 0$ breaks the gauge group $U(1)$ to $\mathbb{Z}_N$ whose elements are given by $g = e^{2\pi\mathrm{i}k/N}$ $(k = 0, \cdots, N-1)$.

The Lagrangian describing the Higgs mechanism for $U(1)$ gauge theory is given by

$$\mathcal{L} = t^2(\mathrm{d}\phi - NA)\wedge *(\mathrm{d}\phi - NA) + \frac{1}{2e^2}\mathrm{d}A\wedge *\mathrm{d}A. \tag{4.35}$$

with the phase ϕ of the Higgs field $H = |H|\mathrm{e}^{\mathrm{i}\phi}$. ϕ is a 2π periodic scalar, $\phi \sim \phi + 2\pi$. We will see that the low energy limit of the theory realized by $t \to \infty$ is topological $\mathbb{Z}_N$ gauge theory. To this end, we dualize the scalar ϕ. We view $C = \mathrm{d}\phi$ as independent field by introducing a two-form B, which plays a role of Lagrange multiplier,

$$\begin{aligned}
&t^2(C-NA)\wedge *(C-pA)+\frac{\mathrm{i}}{2\pi}C\wedge \mathrm{d}B+\frac{1}{2e^2}\mathrm{d}A\wedge *\mathrm{d}A\\
&=t^2(C-NA-\frac{\mathrm{i}}{4\pi t^2}*\mathrm{d}B)\wedge *(C-NA+\frac{\mathrm{i}}{4\pi t^2}*\mathrm{d}B)\\
&+\frac{\mathrm{i}N}{2\pi}B\wedge \mathrm{d}A+\frac{1}{(4\pi t)^2}\mathrm{d}B\wedge *\mathrm{d}B+\frac{1}{2e^2}\mathrm{d}A\wedge *\mathrm{d}A. \qquad (4.36)
\end{aligned}$$

By integrating out C, we obtain

$$\frac{\mathrm{i}N}{2\pi}B\wedge \mathrm{d}A+\frac{1}{(4\pi t)^2}\mathrm{d}B\wedge *\mathrm{d}B+\frac{1}{2e^2}\mathrm{d}A\wedge *\mathrm{d}A. \qquad (4.37)$$

In the $t\to\infty$ limit, the second term is gone, the equation of motion locally leads to a flat connection $\mathrm{d}A=0$, and thus we obtain the BF theory,

$$\mathcal{L}_{BF}=\frac{\mathrm{i}N}{2\pi}B\wedge \mathrm{d}A, \qquad (4.38)$$

which is a topological $\mathbb{Z}_N$ gauge theory. In fact, the equations of motion $\mathrm{d}A=0=\mathrm{d}B$ mean that there exists no local degree of freedom.

While the Higgs Lagrangian (4.35) allows us to understand the origin of the 0-form $\mathbb{Z}_N$ gauge symmetry, it does not explain the one-form $\mathbb{Z}_N$ gauge symmetry, that shows up in the BF theory as well as the 0-form symmetry according to the discussion in the last section. To figure this out, we dualize A in (4.37) by introducing an independent field F and a Lagrange multiplier V,

$$\begin{aligned}
&\frac{\mathrm{i}N}{2\pi}B\wedge F+\frac{1}{(4\pi t)^2}\mathrm{d}B\wedge *\mathrm{d}B+\frac{1}{2e^2}F\wedge *F+\frac{\mathrm{i}}{2\pi}\mathrm{d}V\wedge F\\
&=\frac{1}{(4\pi t)^2}\mathrm{d}B\wedge *\mathrm{d}B+\frac{e^2}{8\pi^2}\left(\mathrm{d}V+NB\right)\wedge *\left(\mathrm{d}V+NB\right)\\
&+\frac{1}{2e^2}\left(*F+\frac{\mathrm{i}e^2}{2\pi}\mathrm{d}V+\frac{\mathrm{i}e^2N}{2\pi}B\right)\wedge *\left(*F+\frac{\mathrm{i}e^2}{2\pi}\mathrm{d}V+\frac{\mathrm{i}e^2N}{2\pi}B\right). \qquad (4.39)
\end{aligned}$$

Integrating out F we obtain

$$\frac{1}{(4\pi t)^2}\mathrm{d}B\wedge *\mathrm{d}B+\frac{e^2}{8\pi^2}\left(\mathrm{d}V+NB\right)\wedge *\left(\mathrm{d}V+NB\right). \qquad (4.40)$$

The matter field is given by one-form (vector), which has charge N under one-form gauge group with two-form gauge field B. Therefore, this dual Lagrangian describes the Higgs mechanism breaking $U(1)$ one-form gauge group down to $\mathbb{Z}_N$ one-form gauge group. The BF theory can be understood to emerge as a low energy limit of the Higgs Langrangians (4.35) and (4.40).

It is noted that the BF Lagrangian can be further dualized. Dualizing B results in

$$-\frac{\mathrm{i}N}{2\pi}H\wedge A+\frac{\mathrm{i}}{2\pi}H\wedge \mathrm{d}\varphi=\frac{\mathrm{i}}{2\pi}H\wedge(\mathrm{d}\varphi-NA), \qquad (4.41)$$

with only 0-form gauge symmetry is remained. This is of course directly obtained from the original Higgs Lagrangian (4.35) by taking the low energy limit $t\to\infty$, which forces $\mathrm{d}\phi-NA=0$. This is precisely reproduced by integrating out H and identifying φ with ϕ (4.41). On the other hand, dualizing A results in

$$\frac{\mathrm{i}N}{2\pi}B\wedge F+\frac{\mathrm{i}}{2\pi}\mathrm{d}\widehat{A}\wedge F=\frac{\mathrm{i}}{2\pi}(NB+\mathrm{d}\widehat{A})\wedge F. \qquad (4.42)$$

with one-form gauge symmetry is remained. We notice that there is also an emergent 0-form $U(1)$ gauge invariance under $\widehat{A} \to \widehat{A} + \mathrm{d}\widehat{\alpha}_0$.

Finally, in order to convince ourselves that the one-form $\mathbb{Z}_N$ symmetry is indeed gauged by coupling a target theory to the BF theory, we have a look at the coupled partition function,

$$\mathcal{Z}_{\text{total}}[B] = \int DA \exp\left[\frac{\mathrm{i}N}{2\pi}\int_M B \wedge \mathrm{d}A\right] \mathcal{Z}[B]. \tag{4.43}$$

The target theory is described by the partition function $\mathcal{Z}[B]$ and we assume that it has a one-form $U(1)$ gauge symmetry with a background two-form gauge field B. The closed two-form $\mathrm{d}A/2\pi \in H^2[M,\mathbb{Z}]$ is decomposed as

$$\mathrm{d}A = \mathrm{d}\alpha + \sum_i m_i \omega_i, \tag{4.44}$$

with a one-form α, integers m_i, and harmonic two-forms ω_i.[8] By substituting this expression, we obtain the partition function,

$$\begin{aligned}
&\mathcal{Z}_{\text{total}}[B] \\
&= \int DA \exp\left[\frac{\mathrm{i}N}{2\pi}\int_M B \wedge \mathrm{d}\alpha\right] \exp\left[\sum_i \frac{\mathrm{i}Nm_i}{2\pi}\int_M B \wedge \omega_i\right] \mathcal{Z}[B] \\
&= \int D\alpha \exp\left[-\frac{\mathrm{i}N}{2\pi}\int_M \mathrm{d}B \wedge \alpha\right] \prod_i \exp\left[\sum_{m_i} \frac{\mathrm{i}Nm_i}{2\pi}\int_{\mathcal{P}[\omega_i]} B\right] \mathcal{Z}[B].
\end{aligned} \tag{4.45}$$

where $\mathcal{P}[\omega_i]$ is a two-cycle that is Poincare dual to a harmonic two-form ω_i. The integration over α leads to $N\mathrm{d}B = 0$, and sums over m_i result in the constraint on the holonomy $\int_{\mathcal{P}[\omega_i]} B \in (2\pi/N)\mathbb{Z}$. Consequently, the two-form gauge fields B is restricted to a flat $\mathbb{Z}_N$ gauge field, i.e., $NB/2\pi \in H^2[M,\mathbb{Z}_N]$. Actually, this result is immediately obtained from (4.42) by integrating out F, which yields $NB = -\mathrm{d}\widehat{A}$, and hence the properties stated above are quite obvious. Gauging 0-form $\mathbb{Z}_N$ symmetry is similarly done by integrating out two-form gauge field B instead of A in (4.43).

Another way to check whether the discrete symmetry is gauged is by comparing the line operator spectrum before and after gauging, which will be done in the next chapter.

[8] A is generally not a globally defined one-form on M, but $\mathrm{d}A$ is a globally defined closed two-form, which is not exact if A is not globally defined one-form. α appearing in the decomposition is a globally defined one-form.

Chapter 5

’t Hooft anomaly and global inconsistency in $SU(N)$ pure Yang-Mills theory

Dynamics of non-Abelian gauge theories depends not only on the gauge coupling constant but also on the topological θ angle. Since its discovery, the dependence of vacua and excitations on the parameter θ has been a key issue to understand the topological nature of gauge theories [83–88]. The strongly interacting sector in the Standard Model of particle physics is the $SU(3)$ vector-like gauge theory, and thus all the interactions preserve CP invariance (or equivalently time-reversal invariance due to the CPT theorem) except for this topological term. It is widely believed that the θ angle of QCD is quite small because CP symmetry in the strong sector is well maintained in our universe according to the experiment on neutron’s electric dipole moment [89].

For four-dimensional $SU(N)$ Yang-Mills theory, the angle θ is periodic in 2π, and thus the requirement of time reversal invariance of theory raises two candidates: $\theta = 0$ and $\theta = \pi$. Understandings on the vacuum structure at $\theta = \pi$ are of particular importance, and many studies have been devoted to it using various techniques, including large N limit, effective models, and chiral perturbation [51–63, 90–92]. In certain limits (for example large N), one can show that $SU(N)$ Yang-Mills theory possesses the first-order phase transition at $\theta = \pi$ and breaks time reversal symmetry spontaneously. This tells us that physics at $\theta = \pi$ is dramatically different from that at $\theta = 0$, and it is not known what would happen in generic cases. Recently, in Ref. [21], a new technique has been developed for $SU(N)$ pure Yang-Mills theory and also for $SU(N)$ Yang-Mills theory with adjoint matter fields, which gives a rigorous constraint on the vacuum structure at $\theta = \pi$ by discussing the ’t Hooft anomaly matching. More interestingly, it reveals under reasonable assumptions that the first-order phase transition at $\theta = \pi$ survives at finite temperatures at least until the deconfinement transition happens.

In this chapter, we give a review on four-dimensional Yang-Mills theory with the gauge Lie algebra $\mathfrak{su}(N)$ based on the preparation made in Chapter 4. We focus on the case where the gauge group is $SU(N)$ or $PSU(N) = SU(N)/\mathbb{Z}_N$. We mainly follow our paper [2], where the subject presented in this chapter is given as a review of Ref. [21]. For more details of this subject, see e.g. Refs. [16, 17, 78, 93].

5.1 $SU(N)$ Yang-Mills theory and its genuine line operators

The four-dimensional $SU(N)$ pure Yang-Mills theory is described by,

$$S = -\frac{1}{2g^2}\int \mathrm{Tr}(G \wedge *G) + \frac{\mathrm{i}\theta}{8\pi^2}\int \mathrm{Tr}(G \wedge G), \tag{5.1}$$

where G is the field strength of the $SU(N)$ gauge field a:

$$G = \mathrm{d}a + \mathrm{i}a \wedge a. \tag{5.2}$$

$a = a_{i\mu}T^i \mathrm{d}x^\mu$ is locally an $n \times n$ Hermitian matrix-valued one-form, and $\mathrm{Tr}(T^iT^j) = \frac{1}{2}\delta^{ij}$. The theory is invariant under the $SU(N)$ gauge transformation

$$a \mapsto a^g = gag^{-1} - \mathrm{i}g\mathrm{d}g^{-1}, \tag{5.3}$$

and the physical observables must respect the gauge invariance.

Let us recall the center symmetry in $SU(N)$ Yang-Mills theory, which plays an central role in the following discussion. Since all the fields in the theory are invariant under the center $\mathbb{Z}_N$ of the gauge group $SU(N)$, the gauge transformation is determined only up to elements of $\mathbb{Z}_N$ [94].[1] Given a nontrivial cycle C parametrized by an angle θ with $0 \le \theta < \pi$, $g \in SU(N)$ is not single-valued but may satisfy

$$g(2\pi) = \mathrm{e}^{2\pi\mathrm{i}/N} g(0). \tag{5.4}$$

This has an interesting consequence on line operators: The Wilson line in the fundamental representation along a closed line C is a gauge invariant object,

$$W(C) = \mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a\right)\right]. \tag{5.5}$$

It transforms under the gauge transformation (5.3) satisfying (5.4) as

$$\begin{aligned} W^g(C) &= \mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a^g\right)\right] = \mathrm{Tr}\left[\prod_{i=0}^{N-1}\exp\left(\mathrm{i}\int_{x_i}^{x_{i+1}} a^g\right)\right] \\ &= \mathrm{Tr}\left[\prod_{i=0}^{N-1}(1 + \mathrm{i}\Delta x a^g + \mathcal{O}(\Delta x^2))\right] \\ &= \mathrm{Tr}\left[\prod_{i=0}^{N-1} g(1 + \mathrm{i}\Delta x a)(g^{-1} + \Delta x \mathrm{d}g^{-1}) + \mathcal{O}(\Delta x^2)\right] \\ &= \mathrm{Tr}\left[\prod_{i=0}^{N-1} g(x_i)\exp\left(\mathrm{i}\int_{x_i}^{x_{i+1}} a\right) g^{-1}(x_{i+1})\right] \\ &= \mathrm{Tr}\left[g(x_0)\mathcal{P}\exp\left(\mathrm{i}\oint_C a\right) g^{-1}(x_N)\right] \\ &= \mathrm{e}^{-2\pi\mathrm{i}/N} W(C), \end{aligned} \tag{5.6}$$

where $\Delta x = x_{i+1} - x_i$ and $x_N = x_0$. We have used (5.4) in the last equality. Hence, this transformation acts on the Wilson line by the center $\mathbb{Z}_N$ and is called the center transformation. It should

[1] This is certainly not true if fermions in the fundamental representations exist. But the same argument works out in the case with fermions in the adjoint representation because they are invariant under $\mathbb{Z}_N \subset SU(N)$. The theories with fundamental fermions will be discussed in Chapter 6 and 7.

be emphasized that the center symmetry associated to this transformation is a global symmetry because it acts on the Wilson loops, which are gauge invariant. Since this symmetry transformation acts on line objects, it can be interpreted as a one-form symmetry associated to the transformation of a fundamental field

$$a \mapsto a + \frac{1}{N}\epsilon, \tag{5.7}$$

with $U(1)$ one-form gauge field ϵ satisfying $\oint \epsilon \in 2\pi\mathbb{Z}$ [17]. One can measure the electric charge z_e of the Wilson line by introducing a topological surface operator [17, 95]. In this sense, $SU(N)$ Yang-Mills theory has a global center symmetry that is called electric $\mathbb{Z}_N$ one-form symmetry (see Chapter 4).

Since theory has fundamental Wilson lines in the spectrum of genuine line operators as mentioned above, there is no magnetic line operator as a genuine line object. Indeed, let (z_e, z_m) be a charge of the line operator, then the Dirac quantization condition with the fundamental Wilson line with charge $(1, 0)$ claims

$$z_m = 0 \mod N. \tag{5.8}$$

Thus, there is no magnetic or dyonic genuine line. The genuine line operators with different electric charges are given by $W(C)^k$ with $z_e = k = 0, 1, \ldots, N-1$.

5.2 $SU(N)/\mathbb{Z}_N$ Yang-Mills theory

Let us next consider the $SU(N)/\mathbb{Z}_N$ gauge theory, and the general argument on the electric charge shows that the purely electric line operators must be invariant under $\mathbb{Z}_N$, such as $W(C)^N$. Since the Dirac quantization condition with allowed electric particles does not give any constraint on z_m, the genuine line with $z_m = 1$ is possible. Let us assume that we have a theory with a magnetic or dyonic line with charge $(z_e, z_m) = (-p, 1)$ with some $p = 0, 1, \ldots, N-1$. The Dirac quantization says that the charge (z'_e, z'_m) of other genuine line operators must satisfy

$$z'_e = -p z'_m \mod N. \tag{5.9}$$

Therefore, the electric charge of line operators with $z_m = 1$ is fixed to $-p$ once the line with $(z_e, z_m) = (-p, 1)$ exists. p is a new parameter of $SU(N)/\mathbb{Z}_N$ gauge theories, which is called the discrete theta angle [16, 17, 78, 93], and it specifies the spectrum of genuine line operators.

We can construct $SU(N)/\mathbb{Z}_N$ Yang-Mills theory by coupling $SU(N)$ Yang-Mills theory (5.1) to the following $\mathbb{Z}_N$ topological field theory [16],

$$S_{\mathrm{TFT}} = \frac{\mathrm{i}}{2\pi}\int F \wedge (\mathrm{d}A + NB) + \frac{\mathrm{i}Np}{4\pi}\int B \wedge B. \tag{5.10}$$

This topological field theory is a low-energy effective description of the spontaneous (one-form) gauge symmetry breaking $U(1) \to \mathbb{Z}_N$ when the fields with charge N are condensed [82]. Here, A and B are one-form and two-form $U(1)$ gauge fields, respectively, and F is a two-form auxiliary field (see Chapter 4 for gauging the center symmetry). We require that the action is invariant under the one-form $U(1)$ gauge transformation,

$$A \mapsto A - N\lambda,\ B \mapsto B + \mathrm{d}\lambda,\ F \mapsto F - p\,\mathrm{d}\lambda, \tag{5.11}$$

and then p must be an integer due to the gauge invariance[2].

[2]We implicitly assume that we consider field theories only on spin manifolds

In order to couple $SU(N)$ Yang-Mills theory to $\mathbb{Z}_N$ topological field theory (5.10), we first extend the gauge group from $SU(N)$ to $U(N) = (SU(N) \times U(1))/\mathbb{Z}_N$, and identify this $U(1)$ factor with that of the $U(1)$ gauge field A in (5.10). Correspondingly, the $SU(N)$ gauge field a is replaced by the $U(N)$ gauge field,

$$\mathcal{A} = a + \frac{1}{N} A \mathbf{1}_N, \tag{5.12}$$

and the gauge field strength becomes

$$\mathcal{G} = \mathrm{d}\mathcal{A} + \mathrm{i}\mathcal{A} \wedge \mathcal{A}. \tag{5.13}$$

Under the $U(1)$ one-form gauge transformation, $\mathcal{G}$ is transformed as

$$\mathcal{G} \mapsto \mathcal{G} - \mathrm{d}\lambda \mathbf{1}_N. \tag{5.14}$$

In order to obtain the $SU(N)/\mathbb{Z}_N$ gauge theory instead of $U(N)$ gauge theory, we postulate the invariance under the $U(1)$ one-form gauge transformation, and then the gauge invariant combination is given by $\mathcal{G} + B\mathbf{1}_N$ (for notational simplicity, the identity matrix $\mathbf{1}_N$ will be omitted below). As a result, the classical action for the $SU(N)/\mathbb{Z}_N$ Yang-Mills theory is given by

$$\begin{aligned} S &= -\frac{1}{2g^2} \int \mathrm{Tr}((\mathcal{G} + B) \wedge *(\mathcal{G} + B)) + \frac{\mathrm{i}\theta}{8\pi^2} \int \mathrm{Tr}((\mathcal{G} + B) \wedge (\mathcal{G} + B)) \\ &\quad + \frac{\mathrm{i}}{2\pi} \int F \wedge (\mathrm{d}A + NB) + \frac{\mathrm{i}Np}{4\pi} \int B \wedge B. \end{aligned} \tag{5.15}$$

Locally, we obtain $B = -\frac{1}{N}\mathrm{d}A$ by integrating out F, and its substitution recovers the original $SU(N)$ Yang-Mills action but this operation is ill-defined globally. Spectrum of local operators on topologically trivial manifolds is unchanged by this gauging procedure, but there is a crucial difference on nontrivial topologies or with non-local operators as we shall see below.

If one tries to define the Wilson line by (5.5), it is not gauge invariant under the $U(N)$ gauge transformation. We can define two kinds of gauge-invariant line operators with $(z_e, z_m) = (1, 0)$ and $(0, 1)$ but they need a topological surface ($\partial\Sigma = C$) in general in order to maintain the $U(N)$ 0-form and $U(1)$ 1-form gauge invariance:

$$\begin{aligned} W(C, \Sigma) &= \mathrm{Tr}\left[\mathcal{P} \exp\left(\mathrm{i} \oint_C a\right)\right] \exp\left[\frac{\mathrm{i}}{N} \oint_C A + \mathrm{i} \int_\Sigma B\right], &(5.16) \\ H(C, \Sigma) &= \exp\left[\mathrm{i} \int_\Sigma (F + pB)\right]. &(5.17) \end{aligned}$$

We now claim that the action (5.15) indeed describes the Yang-Mills theory of gauge group $SU(N)/\mathbb{Z}_N$ with the discrete theta angle p. Indeed, let us consider $H(C, \Sigma)W(C, \Sigma)^{-p}$ that has charge $(z_e, z_m) = (-p, 1)$:

$$H(C, \Sigma)W(C, \Sigma)^{-p} = \exp\left(\mathrm{i} \int_\Sigma F\right) \left(\mathrm{Tr}\left[\mathcal{P} \exp\left(\mathrm{i} \oint_C a\right)\right] \exp\left[\frac{\mathrm{i}}{N} \oint_C A\right]\right)^{-p}. \tag{5.18}$$

Superficially, it depends on the surface Σ, but the equation of motion of A claims that $\frac{1}{2\pi}F \in H^2(X, \mathbb{Z})$, and thus $\exp(\int_\Sigma F)$ does not depend on the choice of surfaces Σ satisfying $\partial\Sigma = C$. Therefore, the theory (5.15) have the dyonic genuine line operator with charge $(z_e, z_m) = (-p, 1)$, which is concretely given by $H(C, \Sigma)W(C, \Sigma)^{-p}$. This is precisely the spectrum of genuine line operators in the $SU(N)/\mathbb{Z}_N$ gauge theory.

Let us also explain why p is called the discrete theta angle. For this purpose, we consider the shift $\theta \mapsto \theta + 2\pi$. The change of the action (5.15) under this shift is given by

$$\Delta S = \frac{\mathrm{i}}{4\pi} \int \mathrm{Tr}((\mathcal{G} + B) \wedge (\mathcal{G} + B)) = \frac{\mathrm{i}}{4\pi} \int \mathrm{Tr}(\mathcal{G} \wedge \mathcal{G}) - \frac{\mathrm{i}n}{4\pi} \int B \wedge B. \tag{5.19}$$

The first term is in $2\pi\mathrm{i}\mathbb{Z}$ on spin four-manifolds due to the index theorem, and thus the 2π shift of θ changes p to $p - 1$ (mod N). This is a consequence of the fact that the electric charge of dyons is shifted by $\theta/2\pi$ because of the θ angle, i.e., the Witten effect [77]. As a result, the periodicity of θ is extended to $2\pi N$ from 2π. Since N different choices of p for the $SU(N)/\mathbb{Z}_N$ gauge theory is related by 2π shifts of θ, p is called the discrete theta angle, although this is not always true for other gauge groups [78].

5.3 Spontaneous time reversal symmetry breaking at $\theta = \pi$ of $SU(N)$ Yang-Mills theory

We see how one can claim the spontaneous breaking of time reversal symmetry T at $\theta = \pi$ following the procedure with use of an 't Hooft anomaly, which was recently developed in Ref. [21]. We assume that $SU(N)$ Yang-Mills theory at $\theta = 0$ is trivially gapped with unbroken T, and also that the first-order phase transition does not happen at any $0 < \theta < \pi$. Let us couple the theory to background $\mathbb{Z}_N$ two-form gauge fields B as we have done in Sec. 5.2.

Even after this coupling, T must be still unbroken by choosing appropriate p at $\theta = 0$. If T is broken after gauging the $\mathbb{Z}_N$ one-form symmetry, then this means that there is a mixed 't Hooft anomaly between the T symmetry and $\mathbb{Z}_N$ one-form symmetry. Since an 't Hooft anomaly is renormalization group invariant [9, 17], there must be a certain degree of freedom carrying the same anomaly and surviving in the infrared limit. The assumption on the trivially gapped state claims that there is no such degree of freedom, and thus there must be a way to couple to the $\mathbb{Z}_N$ two-form gauge field B without breaking T.

Note that the T transformation flips the sign of the $\int B \wedge B$ term, and effectively p is mapped to $-p$ under the T transformation. Accordingly, the partition function changes as

$$\mathcal{Z} \mapsto \mathcal{Z} \exp\left[-2p \frac{\mathrm{i}N}{4\pi} \int B \wedge B\right]. \tag{5.20}$$

Therefore, above discussion claims that we can choose the discrete theta angle satisfying[3]

$$2p = 0 \mod N \tag{5.21}$$

in order not to break T at $\theta = 0$. Since this has a solution (e.g., $p = 0$ mod N is always a solution), the assumption on the gap and unbroken T at $\theta = 0$ is consistent.

Let us discuss the fate of T symmetry at $\theta = \pi$. In the $SU(N)$ Yang-Mills theory, $\theta = \pi$ is T-invariant because T flips $\theta = \pi$ to $\theta = -\pi$ and one can shift θ to $\theta + 2\pi$. After considering the coupling to the $\mathbb{Z}_N$ two-form gauge field B, this procedure changes p to $-p - 1$ because T flips p to $-p$ and the 2π shift of θ changes $-p$ to $-p - 1$ due to the Witten effect. The variation of the partition function is given by

$$\mathcal{Z} \mapsto \mathcal{Z} \exp\left[(-2p - 1) \frac{\mathrm{i}N}{4\pi} \int B \wedge B\right]. \tag{5.22}$$

[3] The condition derived here is different from and weaker than that given in Ref. [21] since we only consider theories on spin manifolds while they consider theories on non-spin manifolds as well as spin ones. Nevertheless, it does not affect the consequence about the fate of T symmetry at $\theta = \pi$.

In order not to break T due to the coupling to B at $\theta = \pi$, we must choose the discrete theta angle satisfying

$$2p + 1 = 0 \mod N, \tag{5.23}$$

but this is inconsistent with our choice at $\theta = 0$. If N is even, $2p + 1 = 0$ (mod N) does not have any integer solution, and thus there is a mixed 't Hooft anomaly and the anomalous phase factor $\mathcal{A}$ is given by

$$\mathcal{A} = -(2p+1)\frac{N}{4\pi}\int B \wedge B. \tag{5.24}$$

This claims that T is broken because of the $\mathbb{Z}_N$ gauge fields. For odd N, the condition $2p + 1 = 0$ (mod N) can be solved by putting $p = (N-1)/2$, and thus there is no 't Hooft anomaly. We are tempted to conclude that T invariance is preserved in this case. However, $p = (N-1)/2$ is not the T-invariant choice at $\theta = 0$; the theory is not T-invariant at $\theta = \pi$ since we have assumed that there is no first-order phase transition at $0 < \theta < \pi$. This is a consequence of the global inconsistency, namely, there is no choice of p such that the anomalous phase factors at $\theta = 0$ and π are simultaneously removed. Therefore, T is broken after coupling $SU(N)$ Yang-Mills theory at $\theta = \pi$ to $\mathbb{Z}_N$ background two-form gauge fields for any choice of the discrete theta angle p preserving T at $\theta = 0$.

For consistency, there must be some low-energy degrees of freedom in the $SU(N)$ Yang-Mills theory at $\theta = \pi$ that explains the T breaking after coupling it to $\mathbb{Z}_N$ two-form gauge fields [21]. There are several possible candidates for this:

- The vacua are trivially gapped but degenerate. Each of them breaks T spontaneously.
- The vacuum is gapped with unbroken T symmetry but described by a nontrivial topological field theory.
- The theory contains massless excitations.

If one further assumes or proves that the gap does not close at finite θ and the theory does not show the topological phase transition, T is broken spontaneously and there is a first-order phase transition at $\theta = \pi$. This interesting discussion given in Ref. [21] does not rely on any specific microscopic details, and thus the consequence is very general as long as the theory has the $\mathbb{Z}_N$ one-form symmetry (i.e., matters are in the adjoint representation) and satisfies the assumption about the mass gap or topological excitations.

5.4 Finite temperature

A natural question one may next ask is whether the 't Hooft anomaly and global inconsistency are robust against thermal fluctuation. In order to answer the question we perform circle compactification in the imaginary time direction with radius β, i.e., $M^4 \to M^3 \times S^1$ where S^1 is parametrized by τ. In this section, we consider the high temperature phase, where the inverse temperature β is much smaller than the dynamical scale in Yang-Mills theory, and study the three dimensional theory on M^3 obtained by the dimensional reduction [21, 24, 26, 96]. After the circle compactification, the one-form $\mathbb{Z}_N$ symmetry on M^4 splits into the one-form $\mathbb{Z}_N$ symmetry on M^3 acting on the Wilson loops on M^3 and the zero-form $\mathbb{Z}_N$ symmetry acting on the Polyakov loops wrapping S^1,

$$\mathrm{tr}\left[\mathcal{P}\exp\left(\mathrm{i}\int_{M^4} a\right)\right] \xrightarrow{\text{compactify}} \begin{cases} \mathrm{tr}\left[\mathcal{P}\exp\left(\mathrm{i}\int_{M^3} a\right)\right], \\ \mathrm{tr}\left[\mathcal{P}\exp\left(\mathrm{i}\int_{S^1} a\right)\right]. \end{cases} \tag{5.25}$$

It is noted that from three dimensional point of view the Polyakov loops are no longer extended objects but point-like objects, and hence, charged under the zero-form symmetry.

To study the dimensional reduction of the 't Hooft anomaly and global inconsistency, we introduce a background two-form gauge field $B^{(2)}$ for the one-form symmetry and background one-form gauge field $B^{(1)}$ for the zero-form symmetry, neither of which depends on τ. If we regard B as the two-form gauge field on four-dimensional manifold $M^3 \times S^1$, it may be expressed in terms of gauge fields on M^3 as

$$B = B^{(2)} + B^{(1)} \wedge \beta^{-1} \mathrm{d}\tau. \tag{5.26}$$

Now, the anomaly term on M^3 can be obtained simply by plugging (5.26) into four-dimensional anomaly term (5.24),

$$-(2p+1)\frac{N}{4\pi}\int_{M^4} B \wedge B \xrightarrow{\text{compactify}} -(2p+1)\frac{N}{2\pi}\int_{M^3} B^{(2)} \wedge B^{(1)}. \tag{5.27}$$

Hence, this term reproduce the same 't Hooft anomaly and global inconsistency as those in four-dimensional theory.

A crucial observation is that upon the circle compactification the Polyakov loop wrapping around S^1 is stable because it cannot be unwind without changing the topology of spacetime, which in turn leads to the reduced zero-form center symmetry. Therefore, the anomalies involving higher-form symmetry is unlikely to be wiped out at finite temperatures. In other words, they are robust against thermal fluctuations. It is not generically true that an 't Hooft anomaly or global inconsistency for zero-form global symmetry survives after circle compactification because point-like objects are localized on M^3 if the radius of circle is much smaller than the diameter of M^3 and gauge fields cannot be introduced in the compactified direction unless it carries nontrivial topology via twisted boundary condition.

A consequence of the 't Hooft anomaly (5.26) is that at least one of the involved symmetries has to be broken at any temperature [21]. Particularly, the center symmetry being broken at higher temperature and the T symmetry being broken at lower temperature, their transition temperatures T_{deconf} and T_{T} must satisfy the relation $T_{\mathrm{decomf}} \leq T_{\mathsf{T}}$.

We will further discuss the anomalies and circle compactification in QCD in Chapter 7, where the involvement of two-form gauge field plays a key role although there is no higher-form global symmetry in the original theory.

Chapter 6

Vacuum structure of $SU(N)\times SU(N)$ bifundamental gauge theories

The purpose of this chapter is to extend and apply the technique developed in the last chapter to give a rigorous constraint on the vacuum structure of $SU(N)\times SU(N)$ Yang-Mills theory with bifundamental matter fields at finite topological angles, which was reported in our work [2]. A detailed analysis of the symmetries and 't Hooft anomalies in the bifundamental gauge theory will help us understand those in QCD because $SU(N)\times SU(N)$ can be regarded as color and flavor gauge groups upon introducing flavor background gauge field in QCD to study the 't Hooft anomaly involving flavor symmetry. We will discuss in full detail in Chapter 7. Since the theory has two $SU(N)$ gauge groups, it has two topological angles θ_1 and θ_2. The theory is T invariant at $(\theta_1,\theta_2)=(0,0)$, $(\pi,0)$, $(0,\pi)$ and (π,π), and we discuss the global consistency of 't Hooft anomalies to see whether the vacuum is continuously connected without breaking T at those points. We propose phase diagrams in the (θ_1,θ_2) plane that are consistent with the constraints, and give its heuristic interpretation based on the dual superconductor model of confinement. The summary for this chapter is given in Section 6.4.

6.1 $SU(N)\times SU(N)$ bifundamental gauge theories with finite theta angles

We consider a gauge theory with the gauge group $SU(N)_1\times SU(N)_2$ and bifundamental matter fields. We use the convention that the gauge fields a_i of $SU(N)_i$ are realized as the traceless and Hermitian $N\times N$ matrix-valued local one-form. Our argument in the following is valid for any kinds of the bifundamental matter fields, but, as a specific example, one can consider single bifundamental Dirac field Ψ: Ψ belongs to the fundamental representation of $SU(N)_1$ and to the anti-fundamental representation of $SU(N)_2$, and it is realized as an $N\times N$ matrix-valued four-component Dirac fields. The $SU(N)_1\times SU(N)_2$ gauge transformation (u_1,u_2) acts on Ψ and a_i as $\Psi\mapsto u_1\Psi u_2^\dagger$ and $a_i\mapsto u_i a_i u_i^\dagger - \mathrm{i}u_i \mathrm{d}u_i^\dagger$. The classical action of the theory is given by

$$\begin{aligned} S &= -\frac{1}{2g_1^2}\int \mathrm{Tr}(G_1\wedge *G_1) - \frac{1}{2g_2^2}\int \mathrm{Tr}(G_2\wedge *G_2) + \int \mathrm{Tr}\,\overline{\Psi}(\not{D}+m)\Psi \\ &\quad + \frac{\mathrm{i}\theta_1}{8\pi^2}\int \mathrm{Tr}(G_1\wedge G_1) + \frac{\mathrm{i}\theta_2}{8\pi^2}\int \mathrm{Tr}(G_2\wedge G_2), \end{aligned} \tag{6.1}$$

where G_i is the field strength of the $SU(N)_i$ gauge group,

$$G_i = \mathrm{d}a_i + \mathrm{i}a_i \wedge a_i, \tag{6.2}$$

and

$$\not{D}\Psi = \gamma^\mu(\partial_\mu \Psi + \mathrm{i}a_{1\mu}\Psi - \mathrm{i}\Psi a_{2\mu}). \tag{6.3}$$

We assume that $m > 0$, and the matter part does not break time reversal symmetry explicitly. We denote the electric and magnetic charge of the $SU(N)_1 \times SU(N)_2$ gauge group as $(z_{e_1}, z_{m_1}) \oplus (z_{e_2}, z_{m_2})$, then the bifundamental Dirac field has the charge $(1,0) \oplus (N-1, 0)$ mod N. This theory has fundamental Wilson lines

$$W_1(C) = \mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a_1\right)\right], \quad W_2(C) = \mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a_2\right)\right], \tag{6.4}$$

and they have charge $(1,0) \oplus (0,0)$ and $(0,0) \oplus (1,0)$, respectively. $W_1 W_2^{-1}$ has the same charge with the dynamical fermion of this theory.

Let us describe the (0-form) symmetries of this theory. $U(1)_V$ is the phase rotation of the fermionic field

$$\Psi \mapsto \mathrm{e}^{\mathrm{i}\phi}\Psi, \overline{\Psi} \mapsto \mathrm{e}^{-\mathrm{i}\phi}\overline{\Psi}, \tag{6.5}$$

and this does not act on gauge fields a_i. If $g_1^2 = g_2^2$ and $\theta_1 = \theta_2$, there is the $(\mathbb{Z}_2)_I$ symmetry, which interchanges two gauge fields

$$a_1 \leftrightarrow -a_2^t \tag{6.6}$$

and acts on fermions as $\Psi \mapsto \Psi^t$. Except for these internal symmetries, there exist usual charge conjugation C, parity P, and time reversal T symmetries. When the $(\mathbb{Z}_2)_I$ symmetry and charge conjugation is combined, the gauge fields are transformed as $a_1 \leftrightarrow a_2$.

Recall that the $SU(N)$ pure Yang-Mills theory has the electric $\mathbb{Z}_N$ one-form symmetry, and thus this theory has $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry when the mass m of the Dirac fermion is infinitely large. At finite m, the bifundamental Dirac fermion becomes dynamical, and it breaks $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry to the stabilizer subgroup of $W_1 W_2^{-1}$. The $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry is explicitly broken to the diagonal $\mathbb{Z}_N$ one-form symmetry. Under this electric one-form symmetry, W_1 and W_2 have the same charge.

We study the consistency on the dynamics at $\theta = \pi$ using a mixed 't Hooft anomaly with this electric one-form symmetry, and constrain structures of the phase diagram. For that purpose, we first discuss gauging of $\mathbb{Z}_N$ one-form symmetry. For simplicity of discussion, we assume that the vacua are always trivially gapped, and we will study how the first-order phase transition happens as a function of θ_1 and θ_2.

6.2 Coupling with the $\mathbb{Z}_N$ two-form gauge fields

We couple the above $SU(N)_1 \times SU(N)_2$ bifundamental gauge theory to $\mathbb{Z}_N$ gauge fields in order to obtain $(SU(N)_1 \times SU(N)_2)/(\mathbb{Z}_N)_{\text{diagonal}}$ gauge theory. First, we discuss the possible charges of genuine line operators with dynamical matter fields. Bifundamental matters have the charge $(1,0) \oplus (-1,0)$, and thus, in order for the line with charge $(z_{e_1}, z_{m_1}) \oplus (z_{e_2}, z_{m_2})$ to be a genuine line, the Dirac quantization condition requires

$$z_{m_1} - z_{m_2} = 0 \mod N. \tag{6.7}$$

This means that two magnetic charges modulo N for $SU(N)_{1,2}$ gauge groups must be the same. Let us consider the case where the theory have a genuine line operator with the magnetic charge

$z_{m_1} = z_{m_2} = 1$. The Dirac quantization further restricts the possible purely electric genuine lines. To see it, let $(z_{e_1}, 0) \oplus (z_{e_2}, 0)$ be a charge of the genuine line, then we get

$$z_{e_1} + z_{e_2} = 0 \mod N, \tag{6.8}$$

from the Dirac quantization. As a result, the purely electric lines are given by $(W_1 W_2^{-1})^k$ for $k = 0, 1, \ldots, N-1$.[1]

We shall obtain such a theory by coupling the $SU(N) \times SU(N)$ bifundamental gauge theory to a $\mathbb{Z}_N$ topological field theory. We introduce the $\mathbb{Z}_N$ two-form gauge field B, and its classical action is given by the same action in (5.10):

$$S_{\mathrm{TFT}} = \frac{\mathrm{i}}{2\pi} \int F \wedge (\mathrm{d}A + NB) + \frac{\mathrm{i}Np}{4\pi} \int B \wedge B. \tag{6.9}$$

Here, A and B are $U(1)$ one-form and two-form gauge fields, and the equation of motion for F requires $NB = -\mathrm{d}A$, which makes B a $\mathbb{Z}_N$ two-form gauge field. We consider theories only on spin manifolds since we would like to include the case where bifundamental matters are Dirac fermions, then the parameter p must be an integer mod N: The condition on p being an integer comes from the requirement on the $U(1)$ one-form gauge invariance of (6.9). p is identified with $p + N$ since integrating out F yields

$$S_{\mathrm{TFT}} = 2\pi \mathrm{i} \frac{p}{N} \left(\frac{1}{2} \int \frac{\mathrm{d}A}{2\pi} \wedge \frac{\mathrm{d}A}{2\pi} \right), \tag{6.10}$$

and difference of p by multiples of N gives the difference of S_{TFT} in $2\pi\mathrm{i}\mathbb{Z}$. Hence, it does not affect the result in quantum theories.

To couple $SU(N)$ gauge fields a_1, a_2 to B, we first extend the gauge group $SU(N)_1 \times SU(N)_2$ to

$$\frac{SU(N)_1 \times SU(N)_2 \times U(1)}{\mathbb{Z}_N}, \tag{6.11}$$

and replace the $SU(N)$ gauge fields a_1 and a_2 by $U(N)$ gauge fields

$$\mathcal{A}_1 = a_1 + \frac{1}{N} A \mathbf{1}_N,\ \mathcal{A}_2 = a_2 + \frac{1}{N} A \mathbf{1}_N. \tag{6.12}$$

The $U(1)$ gauge field $A = \mathrm{Tr}(\mathcal{A}_1) = \mathrm{Tr}(\mathcal{A}_2)$ is the same with the one that appears in (6.9), and this creates the coupling of theories that we want. This $U(1)$ gauge field A does not couple to bifundamental fields, and it can be easily checked by an explicit form of the covariant derivative (6.3).

We construct the Wilson and 't Hooft line operators, which need not be genuine but must be gauge invariant. After that, we study the spectrum of genuine line operators to check whether we have obtained the $(SU(N)_1 \times SU(N)_2)/\mathbb{Z}_N$ gauge theory. The former definitions of Wilson lines in (6.4) are no longer gauge invariant after gauging the $\mathbb{Z}_N$ one-form symmetry. Let Σ be a two-dimensional surface with $C = \partial\Sigma$, and the gauge invariant Wilson loops are defined by

$$W_1(C, \Sigma) = \mathrm{Tr} \left[\mathcal{P} \exp \left(\mathrm{i} \oint_C a_1 \right) \right] \exp \left(\frac{\mathrm{i}}{N} \oint_C A + \mathrm{i} \int_\Sigma B \right), \tag{6.13}$$

$$W_2(C, \Sigma) = \mathrm{Tr} \left[\mathcal{P} \exp \left(\mathrm{i} \oint_C a_2 \right) \right] \exp \left(\frac{\mathrm{i}}{N} \oint_C A + \mathrm{i} \int_\Sigma B \right). \tag{6.14}$$

[1]In other words, with the gauge group $(SU(N)_1 \times SU(N)_2)/(\mathbb{Z}_N)_{\mathrm{diagonal}}$, the allowed Wilson (purely electric genuine) lines must be invariant under $(\mathbb{Z}_N)_{\mathrm{diagonal}}$ and thus have the vanishing electric charge. If the gauge group is $SU(N)_1 \times SU(N)_2$ instead of $(SU(N)_1 \times SU(N)_2)/(\mathbb{Z}_N)_{\mathrm{diagonal}}$, there is no constraint on the electric charge of the Wilson lines. Then, the Dirac quantization between the dyonic line $(z_{e_1}, z_{m_1}) \oplus (z_{e_2}, z_{m_2})$ and the Wilson line $(z'_{e_1}, 0) \oplus (z'_{e_2}, 0)$ leads to $z'_{e_1} z_{m_1} + z'_{e_2} z_{m_2} = 0 \pmod N$ with $z_{m_1} = z_{m_2}$ for an arbitrary $z'_{e_1} + z'_{e_2}$. Therefore, we have to choose $z_{m_1} = z_{m_2} = 0 \pmod N$ and magnetically charged line operators are excluded from the genuine line operator spectrum.

The magnetic one with charge $(0,1)\oplus(0,1)$ is also defined by

$$H(C,\Sigma)=\exp\left(\mathrm{i}\int_\Sigma (F+pB)\right). \tag{6.15}$$

Using Wilson lines, the genuine line operator of charge $(1,0)\oplus(-1,0)$ is given by

$$W_1(C,\Sigma)W_2(C,\Sigma)^{-1}=\mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a_1\right)\right]\left(\mathrm{Tr}\left[\mathcal{P}\exp\left(\mathrm{i}\oint_C a_2\right)\right]\right)^{-1}. \tag{6.16}$$

We can also construct a dyonic genuine line object,

$$H(C,\Sigma)W_1(C,\Sigma)^{-p}=\exp\left(\mathrm{i}\int_\Sigma F\right)\left(\mathrm{Tr}\left[\mathcal{P}\exp\mathrm{i}\oint_C\left(a_1+\frac{1}{N}A\right)\right]\right)^{-p}, \tag{6.17}$$

which has the charge $(-p,1)\oplus(0,1)$. By multiplying $(W_1W_2^{-1})^k$ to it, we can generally obtain the genuine line operator $HW_1^{-p}(W_1W_2^{-1})^k$ with the charge $(-p+k,1)\oplus(-k,1) \bmod N$. The discrete theta angle p designates the sum of electric charge for the genuine dyonic particles with the magnetic charges 1.

Since the topological θ angle is the central issue of our discussion, we compute how it is changed after the gauging in an explicit manner. In oder to maintain the 1-form gauge invariance, we should replace the gauge field strength G_1 and G_2 by $\mathcal{G}_1+B$ and $\mathcal{G}_2+B$, respectively, where $\mathcal{G}_i$ are the $U(N)$ field strengths of $\mathcal{A}_i$; $\mathcal{G}_i=\mathrm{d}\mathcal{A}_i+\mathrm{i}\mathcal{A}_i\wedge\mathcal{A}_i$. As a result, the topological θ term becomes

$$\begin{aligned}S_\theta &= \sum_{i=1,2}\frac{\mathrm{i}\theta_i}{8\pi^2}\int \mathrm{Tr}\left[(\mathcal{G}_i+B)\wedge(\mathcal{G}_i+B)\right]\\ &= \sum_{i=1,2}\frac{\mathrm{i}\theta_i}{8\pi^2}\int\left\{\mathrm{Tr}(\mathcal{G}_i\wedge\mathcal{G}_i)+2B\wedge\mathrm{Tr}(\mathcal{G}_i)+NB\wedge B\right\}.\end{aligned} \tag{6.18}$$

Using the equation of motion of F, $\mathrm{Tr}(\mathcal{G}_i)=\mathrm{d}A=-NB$, we obtain

$$S_\theta=\sum_{i=1,2}\frac{\mathrm{i}\theta_i}{8\pi^2}\int\left\{\mathrm{Tr}(\mathcal{G}_i\wedge\mathcal{G}_i)-NB\wedge B\right\}. \tag{6.19}$$

Using the consistency of the local counter term p with the T symmetry, we will discuss the possible phase structure of the $SU(N)\times SU(N)$ bifundamental gauge theories in the following sections.

6.2.1 Spontaneous T breaking at $(\theta_1,\theta_2)=(\pi,0)$ and $(0,\pi)$

We first constrain the possible dynamics of bifundamental gauge theories at $(\theta_1,\theta_2)=(\pi,0)$ or $(\theta_1,\theta_2)=(0,\pi)$. We follow the same logic given in Ref. [21], and start with the assumption that the vacuum at $\theta_1=\theta_2=0$ is trivially gapped without breaking the T symmetry. Therefore, there must be a way to gauge other symmetries without breaking the T symmetry at θ=0 by using the 't Hooft anomaly matching condition. Particularly when gauging the electric $\mathbb{Z}_N$ symmetry, the local counter term $\frac{\mathrm{i}Np}{4\pi}\int B\wedge B$ can be chosen to be T invariant from this argument, and such p must satisfy

$$2p=0 \mod N, \tag{6.20}$$

since $\int B\wedge B$ flips its sign under the T transformation.

We further assume that the vacua are always trivially gapped and that there is a way to continuously connect $(\theta_1, \theta_2) = (0,0)$ and $(\theta_1, \theta_2) = (\pi, 0)$, $(0, \pi)$ without phase transitions. We will show that there exists first-order phase transition associated with the spontaneous T breaking at $(\theta_1, \theta_2) = (\pi, 0)$ and at $(\theta_1, \theta_2) = (0, \pi)$ under this assumption.

Let us discuss the T symmetry at $\theta_1 = \pi$ with $\theta_2 = 0$ after gauging the $\mathbb{Z}_N$ one-form symmetry. Since T flips the orientation, p and θ_i change their signs and become $-p$ and $-\theta_i$, respectively. In oder to consider the theory at $\theta_1 = \pi$, we must consider not only the change $\theta_1 = \pi \mapsto \theta_1' = -\pi$ but also the shift $\theta_1' = -\pi \mapsto \theta_1' + 2\pi = \pi$ to discuss its T invariance. Under these transformations, the topological θ term is changed by

$$\Delta S_\theta = \frac{2\pi \mathrm{i}}{8\pi^2} \int \mathrm{Tr}(\mathcal{G}_1 \wedge \mathcal{G}_1) - \frac{\mathrm{i}N}{4\pi} \int B \wedge B. \tag{6.21}$$

The first term is in $2\pi\mathrm{i}\mathbb{Z}$, and thus does not affect the path integral. The second term shifts the value of p by -1. As a result, p is changed to $p \mapsto -p-1$ under the T transformation at $(\theta_1, \theta_2) = (\pi, 0)$, or equivalently, the partition function is changed as

$$\mathcal{Z} \mapsto \mathcal{Z} \exp\left[(-2p-1)\frac{\mathrm{i}N}{4\pi} \int B \wedge B\right]. \tag{6.22}$$

Thus the condition for the T invariance at $(\theta_1, \theta_2) = (\pi, 0)$ after gauging is given by

$$2p + 1 = 0 \mod N. \tag{6.23}$$

For even N, there is no such integer p. Therefore, there is an 't Hooft anomaly, and all the quasi-vacua must form pairs under T or become gapless to saturate the anomaly. For odd N, $p = (N-1)/2$ satisfies this condition, but it is inconsistent with the choice of p at $\theta_1 = \theta_2 = 0$. Since we put an assumption that a vacuum at $(\theta_1, \theta_2) = (\pi, 0)$ is continuously connected to the T-invariant vacuum at $\theta_1 = \theta_2 = 0$, consistency condition requires the existence of low-energy degrees of freedom to saturate this inconsistency, such as degenerate vacua or massless excitations. Since we have also assumed that the mass gap does not close, there exists the first-order phase transition at $(\theta_1, \theta_2) = (\pi, 0)$ in both cases associated with the spontaneous T breaking.

The same argument holds for $(\theta_1, \theta_2) = (0, \pi)$, and we can argue the spontaneous T breaking there.

6.2.2 Vacuum structure around $\theta_1 = \theta_2 = \pi$

Let us next discuss the consistency condition for T at $\theta_1 = \theta_2 = \pi$. This case is somewhat tricky, since there are two topologically distinct ways that connect $(\theta_1, \theta_2) = (\pi, \pi)$ and $(\theta_1, \theta_2) = (0, 0)$ (See Fig. 6.1). Since both θ_1 and θ_2 are 2π periodic for the gauge group $SU(N) \times SU(N)$, $\theta_1 = \theta_2 = \pi$ and $\theta_1 = -\theta_2 = -\pi$ are equivalent. We will discuss whether we encounter the first-order phase transition when changing θ_1 and θ_2 continuously from $(\theta_1, \theta_2) = (0, 0)$ to $(\theta_1, \theta_2) = (\pi, \pi)$ or $(-\pi, \pi)$.

Let us consider the case $(\theta_1, \theta_2) = (\pi, \pi)$. After gauging the $\mathbb{Z}_N$ symmetry, we must use the 2π periodicity of both θ_1 and θ_2 to discuss the T symmetry, and under these shifts the topological term (6.19) is changed by

$$\Delta S_\theta = \frac{2\pi \mathrm{i}}{8\pi^2} \sum_{i=1,2} \int \mathrm{Tr}(\mathcal{G}_i \wedge \mathcal{G}_i) - \frac{2\mathrm{i}N}{4\pi} \int B \wedge B. \tag{6.24}$$

On spin manifolds, the first term is in $2\pi\mathrm{i}\mathbb{Z}$ and does not affect the path integral. It thus changes p to $p-2$. One can understand this from the spectrum of genuine line operators. Originally, spectrum of genuine line operators are given by $(-p+k, 1) \oplus (-k, 1) \mod N$ with $k = 0, \dots, N-1$. Since they

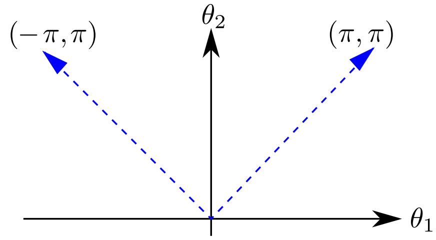


Figure 6.1: Two different paths connecting $(\theta_1,\theta_2)=(0,0)$ and $(\theta_1,\theta_2)=(\pi,\pi)\sim(-\pi,\pi)$. This figure is taken from Ref. [2].

have the monopole charge 1, the 2π shift of $\theta_{1,2}$ causes the shift of charge $(-p+k+1,1)\oplus(-k+1,1)$ due to the Witten effect, and they become $(-p+2+k',1)\oplus(-k',1)$ mod N with $k'=0,\ldots,N-1$ by putting $k'=k-1$. Notice that the spectrum is not changed only when $N=2$, and this will become important for our result.

Let us consider whether there is a way to gauge the electric one-form symmetry without breaking the T invariance at $\theta_1=\theta_2=\pi$. After gauging, we have a local counter term $\frac{\mathrm{i}Np}{4\pi}\int B\wedge B$, which flips the sign under T. It can be described effectively by the map $p\mapsto -p$, and $\theta_{1,2}=\pi\mapsto-\pi$. To get the original topological angle, we perform the 2π shift of both θ_1 and θ_2 that changes $-p\mapsto -p-2$. As a result, the T invariance at $\theta_1=\theta_2=\pi$ after gauging requires to choose p satisfying

$$p=-p-2 \mod N. \tag{6.25}$$

This always has the integer solution, and thus there is an T-invariant quasi-vacuum which may or may not be the true vacuum. Let us next discuss the global inconsistency. If $\theta_1=\theta_2=0$ and $\theta_1=\theta_2=\pi$ can be continuously connected without breaking T at $\theta_1=\theta_2=\pi$, then the integer solution p at $\theta_1=\theta_2=\pi$ must also be consistent with the T-invariant regularization at $\theta_1=\theta_2=0$; this says that

$$2=0 \mod N. \tag{6.26}$$

The vacuum at $\theta_1=\theta_2=0$ can be continuously changed to the T-invariant vacuum at $\theta_1=\theta_2=\pi$ without closing the mass gap only if this condition holds.

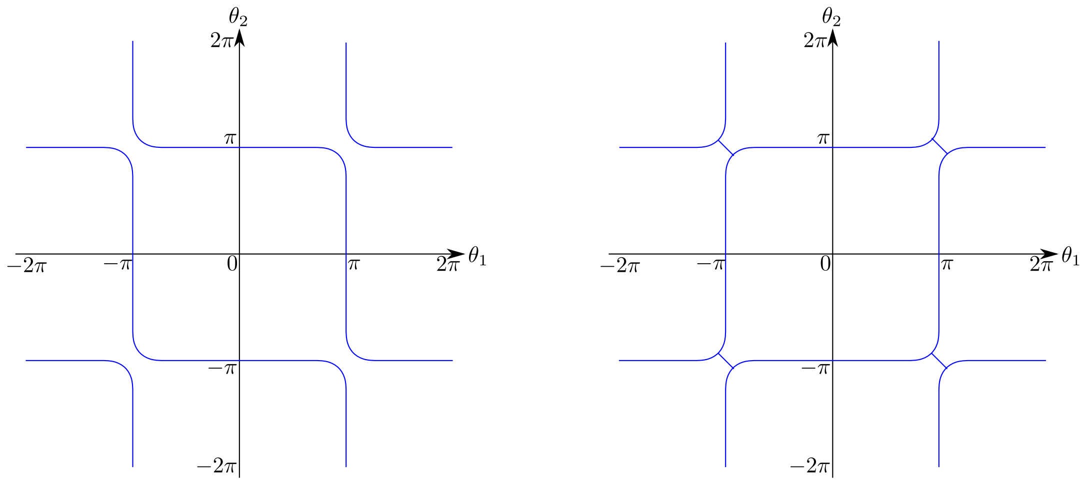
For $N\geq 3$, the above global consistency relation cannot be true. One possibility is that the vacua at $\theta_{1,2}=0$ and $\theta_{1,2}=\pi$ are separated by first-order phase transitions. Another possibility is that the vacuum at $\theta_{1,2}=\pi$ breaks T spontaneously to saturate the inconsistency. For $N=2$, we cannot impose any constraints on the state at $\theta_1=\theta_2=\pi$ from our argument, and basically any possibilities are allowed[2].

Next, let us consider what happens when connecting $\theta_1=\theta_2=0$ and $\theta_1=-\theta_2=-\pi$. In this case, the T transformation at $(\theta_1,\theta_2)=(-\pi,\pi)$ is associated with the shift $\theta_1\mapsto\theta_1-2\pi$ and $\theta_2\mapsto\theta_2+2\pi$. The change of the topological term under these shifts is given by

$$\Delta S_\theta=\frac{\mathrm{i}}{4\pi}\int \mathrm{Tr}(-\mathcal{G}_1\wedge\mathcal{G}_1+\mathcal{G}_2\wedge\mathcal{G}_2)\in 2\pi\mathrm{i}\mathbb{Z}. \tag{6.27}$$

Therefore it does not affect the path integral at all. In this case, the T transformation changes $p\mapsto -p$ mod N as in the case of $\theta_1=\theta_2=0$. When connecting $\theta_1=\theta_2=0$ and $\theta_1=-\theta_2=-\pi$,

[2] This might be because we consider theories defined only on spin manifolds. If we restrict our attention to theories without fermions, then theories can be defined also on non-spin manifolds. We can repeat the same argument for non-spin cases at least formally just by changing the identification of the discrete theta angle from $p\sim p+N$ to $p\sim p+2N$. The necessary condition for unbroken T given by (6.26) becomes $2=0$ mod $2n$, and then we would find that T must be broken for all $N\geq 2$.



(a) T is unbroken at $(\theta_1, \theta_2) = (\pi, \pi)$

(b) T is broken at $(\theta_1, \theta_2) = (\pi, \pi)$

Figure 6.2: Possible phase boundaries of $SU(N) \times SU(N)$ bifundamental gauge theories in the θ_1-θ_2 plane ($N \geq 3$). These figures are taken from Ref. [2].

the global consistency holds and thus the vacua can be continuously connected without the phase transition and T needs not be broken at $\theta_1 = -\theta_2 = -\pi$.

By combining the result and respecting the 2π periodicity of $\theta_{1,2}$, we obtain Fig. 6.2 as a possible phase boundary of the first-order phase transition in the (θ_1, θ_2) plane when $N \geq 3$. Whether the phase boundary opens at $(\theta_1, \theta_2) = (\pi, \pi)$ depends on details of the dynamics such as matter contents. In Fig. 6.2a, we consider the possibility when the T symmetry is unbroken at $\theta_1 = \theta_2 = \pi$. In this case, the first-order phase transition line must separate $\theta_1 = \theta_2 = 0$ and $\theta_1 = \theta_2 = \pi$, but $\theta_1 = \theta_2 = 0$ and $\theta_1 = -\theta_2 = \pi$ can be smoothly connected. In Fig. 6.2b, the T symmetry is spontaneously broken at $\theta_1 = \theta_2 = \pi$, and thus there is a first-order phase transition line around it. In this case, $\theta_1 = \theta_2 = 0$ and $\theta_1 = -\theta_2 = \pi$ would be separated by another first-order line because the vacuum at $\theta_1 = \theta_2 = 0$ is continuously connected to the T-invariant quasi-vacuum at $\theta_1 = -\theta_2 = \pi$ but not to the true T-broken vacuum according to the global consistency relation.

One may wonder whether the first-order phase transition line in Fig. 6.2 can terminate so that one can smoothly change (θ_1, θ_2) from $(0, 0)$ to $(2\pi, 0)$ without phase transitions. In Fig. 6.2, this is impossible and we claim that it is a general result for $N \geq 3$. By repeating the same argument on the T transformation after gauging the $\mathbb{Z}_N$ symmetry, the condition for the T invariance at $(\theta_1, \theta_2) = (2\pi, 0)$ is given by $p = -p - 2 \bmod N$. Of course this has the solution, but it is inconsistent with the T invariant choice at $(\theta_1, \theta_2) = (0, 0)$ when $N \geq 3$. This means that if we could connect $(\theta_1, \theta_2) = (0, 0)$ and $(2\pi, 0)$ without any phase transition and without closing the mass gap, then T must be spontaneously broken at $(\theta_1, \theta_2) = (2\pi, 0)$ but this is the contradiction because $(\theta_1, \theta_2) = (0, 0)$ and $(2\pi, 0)$ must be equivalent for $SU(N) \times SU(N)$ gauge theories. If we further assume that the mass gap does not close at generic (θ_1, θ_2), then $(\theta_1, \theta_2) = (0, 0)$ and $(2\pi, 0)$ must be separated by the first-order phase transition line when $N \geq 3$. For $N = 2$, this is not the case.

6.3 Interpretation via the dual superconductor picture

The purpose of this section is to understand the result intuitively from the dual superconductor model of confinement [52, 97, 98]. Let us first consider the case $m \to \infty$ where the bifundamental matters decouple. Then, we have two decoupled $SU(N)$ Yang-Mills theories, so let us start with the

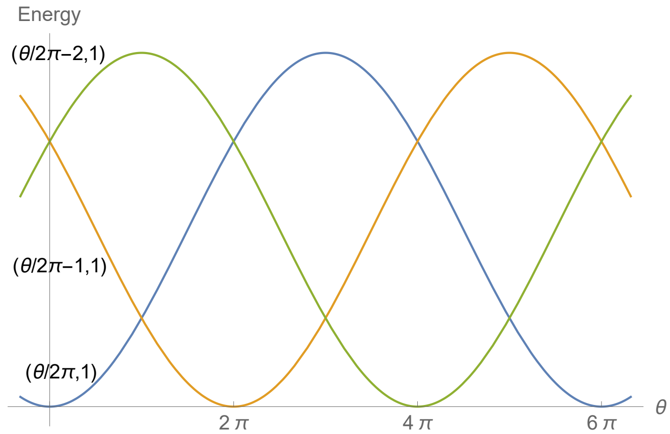


Figure 6.3: Schematic figure on the ground state energy $E(\theta)$ of the $SU(N)$ Yang-Mills theory based on the dual superconductor model for $N = 3$. There are N different branches labeled by the condensed charge $(\theta/2\pi - k, 1)$ of dyons ($k = 0, 1, \dots, N-1$), and each branch is $2\pi N$ periodic. This figure is taken from Ref. [2].

discussion for the $SU(N)$ Yang-Mills theory.

6.3.1 $SU(N)$ Yang-Mills theory

Following the dual superconductor model, we assume that confinement of $SU(N)$ Yang-Mills theory on $\mathbb{R}^4$ is caused by condensation of magnetic monopoles or dyons. Let us say that their charges are given by $(-k, 1)$ mod N with $k = 0, 1, \dots, N-1$ at $\theta = 0$. This assumes that all the Wilson loops with nontrivial center elements obey the area law. There are N candidates of condensed particles, and correspondingly there are N different quasi-vacua. To be specific, let us assume that the magnetic monopole with charge $(0, 1)$ condenses at $\theta = 0$ in the true vacuum. Now, we turn on the finite topological θ angle, and the Witten effect shifts charges of dyons to $(-k + \theta/2\pi, 1)$. Since the charge of each dyon goes back to its original value only after the shift of $2\pi N$, each branch of quasi-vacua are $2\pi N$ periodic in θ instead of 2π periodic. However, the true vacuum must be 2π periodic in terms of θ, so there must be some jump among quasi-vacua between $0 < \theta < 2\pi$.

Let us pay attention to the charge at $\theta = \pi$. Assuming that no phase transition occurs for $0 < \theta < \pi$, then the charge of condensed particles (magnetic monopole at $\theta = 0$) becomes $(\theta/2\pi, 1)$ due to the Witten effect. It is not invariant under the T transformation at $\theta = \pi$ although the theory is T invariant. Under the T transformation, the charge $(1/2, 1)$ is mapped to $(-1/2, 1)$, and thus the quasi-vacua with charges $(\pm 1/2, 1)$ must have the same energy because of the T symmetry of the theory. Therefore, the first-order phase transition occurs at $\theta = \pi$, and the true vacuum jumps from the branch with the condensed charge $(\theta/2\pi, 1)$ to the another branch with the condensed charge $(-1 + \theta/2\pi, 1)$ (see Fig. 6.3). This is how the T symmetry is spontaneously broken at $\theta = \pi$ for pure $SU(N)$ Yang-Mills theory in the dual superconductor scenario.

To summarize the case for the $SU(N)$ Yang-Mills theory, let us denote $E_0(\theta)$ as the energy of the quasi-vacuum with the condensed charge $(\theta/2\pi, 1)$. T symmetry tells us that $E_0(\theta) = E_0(-\theta)$, and N-ality shows that $E_0(\theta + 2\pi N) = E_0(\theta)$. There are N candidates for the condensate, $(-k + \theta/2\pi, 1)$ with $k = 0, \dots, N-1$, and the energy of the true vacuum is

$$E_{SU(N)}(\theta) = \min\{E_0(\theta - 2\pi k) \,|\, k = 0, 1, \dots, N-1\}. \tag{6.28}$$

If E_0 is smooth, it is natural to have the bump for $E(\theta)$ at $\theta = \pi$, at which the branch jumps from

$E_0(\theta)$ to $E_0(\theta - 2\pi)$ with the first-order phase transition (see Fig. 6.3). In the large-N limit, it is well established that the Yang-Mills vacuum is described by the minimum of N branches [51, 61].

Before going to the case of bifundamental gauge theories, let us deepen our understandings on the meaning of 't Hooft anomaly and global inconsistency. T invariance at $\theta = \pi$ requires that $p = -p-1$ mod N, and it cannot be solved for even N. In the dual superconductor picture, the condensed particles at $\theta = \pi$ have charges $(\pm 1/2, 1), \dots, (\pm(N-1)/2, 1)$ and they form $N/2$ T-invariant pairs. Including quasi-vacua, no states can be invariant under T, and this is suggested by the 't Hooft anomaly. Next, let us consider the case of odd N, then 't Hooft anomaly does not exist by setting $p = (N-1)/2$. The condensed charges are given by $(\pm 1/2, 1), \dots, (\pm(N-2)/2, 1)$ and $(N/2, 1)$. Since the quasi-vacuum with the condensed charge $(N/2, 1)$ is invariant under T (see Fig. 6.3), one cannot argue the spontaneous T breaking at $\theta = \pi$ without putting another assumption. The point is that the state with the charge $(N/2, 1)$ at $\theta = \pi$ is not continuously connected to the vacuum with the charge $(0, 1)$ at $\theta = 0$, so the absence of the first-order phase transition at $0 < \theta < \pi$ can purge this state from our consideration on vacua. In the language of the global inconsistency, this is implied by the fact that there is no common integer p for the T invariance at $\theta = 0$ and $\theta = \pi$.

6.3.2 $SU(N) \times SU(N)$ bifundamental gauge theories

Let us now discuss $SU(N) \times SU(N)$ Yang-Mills theory. Considering the limit $m \to \infty$ so that bifundamental matters decouple, we just have two copies of the above argument.

We first connect $\theta_1 = \theta_2 = 0$ and $\theta_1 = \theta_2 = \pi$. We select the path $\theta_1 = \theta_2$ for instance and denote the common angle as $\theta := \theta_1 = \theta_2$. We now have N^2 candidates for the condensed particles with the charge $(-k + \theta/2\pi, 1) \oplus (-\ell + \theta/2\pi, 1)$ mod N with $k, \ell = 0, 1, \dots, N-1$, and thus the ground-state energy is given by

$$E_{SU(N)\times SU(N)}(\theta) = \min\{E_0(\theta - 2\pi k) + E_0(\theta - 2\pi\ell) \,|\, k, \ell = 0, \dots, N-1\}. \tag{6.29}$$

By assumption that the monopole $(0, 1)$ condenses for the $SU(N)$ Yang-Mills theory at $\theta = 0$, the quasi-vacuum with $(\theta/2\pi, 1) \oplus (\theta/2\pi, 1)$ is selected when θ is close to zero. At $\theta = \pi$, T is broken and the ground state must be at least two-fold degenerate. In our limit $m \to \infty$, there is four-fold degeneracy at $\theta = \pi$, and the condensed charges for those four states are

$$\begin{aligned} &(\theta/2\pi, 1) \oplus (\theta/2\pi, 1),\ (\theta/2\pi - 1, 1) \oplus (\theta/2\pi - 1, 1), \\ &(\theta/2\pi - 1, 1) \oplus (\theta/2\pi, 1),\ (\theta/2\pi, 1) \oplus (\theta/2\pi - 1, 1). \end{aligned} \tag{6.30}$$

This is because the T symmetry is extended to $\mathbb{Z}_2 \times \mathbb{Z}_2$ from $\mathbb{Z}_2$ in the limit $m \to \infty$ as a result of the decoupling between two $SU(N)$ Yang-Mills theories. If we assume that $E_0(\theta)$ is smooth and monotonically increasing for $0 < \theta < 2\pi$, there is the first-order phase transition from the sate with condensed charge $(\theta/2\pi, 1) \oplus (\theta/2\pi, 1)$ to the another one with $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi - 1, 1)$ at $\theta = \pi$ (see Fig. 6.4).

Let us turn on finite m and make the bifundamental matters dynamical. Then, the T symmetry becomes $\mathbb{Z}_2$ and the accidental four-fold degeneracy at $\theta = \pi$ must be resolved. Let us first notice that states with the charge $(\theta/2\pi, 1) \oplus (\theta/2\pi, 1)$ and $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi - 1, 1)$ cannot be mixed by dynamical bifundamental fields since the difference of their charges is different from the bifundamental charge $(1, 0) \oplus (-1, 0)$. On the other hand, the difference of two charges $(\theta/2\pi, 1) \oplus (\theta/2\pi - 1, 1)$ and $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi, 1)$ is given by $(1, 0) \oplus (-1, 0)$, and this is nothing but the charge of dynamical matter fields Ψ. These states can be mixed as a result of interacting bifundamental matters, which leads to the non-degenerate quasi-vacuum with the mass gap.

If the energy of the mixed states of $(\theta_1/2\pi, 1) \oplus (\theta_2/2\pi - 1, 1)$ and $(\theta_1/2\pi - 1, 1) \oplus (\theta_2/2\pi, 1)$ is lowered by dynamical matter fields as in Fig. 6.5a, then the mixed state is selected as the ground

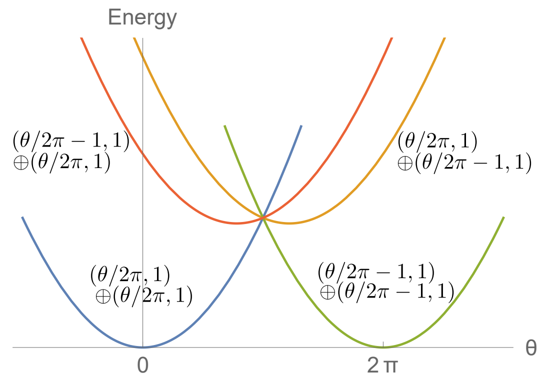
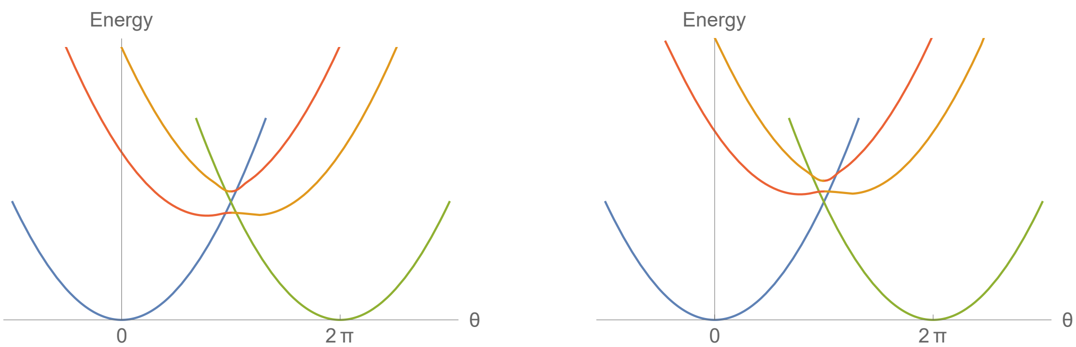


Figure 6.4: Energies of the quasi-vacua of the $SU(N) \times SU(N)$ gauge theory in the limit of $m \to \infty$ when $\theta = \theta_1 = \theta_2$. This figure is taken from Ref. [2].



(a) T is unbroken at $\theta = \pi$

(b) T is broken at $\theta = \pi$

Figure 6.5: Two possibilities of the mixing of states due to dynamical bifundamental matter fields when $\theta = \theta_1 = \theta_2$ are around π. These figures are taken from Ref. [2].

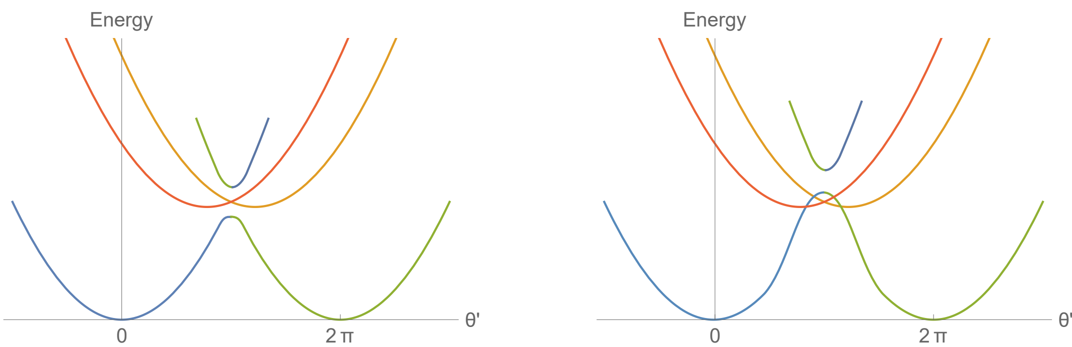
state around $\theta = \pi$. In this case, T (and $(\mathbb{Z}_2)_I$ if exists) need not be broken, but the first-order phase transition happens across the path connecting $\theta = 0$ and $\theta = \pi$. If the energy of the mixed states of $(\theta/2\pi, 1) \oplus (\theta/2\pi - 1, 1)$ and $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi, 1)$ is lifted as in Fig. 6.5b, then they drop out from the consideration and there is the first order phase transition from the state with $(\theta/2\pi, 1) \oplus (\theta/2\pi, 1)$ to the one with $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi - 1, 1)$. In this case, T is spontaneously broken at $\theta = \pi$.

We can also understand why no phase transition is required when connecting $\theta_1 = \theta_2 = 0$ and $\theta_1 = -\theta_2 = -\pi$. For instance, let us pick up a path with $\theta_1 = -\theta_2$, and denote $\theta' = -\theta_1 = \theta_2$. By taking the limit $m \to \infty$, we can again consider possible phases using the dual superconductor picture. The four-fold degeneracy at $\theta' = \pi$ happens at $m = \infty$, and the condensed charges for those four states are given by

$$\begin{aligned} &(-\theta'/2\pi, 1) \oplus (\theta'/2\pi, 1),\ (-\theta'/2\pi + 1, 1) \oplus (\theta'/2\pi - 1, 1), \\ &(-\theta'/2\pi + 1, 1) \oplus (-\theta'/2\pi, 1),\ (-\theta'/2\pi, 1) \oplus (\theta'/2\pi - 1, 1). \end{aligned} \tag{6.31}$$

Figure for the vacuum energy is almost the same with Fig. 6.4 just by replacing the label of charges in a straightforward manner.

Let us turn on dynamical bifundamental fields by making m finite. In this case, the states with charges $(-\theta'/2\pi, 1) \oplus (\theta'/2\pi, 1)$ and $(-\theta'/2\pi + 1, 1) \oplus (\theta'/2\pi - 1, 1)$ can be mixed by dynamical matter fields, while the states with $(-\theta'/2\pi + 1, 1) \oplus (\theta'/2\pi, 1)$ and $(-\theta'/2\pi, 1) \oplus (\theta'/2\pi - 1, 1)$



(a) T is unbroken at $\theta' = \pi$ (b) T is broken at $\theta' = \pi$

Figure 6.6: Two possibilities of the mixing of states due to dynamical bifundamental matter fields when $\theta' = -\theta_1 = \theta_2$ are around π. Thiese figures are taken from Ref. [2].

cannot be mixed. Depending on relative energies of those states, we obtain Fig. 6.6 for quasi-vacua of bifundamental gauge theories as a function of $\theta' = -\theta_1 = \theta_2$. By checking charges of condensed particles, we can notice that Figs. 6.5a and 6.6a are connected, and T is unbroken at $\theta_1 = \theta_2 = \pi$. Similarly, Figs. 6.5b and 6.6b are connected, and T is spontaneously broken at $\theta_1 = \theta_2 = \pi$. These explain two possible phase boundaries shown in Figs. 6.2a and 6.2b, respectively.

We have so far explained how T is spontaneously broken at $\theta = \pi$ for bifundamental theories, but our result suggests that it needs not happen if $N = 2$. We close this section by observing why $N = 2$ can be special. Let us consider the case with $\theta = \theta_1 = \theta_2$ for example, then above conclusion comes from the fact that the states with condensed particles $(\theta/2\pi, 1) \oplus (\theta/2\pi, 1)$ and $(\theta/2\pi - 1, 1) \oplus (\theta/2\pi - 1, 1)$ mod N cannot be mixed. This is because the difference of these charges of condensed particles is $(1, 0) \oplus (1, 0)$ mod N, while the charge of dynamical bifundamental matters is $(1, 0) \oplus (-1, 0)$ mod N. These two are different for $N \geq 3$, but they are the same at $N = 2$. Therefore, for $N = 2$, these two states can also be mixed by dynamical bifundamental fields, and thus we need no first-order phase transition lines that separate $\theta = 0$ and $\theta = \pi$.

6.4 Summary

We have studied the phase structure for $SU(N) \times SU(N)$ bifundamental gauge theories at finite topological angles by applying consistency for mixed 't Hooft anomalies of T and center symmetry. For the gauge group $SU(N) \times SU(N)$, there are two topological angles θ_1 and θ_2, and there are four T invariant points, $(\theta_1, \theta_2) = (0, 0)$, $(\pi, 0)$, $(0, \pi)$, and (π, π). We discuss that there must be a first-order phase transition at $(\theta_1, \theta_2) = (\pi, 0)$ and $(0, \pi)$ associated with spontaneous breaking of the T symmetry, so there is a first-order phase transition line through these points. The global consistency is discussed at $(\theta_1, \theta_2) = (\pi, \pi)$ but there are two different ways to connect $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (\pi, \pi)$ because the point is equivalent to $(\theta_1, \theta_2) = (-\pi, \pi)$. We observe for $N \geq 3$ that the vacua at $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (\pi, \pi)$ cannot be continuously connected without breaking the T symmetry at $(\theta_1, \theta_2) = (\pi, \pi)$, but also that the vacua at $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (-\pi, \pi)$ can without breaking any symmetries. We have proposed phase diagrams in the (θ_1, θ_2) plane that are consistent with these constraints. To understand it better, we have given a heuristic interpretation of the result based on the dual superconductor model of confinement and the role of dynamical bifundamental fields is clarified.

Chapter 7

’t Hooft anomalies in massless QCD

Why has the quantum chromodynamics (QCD) been attracting many people and driven them to explore its phase diagram? One of the most intriguing aspects might be its fascinating phase structure, where various symmetry breaking patterns and appearances of corresponding phases have extensively been studied, particularly on finite-(T,μ) phase diagram, with temperature T and chemical potential μ (see e.g. [99–112] for reviews). Surprisingly, most of the phases have been predicted to be realized nontrivially in the sense that they are not symmetric and gapped phase. Is there any accountability that excludes the trivial phase appearing in the QCD phase diagram? – The answer is affirmative. Indeed, the absence of trivial phase may be attributed to an ’t Hooft anomaly that we have discussed so far. The existence of an ’t Hooft anomaly at finite temperatures and chemical potentials rules out the realization of trivially gapped phase according to the anomaly matching argument. As a consequence, the whole (T,μ)-plane of the QCD phase diagram could be filled with nontrivial phases. This situation is referred to the realization of “persistent order”. In this chapter, we study the phase diagram of QCD-like theory by utilizing the anomaly constraints and show that the persistent order is realized on the finite-(T,μ) phase diagram. After these analyses, we make speculative comments on the real QCD phase diagram.

We first study ’t Hooft anomalies for four-dimensional QCD with massless quarks in the fundamental representation of the gauge group $SU(N_\mathrm{c})$ and the flavor group $SU(N_\mathrm{f})$, which is described by the action

$$S = \frac{2}{g^2}\int \mathrm{tr}_\mathrm{c}[F\wedge *F] + \int \mathrm{d}^4x\, \mathrm{tr}_\mathrm{f}[\bar{\psi}\not{D}\psi], \tag{7.1}$$

where $F = \mathrm{d}a + \mathrm{i}a\wedge a$ with $SU(N_\mathrm{c})$ gauge field a, $\not{D} = \gamma^\mu(\partial_\mu + \mathrm{i}a)$, and ψ is a massless Dirac fermion in a bifundamental representation of $SU(N_\mathrm{c})\times SU(N_\mathrm{f})$. The traces tr_c and tr_f are understood to be the sum over the color and flavor indices. In QCD, the $SU(N_\mathrm{c})$ is a gauge group and the $SU(N_\mathrm{f})$ is a global symmetry group as opposed to the bifundamental gauge theory with the dynamical gauge group $SU(N)\times SU(N)$ that we treated in Chapter 6.

Needless to say, QCD is strongly coupled in infrared energy scale and extracting nonperturbative data of the theory is quite important task to unravel its vacuum structure or phase structure, say, at finite temperature, finite density, and so on. As we have seen so far, ’t Hooft anomalies and associated UV/IR anomaly matching arguments provide rigorous data. It is interesting to see whether there exist ’t Hooft anomalies in addition to the one for the continuous global symmetry $SU(N_\mathrm{f})_\mathrm{L}\times SU(N_\mathrm{f})_\mathrm{R}\times U(1)_\mathrm{V}$ that ’t Hooft originally discovered (see Chapter 2 for discussion). It turns out that there indeed are! These anomalies, in turn, lead to interesting consequences on the QCD vacua and phase structure. We will apply the techniques that we have developed so far, particularly in the last chapter, to explore the anomaly constraints. We will find quite similar arguments between the 't Hooft anomalies in QCD and bifundamental gauge theory. The result reported in this chapter is based on our recent work [3] and some new results from our work in progress.

Let us start with specifying the classical global internal symmetry of the theory:

$$\frac{SU(N_\mathrm{f})_\mathrm{L}\times SU(N_\mathrm{f})_\mathrm{R}\times U(1)_\mathrm{L}\times U(1)_\mathrm{R}}{\mathbb{Z}_{N_\mathrm{c}}\times\mathbb{Z}_{N_\mathrm{f}}\times\mathbb{Z}_{2N_\mathrm{f}}}. \tag{7.2}$$

The quotient by $\mathbb{Z}_{N_\mathrm{c}}$ is due to the overlap between the actions by the center of $SU(N_\mathrm{c})$ and $U(1)_\mathrm{q}$. Concretely, the action of $(U_\mathrm{c},\mathrm{e}^{\mathrm{i}\alpha})\in SU(N_\mathrm{c})\times U(1)_\mathrm{q}$ on the fermion ψ is given by $\psi\mapsto\mathrm{e}^{\mathrm{i}\alpha}U_\mathrm{c}\psi$ contains the $\mathbb{Z}_{N_\mathrm{c}}$ subgroup $(\omega,\omega^{-1})\in SU(N_\mathrm{c})\times U(1)_\mathrm{q}$ with $\omega=\mathrm{e}^{2\pi\mathrm{i}/N_\mathrm{c}}$ that leaves ψ unchanged. Therefore, we need to remove the redundant symmetry so that quarks are in the faithful representation of $(SU(N_\mathrm{c})\times U(1)_\mathrm{q})/\mathbb{Z}_{N_\mathrm{c}}$. The same is also true for the quotient by $\mathbb{Z}_{N_\mathrm{f}}\subset SU(N_\mathrm{f})_\mathrm{V}\times U(1)_\mathrm{q}$. The quotient by $\mathbb{Z}_{2N_\mathrm{f}}$ is slightly trickier and can be understood as follows: The transformation by $\mathbb{Z}_{2N_\mathrm{f}}$ subgroup of $U(1)_\mathrm{A}\subset U(1)_\mathrm{L}\times U(1)_\mathrm{R}$ can actually be reproduced by the center of chiral flavor and $U(1)_\mathrm{q}$ transformation. For instance, a $2\pi/N_\mathrm{f}$ transformation of the center of $SU(N_\mathrm{f})_\mathrm{L}$ followed by a $-\pi/N_\mathrm{f}$ $U(1)_\mathrm{q}$-transformation yields a $\mathbb{Z}_{2N_\mathrm{f}}$ axial transformation:

$$\begin{array}{ccccc}\psi_\mathrm{L} & \xrightarrow{SU(N_\mathrm{f})_\mathrm{L}} & \mathrm{e}^{\mathrm{i}\frac{2\pi}{N_\mathrm{f}}}\psi_\mathrm{L} & \xrightarrow{U(1)_\mathrm{q}} & \mathrm{e}^{\mathrm{i}\frac{\pi}{N_\mathrm{f}}}\psi_\mathrm{L},\\ \psi_\mathrm{R} & \xrightarrow{SU(N_\mathrm{f})_\mathrm{L}} & \psi_\mathrm{R} & \xrightarrow{U(1)_\mathrm{q}} & \mathrm{e}^{-\mathrm{i}\frac{\pi}{N_\mathrm{f}}}\psi_\mathrm{R},\end{array} \tag{7.3}$$

where $\psi_{\mathrm{L/R}}$ is a left/right-handed fermion. Hence, we need to eliminate $\mathbb{Z}_{2N_\mathrm{f}}$ subgroup to obtain the symmetry with faithful representation.

The $U(1)_\mathrm{L}\times U(1)_\mathrm{R}$ symmetry is explicitly broken to $U(1)_\mathrm{V}\times(\mathbb{Z}_{2N_\mathrm{f}})_\mathrm{axial}$ on the quantum level due to the chiral anomaly caused by the $SU(N_\mathrm{c})$ gauge sector as follows: Under the axial $U(1)$ rotation $\psi\mapsto\mathrm{e}^{\mathrm{i}\gamma_5\theta}\psi,\ \bar\psi\mapsto\bar\psi\mathrm{e}^{\mathrm{i}\gamma_5\theta}$ the partition function $\mathcal{Z}$ transforms as

$$\mathcal{Z}\mapsto\mathcal{Z}\exp\left[\mathrm{i}\frac{2N_\mathrm{f}\theta}{8\pi^2}\int\mathrm{tr}_\mathrm{c}F\wedge F\right], \tag{7.4}$$

where $\frac{1}{8\pi^2}\int\mathrm{tr}_\mathrm{c}F\wedge F\in\mathbb{Z}$ is the instanton number. Therefore, although the axial $U(1)$ symmetry is broken for generic θ, $\mathbb{Z}_{2N_\mathrm{f}}$ subgroup is unbroken, which is generated by $\mathrm{e}^{\mathrm{i}\theta}$ with $\theta=\pi/N_\mathrm{f}$. Since $U(1)_\mathrm{V}$ symmetry leads to the quark number conservation, we will denote it $U(1)_\mathrm{q}$.

The symmetry on the quantum level is[1]

$$\frac{SU(N_\mathrm{f})_\mathrm{L}\times SU(N_\mathrm{f})_\mathrm{R}\times U(1)_\mathrm{q}\times(\mathbb{Z}_{2N_\mathrm{f}})_\mathrm{axial}}{\mathbb{Z}_{N_\mathrm{c}}\times\mathbb{Z}_{N_\mathrm{f}}\times\mathbb{Z}_{2N_\mathrm{f}}}. \tag{7.6}$$

We also notice that there is a time-reversal symmetry in addition to these global symmetry, which plays an essential role to show the 't Hooft anomaly with finite imaginary chemical potentials.

[1] Actually, it can simply be expressed as

$$\frac{SU(N_\mathrm{f})_\mathrm{L}\times SU(N_\mathrm{f})_\mathrm{R}\times U(1)_\mathrm{q}}{\mathbb{Z}_{N_\mathrm{c}}\times\mathbb{Z}_{N_\mathrm{f}}}. \tag{7.5}$$

We intentionally write $(\mathbb{Z}_{2N_\mathrm{f}})_\mathrm{axial}$ here to emphasize that the axial symmetry exists, which plays an pivotal role in the following discussion.

7.1 Color-flavor center symmetry and its 't Hooft anomaly

We set $N_\mathrm{f} = N_\mathrm{c} \equiv N$ for the sake of simplicity although we can carry out the same argument as long as $\gcd(N_\mathrm{c}, N_\mathrm{f}) \neq 1$. We explore an 't Hoof anomaly by paying attention to the global symmetry

$$\frac{SU(N)_\mathrm{f}}{\mathbb{Z}_N} \times (\mathbb{Z}_{2N})_\mathrm{axial}, \tag{7.7}$$

where $SU(N)_\mathrm{f}$ is the vector part of the flavor symmetry and $\mathbb{Z}_N$ is the common center subgroup of color $SU(N)_\mathrm{c}$ and flavor $SU(N)_\mathrm{f}$ groups. To detect the 't Hooft anomaly involving the symmetry, we check the fate of axial $\mathbb{Z}_{2N}$ symmetry after coupling the theory to $SU(N)_\mathrm{f}/\mathbb{Z}_N$ background gauge field. We first introduce $SU(N)_\mathrm{f}$ background gauge field A_f by replacing the covariant derivative as

$$D\psi = \mathrm{d}\psi + \mathrm{i}a\psi \mapsto \tilde{D}\psi = \mathrm{d}\psi + \mathrm{i}a\psi + \mathrm{i}\psi A_\mathrm{f}. \tag{7.8}$$

Now, one find the theory quite similar to the bifundamental gauge theory that we studied in the last chapter, where the flavor gauge field was treated as a dynamical field. After the introduction of flavor background gauge field, a one-form center symmetry emerges as a global symmetry even with fundamental fermions. This symmetry transformation acts on the color and flavor Wilson lines and called the color-flavor center symmetry [113]. We repeat the same story to couple the theory to background two-form gauge field for the center symmetry. We first prepare $(U(N)_\mathrm{c} \times U(N)_\mathrm{f})/\mathbb{Z}_N$ gauge group by promoting $SU(N)$ gauge fields to $U(N)$ gauge fields with a $U(1)$ gauge field C,

$$\tilde{a} = a + \frac{1}{N}C, \quad \tilde{A}_\mathrm{f} = A_\mathrm{f} - \frac{1}{N}C, \tag{7.9}$$

and couple to the topological action

$$\frac{\mathrm{i}Np}{4\pi}\int B \wedge B. \tag{7.10}$$

The two-form gauge field B satisfies a constraint $NB = \mathrm{d}C$.[2] Then, we obtain the partition function $\mathcal{Z}[A_\mathrm{f}, B]$ coupled to background $SU(N)/\mathbb{Z}_N$ gauge fields.

In order to detect an 't Hooft anomaly, we see the transformation of $\mathcal{Z}[A_\mathrm{f}, B]$ under the $(\mathbb{Z}_{2N})_\mathrm{axial}$ transformation. Applying (7.4) to the gauged action, we obtain

$$\begin{aligned}\mathcal{Z}[A_\mathrm{f}, B] \mapsto \mathcal{Z}[A_\mathrm{f}, B] \exp\left[\mathrm{i}\frac{1}{4\pi N}\int \mathrm{tr}_{\mathrm{c,f}}(\tilde{F} - \tilde{F}_\mathrm{f}) \wedge (\tilde{F} - \tilde{F}_\mathrm{f})\right] \\ \equiv \mathcal{Z}[A_\mathrm{f}, B]\mathrm{e}^{\mathrm{i}\mathcal{A}[A_\mathrm{f}, B]},\end{aligned} \tag{7.11}$$

with $\tilde{F} = \mathrm{d}\tilde{a} + \mathrm{i}\tilde{a} \wedge \tilde{a}$ and $\tilde{F}_\mathrm{f} = \mathrm{d}\tilde{A}_\mathrm{f} + \mathrm{i}\tilde{A}_\mathrm{f} \wedge \tilde{A}_\mathrm{f}$. Let us calculate the anomaly $\mathcal{A}[A_\mathrm{f}, B]$ to see whether the axial symmetry is preserved.

$$\begin{aligned}\mathcal{A}[A_\mathrm{f}, B] &= \frac{1}{4\pi N}\int \mathrm{tr}_{\mathrm{c,f}}(\tilde{F} \wedge \tilde{F} - 2\tilde{F} \wedge \tilde{F}_f + \tilde{F}_f \wedge \tilde{F}_f) \\ &= \frac{1}{4\pi}\int \mathrm{tr}_\mathrm{c}\tilde{F} \wedge \tilde{F} + \frac{1}{4\pi N}\int 2\mathrm{d}C \wedge \mathrm{d}C + \frac{1}{4\pi}\int \mathrm{tr}_\mathrm{f}\tilde{F}_f \wedge \tilde{F}_f \\ &= \frac{N}{2\pi}\int B \wedge B \ (\mathrm{mod}\ 2\pi),\end{aligned} \tag{7.12}$$

[2] This constraint is equivalent to introducing an additional topological term $(\mathrm{i}/2\pi)\int F \wedge (NB - \mathrm{d}C)$ with a Lagrange multiplier two-form F as we did in (6.9).

where we have used $(1/4\pi)\int \mathrm{tr}_\mathrm{c}\tilde{F}\wedge\tilde{F}\in 2\pi\mathbb{Z}$, $(1/4\pi)\int \mathrm{tr}_\mathrm{f}\tilde{F}_\mathrm{f}\wedge\tilde{F}_\mathrm{f}\in 2\pi\mathbb{Z}$, and the constraint $NB=\mathrm{d}C$ in the last equality. Since this cannot be cancelled by (7.10) for $N>2$, we conclude that there is an 't Hooft anomaly involving $SU(N)_\mathrm{f}/\mathbb{Z}_N$ and $(\mathbb{Z}_{2N})_\mathrm{axial}$ symmetries. For $N=2$ case, the relation $2B=\mathrm{d}C$ leads to

$$\frac{N}{2\pi}\int B\wedge B=\frac{1}{4\pi}\int \mathrm{d}C\wedge\mathrm{d}C\in 2\pi\mathbb{Z}, \tag{7.13}$$

which means that there is no 't Hooft anomaly.

The anomaly matching condition is satisfied in QCD because it has nonvanishing chiral condensate due to the spontaneous chiral symmetry breaking, which breaks $(\mathbb{Z}_{2N})_\mathrm{axial}$ to $\mathbb{Z}_2$.

7.2 Finite temperature

Next, we will pursue further consequences of the 't Hooft anomaly by studying the phases at finite temperature following [96]. As we will see, it is totally nontrivial whether the 't Hooft anomaly survives after circle compactification with small radius compared with the size of M^3 (at high temperature). After compactifying the imaginary time direction on a circle S^1 with circumference β, i.e., $M^4=M^3\times S^1$, we expect that the one-form $\mathbb{Z}_N$ symmetry acting on the Wilson loops on M^4 splits into a one-form $\mathbb{Z}_N$ symmetry acting on the Wilson loops on M^3 and the zero-form $\mathbb{Z}_N$ symmetry acting on Polyakov loops $P_\mathrm{c}=\mathrm{tr}_\mathrm{c}[\mathcal{P}\exp(\mathrm{i}\oint_{S^1}a)]$ and $P_\mathrm{f}=\mathrm{tr}_\mathrm{f}[P\exp(\mathrm{i}\oint_{S^1}A_\mathrm{f})]$ winding S^1, and the two-form gauge field B becomes $B=B^{(2)}+B^{(1)}\wedge\beta^{-1}\mathrm{d}\tau$, where $B^{(2)}$ and $B^{(1)}$ are respectively two- and one-form gauge fields on M^3 (see Section 5.4). This is correct in the bifundamental gauge theory where the color and flavor gauge fields are built in the original theory as dynamical fields. However, the target theory that we want to study is three-dimensional effective theory at finite temperatures obtained as a result of the dimensional reduction via circle compactification. To detect the 't Hooft anomaly of this theory, we first couple the theory to the three dimensional flavor background A_f on M^3. At this point, we notice that the flavor Polyakov loops do not show up as opposed to the color Polyakov loops. Hence under the zero-form $\mathbb{Z}_N$ transformation, the color Polyakov loop transforms as

$$P_\mathrm{c}\to \mathrm{e}^{2\pi\mathrm{i}/N}P_\mathrm{c}. \tag{7.14}$$

One could eliminate the phase by changing the boundary condition of the fermions,

$$\psi(\tau+\beta)=\psi(\tau)\mathrm{e}^{2\pi\mathrm{i}/N}. \tag{7.15}$$

Thus, there is no $\mathbb{Z}_N$ zero-form symmetry in the compactified theory because the transformation changes the boundary condition for fermions in the compactified direction, and hence, changes the theory. In the compactified theory, B is just reduced to $B^{(2)}$ and the 't Hoof anomaly vanishes as [3]

$$\int_{M^4}B\wedge B\xrightarrow{\text{compactify}}\int_{M^3\times S^1}B^{(2)}\wedge B^{(2)}=0. \tag{7.16}$$

However, we can make the anomaly survive by introducing nontrivial flavor holonomies around S^1 or twisted boundary conditions. Here, we consider two possibilities: $SU(N)_\mathrm{f}$- and $U(1)$-twisted boundary conditions. The massless QCD with the boundary condition twisted by flavor $SU(N)_\mathrm{f}$ symmetry is known as $\mathbb{Z}_N$-QCD [113–122]. On the other hand, inserting $U(1)$-twisted boundary condition is equivalent to introducing the imaginary chemical potential. We will discuss the phase structure of these theories from the viewpoint of 't Hooft anomaly constraints.

[3]We will explain this computation in the next subsection in detail.

7.2.1 Anomaly constraints on phase structure of massless $\mathbb{Z}_N$-QCD

We construct so called $\mathbb{Z}_N$-QCD by promoting the $\mathbb{Z}_N$ zero-form transformation to a symmetry in the following procedure. We introduce the $SU(N)_{\mathrm{f}}$-twisted boundary condition in the S^1 direction,

$$\psi(\tau+\beta)=\Omega\psi(\tau), \tag{7.17}$$

where Ω is a $N\times N$ matrix in the flavor space. This boundary condition is changed under the $\mathbb{Z}_N$ zero-form transformation as $\Omega\to\omega\Omega$ with $\omega=\mathrm{e}^{2\pi\mathrm{i}/N}$, which, combined with flavor rotation of the boundary condition, can be made into a symmetry. We prepare a matrix S which acts on the flavor space of fermions such that

$$S\Omega S^{-1}=\omega^{-1}\Omega, \tag{7.18}$$

is satisfied. Ω and S are, for instance, given by

$$\Omega=\omega^{-(N-1)/2}\begin{pmatrix}1&0&0&\cdots&0\\0&\omega&0&\cdots&0\\0&0&\omega^3&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&\omega^{N-1}\end{pmatrix},\quad S=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1\\1&0&0&\cdots&0\end{pmatrix}, \tag{7.19}$$

Let us see how the boundary condition for fermion (7.17) changes under the $\mathbb{Z}_N$ zero-form transformation followed by the flavor shift S. As we noted the $\mathbb{Z}_N$ zero-form transformation induces the boundary condition

$$\tilde{\psi}(\tau+\beta)=\omega\Omega\tilde{\psi}(\tau). \tag{7.20}$$

After the flavor shift,

$$S\tilde{\psi}(\tau+\beta)=\omega S\Omega S^{-1}S\tilde{\psi}(\tau)=\Omega S\tilde{\psi}(\tau+\beta), \tag{7.21}$$

the original boundary condition is restored. Hence this is a symmetry which acts on the local operators on M^3 as

$$\psi\mapsto S\psi,\quad P_{\mathrm{c}}\mapsto\omega P_{\mathrm{c}}, \tag{7.22}$$

and referred to as the "shift symmetry" [96]. The crucial point is the introduction of the twisted boundary condition with a proper choice of the associated shift S. Equivalently, it is realized by introducing nontrivial flavor holonomy $\exp(\mathrm{i}\int_{S^1}A_{\mathrm{f}}^{\tau})$ winding S^1, which will be useful for gauging symmetry later. From the latter point of view, the $\mathbb{Z}_N$-QCD may be interpreted as QCD at finite temperature with flavor dependent imaginary chemical potentials represented by the background flavor holonomy. It is noted that the boundary condition breaks flavor symmetry $SU(N)_{\mathrm{f}}$ to its abelian subgroup $U(1)_{\mathrm{f}}^{N-1}$.

The twisted boundary condition changes the global symmetry from (7.7) to

$$(\mathbb{Z}_N)_{\text{shift}}\times\frac{U(1)_{\mathrm{f}}^{N-1}}{\mathbb{Z}_N}\times(\mathbb{Z}_{2N})_{\text{axial}}. \tag{7.23}$$

This time, we need to introduce the background gauge fields for the shift symmetry and flavor symmetry, by which the background gauge fields on $M^3\times S^1$ can be expressed as follows:

$$A=A_{\mathrm{f}}^{U(1)}+B^{(1)}+A_{\mathrm{f}}^{\tau}, \tag{7.24}$$

$$B=B^{(2)}+B^{(1)}\wedge\beta^{-1}\mathrm{d}\tau. \tag{7.25}$$

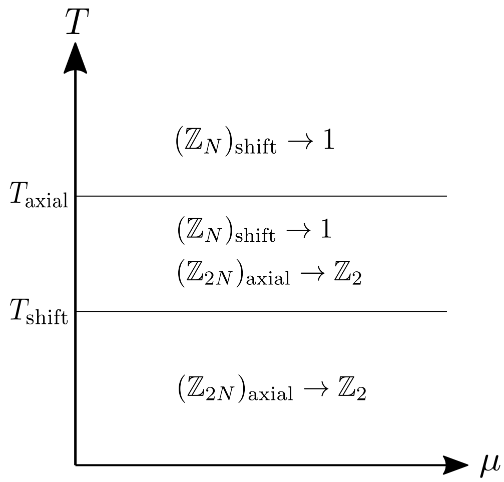


Figure 7.1: A possible finite-(T, μ) phase diagram of the massless $\mathbb{Z}_N$-QCD that satisfies the anomaly constraint $T_{\rm chiral} \geq T_{\rm deconf}$. Symmetry breaking patterns are shown in each phase.

A and B are the one-form flavor gauge field and two-form center gauge field on $M^3 \times S^1$. What show up in the right hand side are background gauge fields on M^3. $A_{\rm f}^{U(1)}$ is for $U(1)_{\rm f}^{N-1}$ symmetry, $A_{\rm f}^{\tau}$ arises because of the twisted boundary condition, $B^{(2)}$ is for the color-flavor center symmetry, and finally, $B^{(1)}$ is interpreted as a background gauge fields both for the flavor shift and the center symmetry and thus appears both in A and B. Once we identify the relations, the computation of an 't Hooft anomaly is completely parallel. Actually, the 't Hooft anomaly for the symmetry (7.23) can be obtained just by substituting B (7.25) into (7.12),

$$\frac{N}{2\pi}\int_{M^4} B \wedge B \xrightarrow{\text{compactify}} \frac{N}{\pi}\int_{M^3} B^{(2)} \wedge B^{(1)}, \tag{7.26}$$

which is indeed nontrivial and the 't Hooft anomaly survives thanks to the twisted boundary condition.

Possible phase structure of massless $\mathbb{Z}_N$-QCD

A consequence of the UV/IR matching for the 't Hooft anomaly (7.26) is that realizable phases at finite temperatures T and chemical potentials μ are described by either by some CFT, TQFT, or spontaneously symmetry broken state [3].[4] We will only focus on the last possibility, in which case at least one of the symmetries among $(\mathbb{Z}_N)_{\rm shift}$, $U(1)_{\rm f}^{N-1}/\mathbb{Z}_N$ and $(\mathbb{Z}_{2N})_{\rm axial}$ has to be broken. Each symmetry is responsible for the following phase identification:

- $(\mathbb{Z}_N)_{\rm shift}$ symmetry, which is originated from the color-flavor center symmetry on M^4 and acts on the color Polyakov loops on M^3, distinguishes between the confinement and deconfinement phases, in each of which the symmetry is unbroken and broken respectively.
- $U(1)_{\rm f}^{N-1}/\mathbb{Z}_N$ symmetry is a remnant of the vector-like flavor symmetry, which can be broken at finite density although is not allowed to be broken in vacuum [70].
- $(\mathbb{Z}_{2N})_{\rm axial}$ is broken to $\mathbb{Z}_2$ by forming chiral condensate, by which the chiral symmetry is broken as well. Hence, $(\mathbb{Z}_{2N})_{\rm axial}$ can be used to distinguish between the chiral symmetry broken and restored phases.

[4] The anomaly constraints are valid at finite density because the finite chemical potential does not change the derivation of the anomaly.

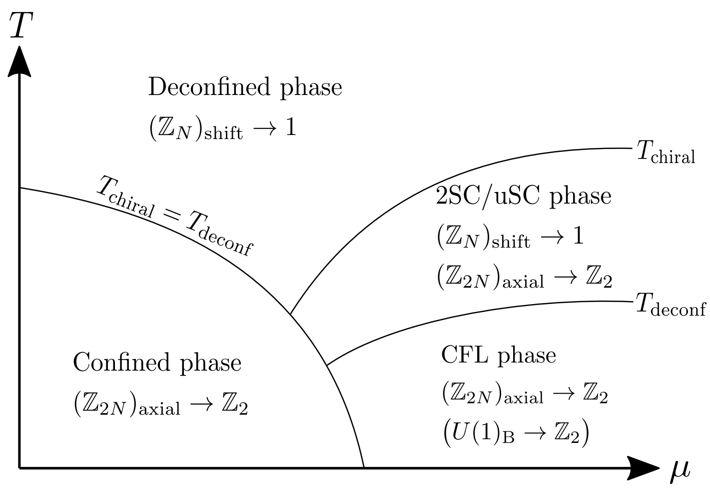


Figure 7.2: Another possible phase diagram of the massless $\mathbb{Z}_N$-QCD that is more similar to the conventional QCD phase diagram. It also satisfies the anomaly constraint. $(\mathbb{Z}_N)_{\text{shift}}$ or $(\mathbb{Z}_{2N})_{\text{axial}}$ is necessarily broken in all phases. See the main text for detailed discussion.

Let us further sharpen the anomaly constraint by excluding the phases with broken $U(1)_{\text{f}}^{N-1}/\mathbb{Z}_N$ symmetry from the circumstantial evidence. The anomaly matching condition under the assumptions mentioned so far imposes the constraint that phases realized on finite-(T,μ) QCD phase diagram has to spontaneously breaks $(\mathbb{Z}_N)_{\text{shift}}$ or $(\mathbb{Z}_{2N})_{\text{axial}}$ symmetry at least. Since $(\mathbb{Z}_N)_{\text{shift}}$ is broken at lower temperatures and $(\mathbb{Z}_{2N})_{\text{axial}}$ is broken at higher temperatures, a consequence is that the chiral restoration temperature T_{chiral} must be equal to or larger than the deconfinement temperature T_{deconf},

$$T_{\text{chiral}} \geq T_{\text{deconf}}, \tag{7.27}$$

in order to avoid the trivial phase with both symmetries restored [21, 24, 26].

Let us take a closer look at typical diagrams allowed by this constraint. Figure 7.1 shows a relatively simple phase diagram satisfying the inequality (7.27). $T = T_{\text{chiral}}(\mu)$ provides the phase boundary of the chiral symmetry broken phase, while $T = T_{\text{deconf}}(\mu)$ separates the deconfined phase from the confined phase. There is no symmetric gapped phase as required by the 't Hooft anomaly.

Finally, we add another ingredient, global symmetry $U(1)_{\text{B}} \equiv U(1)_{\text{q}}/\mathbb{Z}_N$, and deform phase transition lines in such a way that it looks like a conventional QCD phase diagram. The resultant phase structure is shown in Fig. 7.2. We should emphasize that the anomaly constraint (7.27) is still satisfied. Inclusion of $U(1)_{\text{B}}$ in the symmetry breaking pattern allows us to discuss the color superconductivity, such as the color-flavor locking (CFL), 2SC, and uSC phases. The symmetry breaking patters of each phase are shown in Fig. 7.2. It is noted that the phase structure also satisfy the anomaly matching condition of 't Hooft anomaly involving $SU(N_{\text{f}})_{\text{L}} \times SU(N_{\text{f}})_{\text{R}} \times U(1)_{\text{q}}$ discussed in Section 2.1 [71, 72].

We make some speculative remarks on the N-flavor QCD phase diagram in the low temperature limit. Since our derivation of anomaly has the four-dimensional origin, the anomaly matching argument is valid no matter how large the size β of compactification is. By taking the zero-temperature limit $\beta \to 0$, we expect that the effect of boundary condition would disappear. If the vector-like flavor symmetry is unbroken, it is indeed conceivable that the effect of flavor-dependence in the boundary condition disappears. Under this assumption, anomaly matching argument claims that finite-density massless N-flavor QCD shows nontrivial phase at any quark chemical potentials in the zero-temperature limit. We can confirm that this is indeed the case at least for small μ and also for sufficiently large μ in the zero-temperature limits, where the anomaly is satisfied by SSB of discrete axial symmetry. At the zero temperature, anomaly matching for continuous chiral symmetry $SU(N_\mathrm{f})_\mathrm{L} \times SU(N_\mathrm{f})_\mathrm{R} \times U(1)_\mathrm{q}$ imposes further constraints on possible dynamics in cold dense QCD.

7.2.2 Imaginary chemical potential and Roberge-Weiss transition

We first outline the same story with the $U(1)$-twisted boundary condition instead of the flavor twisted one in the S^1 direction,

$$\psi(\tau+\beta) = \mathrm{e}^{\mathrm{i}\phi}\psi(\tau), \tag{7.28}$$

where a scalar constant ϕ is the imaginary chemical potential.[5] Since this boundary condition is changed under the $\mathbb{Z}_N$ zero-form transformation as $\mathrm{e}^{\mathrm{i}\phi} \to \mathrm{e}^{\mathrm{i}(\phi+2\pi/N)}$, we may restrict ϕ in $[-\pi/N, \pi/N)$. Here, we recall that there is the time-reversal symmetry without the twisted boundary condition. The boundary condition breaks the time-reversal symmetry for generic ϕ except for $\phi = 0$ and $-\pi/N$ (one may notice that the situation is very similar to the Yang-Mills theory or bifundamental gauge theory with the θ terms): Under the $\mathbb{Z}_N$ zero-form transformation, the fermion transforms as $\psi \mapsto \tilde{\psi}$ with the boundary condition,

$$\tilde{\psi}(\tau+\beta) = \mathrm{e}^{\mathrm{i}(\phi+2\pi/N)}\tilde{\psi}(\tau). \tag{7.29}$$

Then, the time-reversal transformation T changes the boundary condition as

$$\mathsf{T}\tilde{\psi}(\tau+\beta) = \mathsf{T}\mathrm{e}^{\mathrm{i}(\phi+2\pi/N)}\mathsf{T}^{-1}\mathsf{T}\tilde{\psi}(\tau) = \mathrm{e}^{-\mathrm{i}(\phi+2\pi/N)}\mathsf{T}\tilde{\psi}(\tau). \tag{7.30}$$

If ϕ satisfies $\phi = -\phi - 2\pi/N \pmod{2\pi}$, which is solved by $\phi = -\pi/N$, the time-reversal combined with $\mathbb{Z}_N$ zero-form together forms a symmetry group $(\mathbb{Z}_2)_\mathrm{shift} \subset \mathbb{Z}_N \rtimes \mathbb{Z}_2^\mathsf{T}$ [26]. This is the shift symmetry in the current case. In other words, $\phi = -\pi/N \pmod{2\pi/N}$ is a high-symmetry point with an enhanced symmetry by $(\mathbb{Z}_2)_\mathrm{shift}$.[6]

The global symmetry at high symmetry points is

$$(\mathbb{Z}_2)_\mathrm{shift} \times \frac{SU(N)_\mathrm{f}}{\mathbb{Z}_N} \times (\mathbb{Z}_{2N})_\mathrm{axial}. \tag{7.31}$$

However, it is normally difficult to gauge the T symmetry, which is contained in $(\mathbb{Z}_2)_\mathrm{shift}$ in this case, and indeed it will be easier to gauge $(\mathbb{Z}_{2N})_\mathrm{axial}$ symmetry rather than $(\mathbb{Z}_2)_\mathrm{shift}$ as we will discuss below.

We consider the four dimensional massless QCD and attempt to gauge $(\mathbb{Z}_{2N})_\mathrm{axial}$. By adding the axion term to the QCD action,

$$S_\mathrm{axion} = \frac{\mathrm{i}}{8\pi^2}\int a_5 \mathrm{tr}_\mathrm{c} F\wedge F, \tag{7.32}$$

we restore the $U(1)_A$ symmetry. $a_5(x)$ is the axion field and the transformation law is given by

$$a_5 \mapsto a_5 - 2N_\mathrm{f}\alpha. \tag{7.33}$$

along with $\psi \mapsto \mathrm{e}^{\mathrm{i}\gamma_5\alpha}\psi$ under the $U(1)_A$ transformation, that is designated to cancel the anomalous $U(1)_A$ contribution. Now, we add the background gauge field A_5 for the restored $U(1)_A$ symmetry,

[5] ϕ/β is usually referred to as the imaginary chemical potential.

[6] It is noted that $\phi = 0$ is also a high-symmetry point because the time reversal is symmetry, but the time reversal does not interwined with $\mathbb{Z}_N$ transformation. Therefore, an anomaly does not appear at $\phi = 0$.

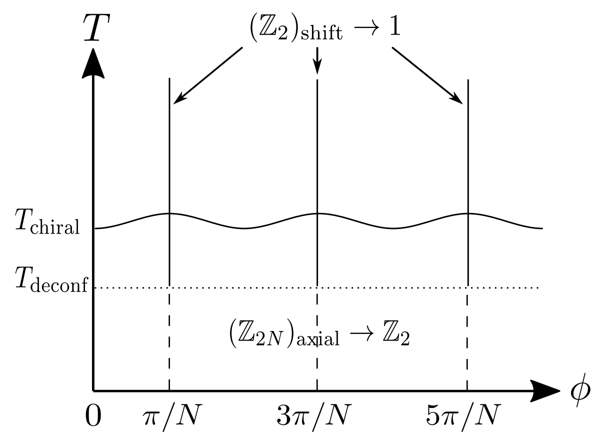


Figure 7.3: A possible phase diagram of the massless QCD at finite temperatures T and finite imaginary chemical potentials ϕ. $(\mathbb{Z}_2)_{\rm shift}$ or $(\mathbb{Z}_{2N})_{\rm axial}$ is always broken at high symmetry points $\phi = \pi/N \pmod{2\pi/N}$. A wavy line represents the phase boundary of the $(\mathbb{Z}_{2N})_{\rm chiral}$ symmetric phase. below which the symmetry is spontaneously broken. The vertical solid lines are the first-order phase transition lines on which the $(\mathbb{Z}_2)_{\rm shift}$ symmetry is spontaneously broken. This lines are bounded by a dotted line $T = T_{\rm deconf}$ and the $(\mathbb{Z}_2)_{\rm shift}$ is restored on vertical dashed lines.

and then require $2N_{\rm f}A_5 = {\rm d}a_5$, so that A_5 plays a role of $(\mathbb{Z}_{2N_{\rm f}})_{\rm axial}$ gauge field. Under the T transformation the total partition function changes as

$$\mathcal{Z} \mapsto \mathcal{Z} \exp\left[-\frac{2{\rm i}}{8\pi^2}\int a_5 {\rm tr}_{\rm c} F \wedge F\right]. \tag{7.34}$$

Furthermore, we introduce background gauge fields $A_{\rm q}$ for $U(1)_{\rm q}$ and $B_{\rm c}$, satisfying $N_{\rm c}B_{\rm c} = {\rm d}C_{\rm c}$, for $\mathbb{Z}_{N_{\rm c}}$ one-form symmetry, the latter of which is the color-$U(1)_{\rm q}$ center symmetry. We obtain the following anomaly

$$\begin{aligned} \mathcal{A} &= -\frac{2}{8\pi^2}\int a_5 {\rm tr}_{\rm c}(\tilde{F} + {\rm d}A_{\rm q}) \wedge (\tilde{F} + {\rm d}A_{\rm q}) \\ &= -\frac{2}{8\pi^2}\int a_5[{\rm tr}_{\rm c}\tilde{F} \wedge \tilde{F} + 2N_{\rm c}B_{\rm c} \wedge {\rm d}A_{\rm q} + N_{\rm c}{\rm d}A_{\rm q} \wedge {\rm d}A_{\rm q}], \end{aligned} \tag{7.35}$$

involving the symmetry

$$\mathsf{T} \times \frac{U(1)_{\rm q}}{\mathbb{Z}_{N_{\rm c}}} \times (\mathbb{Z}_{2N_{\rm f}})_{\rm axial}. \tag{7.36}$$

Next, we would like to ask if this anomaly survives even after the circle compactification with $U(1)_{\rm q}$-twisted boundary condition, or equivalently with a nontrivial $U(1)_{\rm q}$ holonomy along S^1, which is included in $A_{\rm q}$, i.e., $\phi \equiv \exp\left[{\rm i}\int_{S^1} A_{\rm q}\right]$. Then, the resultant anomaly is given by

$$\begin{aligned} \mathcal{A} &= \frac{4N_{\rm c}}{8\pi^2}\int_{S^1} A_{\rm q} \int_{M^3} {\rm d}a_5(B_{\rm c} + {\rm d}A_{\rm q}) \\ &= \frac{N_{\rm c}N_{\rm f}}{\pi^2}\phi \int_{M^3} A_5(B_{\rm c} + {\rm d}A_{\rm q}). \end{aligned} \tag{7.37}$$

If we take $\phi = 0$, then there is no anomaly. Classically, $\phi = -\pi/N_{\rm c}$ is the other T symmetric point and let us take a look at the anomaly,

$$\mathcal{A} = -\frac{N_{\rm f}}{\pi}\int_{M^3} A_5 B_{\rm c} - \frac{1}{2\pi}\int_{M^3} {\rm d}a_5 {\rm d}A_{\rm q}. \tag{7.38}$$

The second term is in $2\pi\mathbb{Z}$. We need to make sure that the second term is not canceled by a local counter term. Let us take $N_\mathrm{c} = N_\mathrm{f} \equiv N$ for simplicity. Then, the following gauge invariant counter term is allowed to add

$$\frac{\mathrm{i}Np}{\pi} \int_{M^3} A_5 B_\mathrm{c}, \tag{7.39}$$

with a parameter $p \in \mathbb{Z}_N$. Since the added counter term (7.39) is odd under the T transformation, the total partition function is changed as

$$\mathcal{Z} \mapsto \mathcal{Z} \exp\left[(-2p-1)\frac{\mathrm{i}N}{\pi}\int_{M^3} A_5 B_\mathrm{c}\right]. \tag{7.40}$$

Therefore, the conditions for the anomaly to vanish are

$$2p + 1 = 0 \pmod N, \tag{7.41}$$

which leads to the 't Hooft anomaly for even N and the global inconsistency between $\phi = 0$ and $-\pi/N$ for odd N. If we assume the trivial phase at $\phi = 0$ and no phase transition between the points, we may conclude the existence of nontrivial phase for odd N.

We extract similar constraints as those in the last subsection. Suppose neither topological order nor CFT emerges as an infrared theory, the global symmetry (7.31) must be broken spontaneously at the high symmetry points $\phi = -\pi/N \pmod{2\pi/N}$. As a result, the chiral restoration temperature T_chiral and "deconfinement" temperature T_deconf has to obey the relation $T_\mathrm{chiral} \geq T_\mathrm{deconf}$ at $\phi = -\pi/N \pmod{2\pi/N}$ unless the vector like flavor symmetry is broken. The $(\mathbb{Z}_2)_\mathrm{deconf}$-broken states appear in the phase diagram 7.3 on vertical solid lines, which are the first-order phase transition lines when crossed horizontally, known as the Roberge-Weiss transition [123]. Below T_shift the $(\mathbb{Z}_2)_\mathrm{shift}$ symmetry is restored. By definition, this symmetry is explicitly broken except for the high symmetry points.

A resultant finite-(T, ϕ) phase structure of the massless QCD is shown in 7.3 satisfying the anomaly constraint, where we used another input that the $(\mathbb{Z}_{2N})_\mathrm{axial}$ symmetry is spontaneously broken below T_chiral at any ϕ. The phase diagram is $2\pi/N$-periodic in ϕ because of the Roberge-Weiss periodicity.

Chapter 8

Conclusion

The 't Hooft anomaly and the global inconsistency impose strong constraints on low-energy effective theory via their UV/IR matching argument, which are powerful data to clarify nonperturbative aspects of QFTs. In particular, the trivial phase is strictly forbidden if we find an 't Hooft anomaly. If the 't Hooft anomaly survives after the circle compactification to obtain the finite temperature system, the consequence of the anomaly matching condition, in turn, leads to the realization of the persistent order, i.e., the system cannot be disordered by either quantum or thermal fluctuation. Furthermore, the global inconsistency may provide the nonperturbative data of infrared theory even without the 't Hooft anomaly. It is almost equally powerful as the 't Hooft anomaly in that it excludes the trivial phase without any phase separation between high symmetry points in the parameter space of theories. We applied these techniques to several quantum mechanical models, the pure Yang-Mills theory and bifundamental gauge theory with topological terms, and massless QCD, all of which yield nontrivial 't Hooft anomalies and global inconsistencies.

In the quantum mechanical models that we have discussed possessed 't Hooft anomaly and global inconsistency due to the topological term analogous to the conventional θ term in four-dimensional gauge theory. The UV/IR matching conditions force all the states to be degenerate in the energy spectra at $\theta = 0$ and π. These degeneracies can also understood by the operator formalism, which is analogous to the Kramers degeneracy for time reversal symmetric systems with half-integer spin fermions.

After brief introduction to the generalized global symmetry we jumped into the 't Hooft anomalies and global inconsistencies involving the center symmetry, which is a $\mathbb{Z}_N$ one-form symmetry in $SU(N)$ pure Yang-Mills theory and $SU(N) \times SU(N)$ gauge theory with bifundamental Dirac fermions. From symmetry point of view, these theories are quite similar to the quantum mechanical models with the topological term and the analyses of anomalies and their UV/IR matching arguments are mostly parallel. The crucial difference is that we need to introduce the two-form background gauge field in order to detect the 't Hooft anomaly and global inconsistency involving $\mathbb{Z}_N$ one-form symmetry. This leads to further interesting results: The anomalies involving higher-form symmetries survive after circle compactification. Hence, we could also extract nontrivial constraints for systems at finite temperatures.

Finally, we attacked QCD and its phase structure from the viewpoint of 't Hooft anomaly. The obvious difficulty is that the existence of fundamental fermions spoils the center symmetry. A remarkable thing was that, even without the higher-form symmetry in the original theory, the emergent color-flavor symmetry plays an essential role in gauging the flavor $SU(N)/\mathbb{Z}_N$ symmetry. This is how we derived the new anomaly in QCD. Moreover, we shed light on the massless $\mathbb{Z}_N$ QCD and massless QCD with imaginary chemical potential by means of anomaly constraints. The former turned out to realize the persistent order on finite-(T, μ) phase diagram. The latter also possess the new 't Hooft anomaly, which shows the robustness of RW first-order transition in phase structure at finite imaginary chemical potentials.

Bibliography

[1] Y. Kikuchi and Y. Tanizaki, "Global inconsistency, 't Hooft anomaly, and level crossing in quantum mechanics," *PTEP* **2017** (2017) 113B05, arXiv:1708.01962 [hep-th].

[2] Y. Tanizaki and Y. Kikuchi, "Vacuum structure of bifundamental gauge theories at finite topological angles," *JHEP* **06** (2017) 102, arXiv:1705.01949 [hep-th].

[3] Y. Tanizaki, Y. Kikuchi, T. Misumi, and N. Sakai, "Anomaly matching for phase diagram of massless $\mathbb{Z}_N$-QCD," arXiv:1711.10487 [hep-th].

[4] K. G. Wilson and J. B. Kogut, "The Renormalization group and the epsilon expansion," *Phys. Rept.* **12** (1974) 75–200.

[5] L. D. Landau, "On the theory of phase transitions. I.," *Zh. Eksp. Teor. Fiz.* **11** (1937) 19.

[6] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," *Zh. eksp. teor. Fiz* **20** no. 1064-1082, (1950) 35.

[7] Y. Nambu, "Quasiparticles and Gauge Invariance in the Theory of Superconductivity," *Phys. Rev.* **117** (1960) 648–663.

[8] J. Goldstone, "Field Theories with Superconductor Solutions," *Nuovo Cim.* **19** (1961) 154–164.

[9] G. 't Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking," in *Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979*, vol. 59, pp. 135–157. 1980.

[10] A. Kapustin and R. Thorngren, "Anomalies of discrete symmetries in three dimensions and group cohomology," *Phys. Rev. Lett.* **112** no. 23, (2014) 231602, arXiv:1403.0617 [hep-th].

[11] A. Kapustin and R. Thorngren, "Anomalies of discrete symmetries in various dimensions and group cohomology," arXiv:1404.3230 [hep-th].

[12] A. J. Niemi and G. W. Semenoff, "Axial Anomaly Induced Fermion Fractionization and Effective Gauge Theory Actions in Odd Dimensional Space-Times," *Phys. Rev. Lett.* **51** (1983) 2077.

[13] A. N. Redlich, "Gauge Noninvariance and Parity Violation of Three-Dimensional Fermions," *Phys. Rev. Lett.* **52** (1984) 18.

[14] A. N. Redlich, "Parity Violation and Gauge Noninvariance of the Effective Gauge Field Action in Three-Dimensions," *Phys. Rev.* **D29** (1984) 2366–2374.

[15] A. Kapustin and R. Thorngren, "Higher symmetry and gapped phases of gauge theories," arXiv:1309.4721 [hep-th].

[16] A. Kapustin and N. Seiberg, "Coupling a QFT to a TQFT and Duality," *JHEP* **04** (2014) 001, arXiv:1401.0740 [hep-th].

[17] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, "Generalized Global Symmetries," *JHEP* **02** (2015) 172, arXiv:1412.5148 [hep-th].

[18] E. Witten, "Fermion Path Integrals And Topological Phases," *Rev. Mod. Phys.* **88** no. 3, (2016) 035001, arXiv:1508.04715 [cond-mat.mes-hall].

[19] N. Seiberg and E. Witten, "Gapped Boundary Phases of Topological Insulators via Weak Coupling," *Prog. Theor. Exp. Phys.* **2016** no. 12, (2016) 12C101, arXiv:1602.04251 [cond-mat.str-el].

[20] E. Witten, "The "Parity" Anomaly On An Unorientable Manifold," *Phys. Rev.* **B94** no. 19, (2016) 195150, arXiv:1605.02391 [hep-th].

[21] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, "Theta, Time Reversal, and Temperature," *JHEP* **05** (2017) 091, arXiv:1703.00501 [hep-th].

[22] C. Wang, A. Nahum, M. A. Metlitski, C. Xu, and T. Senthil, "Deconfined quantum critical points: symmetries and dualities," arXiv:1703.02426 [cond-mat.str-el].

[23] Z. Komargodski, A. Sharon, R. Thorngren, and X. Zhou, "Comments on Abelian Higgs Models and Persistent Order," arXiv:1705.04786 [hep-th].

[24] Z. Komargodski, T. Sulejmanpasic, and M. Ünsal, "Walls, Anomalies, and (De)Confinement in Quantum Anti-Ferromagnets," arXiv:1706.05731 [cond-mat.str-el].

[25] G. Y. Cho, C.-T. Hsieh, and S. Ryu, "Anomaly Manifestation of Lieb-Schultz-Mattis Theorem and Topological Phases," arXiv:1705.03892 [cond-mat.str-el].

[26] H. Shimizu and K. Yonekura, "Anomaly constraints on deconfinement and chiral phase transition," arXiv:1706.06104 [hep-th].

[27] J. Wang, X.-G. Wen, and E. Witten, "Symmetric Gapped Interfaces of SPT and SET States: Systematic Constructions," arXiv:1705.06728 [cond-mat.str-el].

[28] M. A. Metlitski and R. Thorngren, "Intrinsic and emergent anomalies at deconfined critical points," arXiv:1707.07686 [cond-mat.str-el].

[29] J. S. Bell and R. Jackiw, "A PCAC puzzle: pi0 −> gamma gamma in the sigma model," *Nuovo Cim.* **A60** (1969) 47–61.

[30] S. L. Adler, "Axial vector vertex in spinor electrodynamics," *Phys. Rev.* **177** (1969) 2426–2438.

[31] K. Fujikawa, "Path Integral Measure for Gauge Invariant Fermion Theories," *Phys. Rev. Lett.* **42** (1979) 1195–1198.

[32] K. Fujikawa, "Path Integral for Gauge Theories with Fermions," *Phys. Rev.* **D21** (1980) 2848. [Erratum: Phys. Rev.D22,1499(1980)].

[33] E. Witten, "An SU(2) Anomaly," *Phys. Lett.* **117B** (1982) 324–328.

[34] E. Witten, "GLOBAL GRAVITATIONAL ANOMALIES," *Commun. Math. Phys.* **100** (1985) 197.

[35] M. F. Atiyah, V. K. Patodi, and I. M. Singer, "Spectral asymmetry and Riemannian Geometry I," *Math. Proc. Cambridge Phil. Soc.* **77** (1975) 43.

[36] M. F. Atiyah, V. K. Patodi, and I. M. Singer, "Spectral asymmetry and Riemannian geometry II," *Math. Proc. Cambridge Phil. Soc.* **78** (1976) 405.

[37] M. F. Atiyah, V. K. Patodi, and I. M. Singer, "Spectral asymmetry and Riemannian geometry. III," *Math. Proc. Cambridge Phil. Soc.* **79** (1976) 71–99.

[38] M. F. Atiyah and I. M. Singer, "The index of elliptic operators on compact manifolds," *Bull. Am. Math. Soc.* **69** (1969) 422–433.

[39] M. F. Atiyah and I. M. Singer, "The Index of elliptic operators. 1," *Annals Math.* **87** (1968) 484–530.

[40] M. F. Atiyah and G. B. Segal, "The Index of elliptic operators. 2.," *Annals Math.* **87** (1968) 531–545.

[41] M. F. Atiyah and I. M. Singer, "The Index of elliptic operators. 3.," *Annals Math.* **87** (1968) 546–604.

[42] M. F. Atiyah and I. M. Singer, "The Index of elliptic operators. 4," *Annals Math.* **93** (1971) 119–138.

[43] M. F. Atiyah and I. M. Singer, "The Index of elliptic operators. 5.," *Annals Math.* **93** (1971) 139–149.

[44] M. F. Atiyah and I. M. Singer, "Dirac Operators Coupled to Vector Potentials," *Proc. Nat. Acad. Sci.* **81** (1984) 2597–2600.

[45] L. Alvarez-Gaume, S. Della Pietra, and G. W. Moore, "Anomalies and Odd Dimensions," *Annals Phys.* **163** (1985) 288.

[46] A. J. Niemi and G. W. Semenoff, "Spectral asymmetry on an open space," *Phys. Rev.* **D30** (1984) 809.

[47] C. G. Callan, Jr. and J. A. Harvey, "Anomalies and Fermion Zero Modes on Strings and Domain Walls," *Nucl. Phys.* **B250** (1985) 427–436.

[48] Y. Frishman, A. Schwimmer, T. Banks, and S. Yankielowicz, "The Axial Anomaly and the Bound State Spectrum in Confining Theories," *Nucl. Phys.* **B177** (1981) 157–171.

[49] S. R. Coleman and B. Grossman, "'t Hooft's Consistency Condition as a Consequence of Analyticity and Unitarity," *Nucl. Phys.* **B203** (1982) 205–220.

[50] R. F. Dashen, "Some features of chiral symmetry breaking," *Phys. Rev.* **D3** (1971) 1879–1889.

[51] E. Witten, "Large N Chiral Dynamics," *Annals Phys.* **128** (1980) 363.

[52] G. 't Hooft, "Topology of the Gauge Condition and New Confinement Phases in Nonabelian Gauge Theories," *Nucl. Phys.* **B190** (1981) 455–478.

[53] N. Ohta, "Vacuum Structure and Chiral Charge Quantization in the Large N Limit," *Prog. Theor. Phys.* **66** (1981) 1408. [Erratum: Prog. Theor. Phys.67,993(1982)].

[54] J. L. Cardy and E. Rabinovici, "Phase Structure of Z(p) Models in the Presence of a Theta Parameter," *Nucl. Phys.* **B205** (1982) 1–16.

[55] J. L. Cardy, "Duality and the Theta Parameter in Abelian Lattice Models," *Nucl. Phys.* **B205** (1982) 17–26.

[56] U. J. Wiese, "Numerical Simulation of Lattice θ Vacua: The 2-d U(1) Gauge Theory as a Test Case," *Nucl. Phys.* **B318** (1989) 153–175.

[57] I. Affleck, "Nonlinear sigma model at Theta = pi: Euclidean lattice formulation and solid-on-solid models," *Phys. Rev. Lett.* **66** (1991) 2429–2432.

[58] M. Creutz, "Quark masses and chiral symmetry," *Phys. Rev.* **D52** (1995) 2951–2959, `arXiv:hep-th/9505112 [hep-th]`.

[59] M. Creutz, "Anomalies and chiral symmetry in QCD," *Annals Phys.* **324** (2009) 1573–1584, `arXiv:0901.0150 [hep-ph]`.

[60] A. V. Smilga, "QCD at theta similar to pi," *Phys. Rev.* **D59** (1999) 114021, `arXiv:hep-ph/9805214 [hep-ph]`.

[61] E. Witten, "Theta dependence in the large N limit of four-dimensional gauge theories," *Phys. Rev. Lett.* **81** (1998) 2862–2865, `arXiv:hep-th/9807109 [hep-th]`.

[62] I. E. Halperin and A. Zhitnitsky, "Anomalous effective Lagrangian and theta dependence in QCD at finite N(c)," *Phys. Rev. Lett.* **81** (1998) 4071–4074, `arXiv:hep-ph/9803301 [hep-ph]`.

[63] K. Mameda, "QCD θ-vacua from the chiral limit to the quenched limit," *Nucl. Phys.* **B889** (2014) 712–726, `arXiv:1408.1189 [hep-ph]`.

[64] R. Thorngren and C. von Keyserlingk, "Higher SPT's and a generalization of anomaly in-flow," `arXiv:1511.02929 [cond-mat.str-el]`.

[65] S. Weinberg, *The Quantum Theory of Fields, Volume 2: Modern Applications*. Cambridge University Press, 5, 2005.

[66] J. A. Harvey, "TASI 2003 lectures on anomalies," 2005. `arXiv:hep-th/0509097 [hep-th]`.

[67] R. Stora, "Algebraic Structure and Topological Origin of Anomalies," in *Progress in Gauge Field Theory. Proceedings, NATO Advanced Study Institute, Cargese, France*. 1983.

[68] B. Zumino, "Chiral Anomalies and Differential Geometry," in *Relativity, groups and topology: Proceedings, 40th Summer School of Theoretical Physics - Session 40: Les Houches, France, June 27 - August 4, 1983, vol. 2*, pp. 1291–1322. 1983. `http://inspirehep.net/record/192970/files/5461614.pdf`.

[69] G. Y. Cho, J. C. Y. Teo, and S. Ryu, "Conflicting Symmetries in Topologically Ordered Surface States of Three-dimensional Bosonic Symmetry Protected Topological Phases," *Phys. Rev.* **B89** no. 23, (2014) 235103, arXiv:1403.2018 [cond-mat.str-el].

[70] C. Vafa and E. Witten, "Restrictions on Symmetry Breaking in Vector-Like Gauge Theories," *Nucl. Phys.* **B234** (1984) 173–188.

[71] F. Sannino, "A Note on anomaly matching for finite density QCD," *Phys. Lett.* **B480** (2000) 280–286, arXiv:hep-ph/0002277 [hep-ph].

[72] S. D. H. Hsu, F. Sannino, and M. Schwetz, "Anomaly matching in gauge theories at finite matter density," *Mod. Phys. Lett.* **A16** (2001) 1871–1880, arXiv:hep-ph/0006059 [hep-ph].

[73] "Quantised singularities in the electromagnetic field," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **133** no. 821, (1931) 60–72.

[74] J. S. Schwinger, "Magnetic charge and quantum field theory," *Phys. Rev.* **144** (1966) 1087–1093.

[75] D. Zwanziger, "Quantum field theory of particles with both electric and magnetic charges," *Phys. Rev.* **176** (1968) 1489–1495.

[76] D. Zwanziger, "Local Lagrangian quantum field theory of electric and magnetic charges," *Phys. Rev.* **D3** (1971) 880.

[77] E. Witten, "Dyons of Charge $e\theta/2\pi$," *Phys. Lett.* **B86** (1979) 283–287.

[78] O. Aharony, N. Seiberg, and Y. Tachikawa, "Reading between the lines of four-dimensional gauge theories," *JHEP* **08** (2013) 115, arXiv:1305.0318 [hep-th].

[79] P. Goddard, J. Nuyts, and D. I. Olive, "Gauge Theories and Magnetic Charge," *Nucl. Phys.* **B125** (1977) 1–28.

[80] D. Simmons-Duffin, "The Conformal Bootstrap," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 1–74. 2017. arXiv:1602.07982 [hep-th]. https://inspirehep.net/record/1424282/files/arXiv:1602.07982.pdf.

[81] J. M. Maldacena, G. W. Moore, and N. Seiberg, "D-brane charges in five-brane backgrounds," *JHEP* **10** (2001) 005, arXiv:hep-th/0108152 [hep-th].

[82] T. Banks and N. Seiberg, "Symmetries and Strings in Field Theory and Gravity," *Phys. Rev.* **D83** (2011) 084019, arXiv:1011.5120 [hep-th].

[83] A. M. Polyakov, "Compact Gauge Fields and the Infrared Catastrophe," *Phys. Lett.* **B59** (1975) 82–84.

[84] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, "Pseudoparticle Solutions of the Yang-Mills Equations," *Phys. Lett.* **B59** (1975) 85–87.

[85] G. 't Hooft, "Symmetry Breaking Through Bell-Jackiw Anomalies," *Phys. Rev. Lett.* **37** (1976) 8–11.

[86] C. G. Callan, Jr., R. F. Dashen, and D. J. Gross, "The Structure of the Gauge Theory Vacuum," *Phys. Lett.* **B63** (1976) 334–340.

[87] R. Jackiw and C. Rebbi, "Vacuum Periodicity in a Yang-Mills Quantum Theory," *Phys. Rev. Lett.* **37** (1976) 172–175.

[88] S. R. Coleman, "More About the Massive Schwinger Model," *Annals Phys.* **101** (1976) 239.

[89] C. A. Baker *et al.*, "An Improved experimental limit on the electric dipole moment of the neutron," *Phys. Rev. Lett.* **97** (2006) 131801, arXiv:hep-ex/0602020 [hep-ex].

[90] D. Boer and J. K. Boomsma, "Spontaneous CP-violation in the strong interaction at theta = pi," *Phys. Rev.* **D78** (2008) 054027, arXiv:0806.1669 [hep-ph].

[91] S. Aoki and M. Creutz, "Pion Masses in Two-Flavor QCD with η Condensation," *Phys. Rev. Lett.* **112** no. 14, (2014) 141603, arXiv:1402.1837 [hep-lat].

[92] J. J. M. Verbaarschot and T. Wettig, "Dirac spectrum of one-flavor QCD at $\theta = 0$ and continuity of the chiral condensate," *Phys. Rev.* **D90** no. 11, (2014) 116004, arXiv:1407.8393 [hep-th].

[93] Y. Tachikawa, "Magnetic discrete gauge field in the confining vacua and the supersymmetric index," *JHEP* **03** (2015) 035, arXiv:1412.2830 [hep-th].

[94] G. 't Hooft, "On the Phase Transition Towards Permanent Quark Confinement," *Nucl. Phys.* **B138** (1978) 1–25.

[95] S. Gukov and E. Witten, "Rigid Surface Operators," *Adv. Theor. Math. Phys.* **14** no. 1, (2010) 87–178, arXiv:0804.1561 [hep-th].

[96] Y. Tanizaki, T. Misumi, and N. Sakai, "Circle compactification and 't Hooft anomaly," arXiv:1710.08923 [hep-th].

[97] Y. Nambu, "Strings, Monopoles and Gauge Fields," *Phys. Rev.* **D10** (1974) 4262.

[98] S. Mandelstam, "Vortices and Quark Confinement in Nonabelian Gauge Theories," in *Phys. Rep. 23 (1976) 245-249, In *Gervais, J.L. (Ed.), Jacob, M. (Ed.): Non-linear and Collective Phenomena In Quantum Physics*, 12-16*, vol. 23, pp. 245–249. 1976.

[99] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, "QCD and Instantons at Finite Temperature," *Rev. Mod. Phys.* **53** (1981) 43.

[100] B. Svetitsky, "Symmetry Aspects of Finite Temperature Confinement Transitions," *Phys. Rept.* **132** (1986) 1–53.

[101] S. P. Klevansky, "The Nambu-Jona-Lasinio model of quantum chromodynamics," *Rev. Mod. Phys.* **64** (1992) 649–708.

[102] T. Hatsuda and T. Kunihiro, "QCD phenomenology based on a chiral effective Lagrangian," *Phys. Rept.* **247** (1994) 221–367, arXiv:hep-ph/9401310 [hep-ph].

[103] H. Meyer-Ortmanns, "Phase transitions in quantum chromodynamics," *Rev. Mod. Phys.* **68** (1996) 473–598, arXiv:hep-lat/9608098 [hep-lat].

[104] A. V. Smilga, "Physics of thermal QCD," *Phys. Rept.* **291** (1997) 1–106, arXiv:hep-ph/9612347 [hep-ph].

[105] K. Rajagopal and F. Wilczek, "The Condensed matter physics of QCD," in *At the frontier of particle physics. Handbook of QCD. Vol. 1-3*, M. Shifman and B. Ioffe, eds., pp. 2061–2151. 2000. arXiv:hep-ph/0011333 [hep-ph].

[106] D. H. Rischke, "The Quark gluon plasma in equilibrium," *Prog. Part. Nucl. Phys.* **52** (2004) 197–296, arXiv:nucl-th/0305030 [nucl-th].

[107] M. G. Alford, A. Schmitt, K. Rajagopal, and T. SchÃďfer, "Color superconductivity in dense quark matter," *Rev. Mod. Phys.* **80** (2008) 1455–1515, arXiv:0709.4635 [hep-ph].

[108] R. S. Hayano and T. Hatsuda, "Hadron properties in the nuclear medium," *Rev. Mod. Phys.* **82** (2010) 2949, arXiv:0812.1702 [nucl-ex].

[109] H. Satz, "The States of Matter in QCD," arXiv:0903.2778 [hep-ph].

[110] M. Huang, "QCD Phase Diagram at High Temperature and Density," arXiv:1001.3216 [hep-ph].

[111] A. Schmitt, "Dense matter in compact stars: A pedagogical introduction," *Lect. Notes Phys.* **811** (2010) 1–111, arXiv:1001.3294 [astro-ph.SR].

[112] K. Fukushima and T. Hatsuda, "The phase diagram of dense QCD," *Rept. Prog. Phys.* **74** (2011) 014001, arXiv:1005.4814 [hep-ph].

[113] A. Cherman, S. Sen, M. Unsal, M. L. Wagman, and L. G. Yaffe, "Order parameters and color-flavor center symmetry in QCD," arXiv:1706.05385 [hep-th].

[114] H. Kouno, Y. Sakai, T. Makiyama, K. Tokunaga, T. Sasaki, and M. Yahiro, "Quark-gluon thermodynamics with the Z(N(c)) symmetry," *J. Phys.* **G39** (2012) 085010.

[115] Y. Sakai, H. Kouno, T. Sasaki, and M. Yahiro, "The quarkyonic phase and the Z_{N_c} symmetry," *Phys. Lett.* **B718** (2012) 130–135, arXiv:1204.0228 [hep-ph].

[116] H. Kouno, T. Makiyama, T. Sasaki, Y. Sakai, and M. Yahiro, "Confinement and $\mathbb{Z}_3$ symmetry in three-flavor QCD," *J. Phys.* **G40** (2013) 095003, arXiv:1301.4013 [hep-ph].

[117] H. Kouno, T. Misumi, K. Kashiwa, T. Makiyama, T. Sasaki, and M. Yahiro, "Differences and similarities between fundamental and adjoint matters in SU(N) gauge theories," *Phys. Rev.* **D88** no. 1, (2013) 016002, arXiv:1304.3274 [hep-ph].

[118] E. Poppitz and T. Sulejmanpasic, "(S)QCD on $\mathbb{R}^3 \times \mathbb{S}^1$: Screening of Polyakov loop by fundamental quarks and the demise of semi-classics," *JHEP* **09** (2013) 128, arXiv:1307.1317 [hep-th].

[119] T. Iritani, E. Itou, and T. Misumi, "Lattice study on QCD-like theory with exact center symmetry," *JHEP* **11** (2015) 159, arXiv:1508.07132 [hep-lat].

[120] H. Kouno, K. Kashiwa, J. Takahashi, T. Misumi, and M. Yahiro, "Understanding QCD at high density from a Z_3-symmetric QCD-like theory," *Phys. Rev.* **D93** no. 5, (2016) 056009, arXiv:1504.07585 [hep-ph].

[121] T. Hirakida, H. Kouno, J. Takahashi, and M. Yahiro, "Interplay between sign problem and Z_3 symmetry in three-dimensional Potts models," *Phys. Rev.* **D94** no. 1, (2016) 014011, arXiv:1604.02977 [hep-lat].

[122] T. Hirakida, J. Sugano, H. Kouno, J. Takahashi, and M. Yahiro, "Sign problem in Z_3-symmetric effective Polyakov-line model," *Phys. Rev.* **D96** no. 7, (2017) 074031, arXiv:1705.00665 [hep-lat].

[123] A. Roberge and N. Weiss, "Gauge Theories With Imaginary Chemical Potential and the Phases of QCD," *Nucl. Phys.* **B275** (1986) 734–745.