Imperial College London

Lines, Groups, and Topology in SU(N) and PSU(N) Yang-Mills Theory

Supervisor:
Prof. Chris Hull FRS

Author:
Thomas Yan

Theoretical Physics Group
Department of Physics
Imperial College London

Submitted in part fulfilment of the requirements for the degree of Master of Science of Imperial College London

Abstract

Generalized symmetry is a relatively new field of study that has coined new terms in the century-old discussion of gauge theory. In this paper we will examine the interplay between Yang-Mills theory and generalized symmetry through various mathematical tools. The paper loosely centers around the line spectrum of $SU(N)$ and $PSU(N)$ Yang-Mills theories in 4d with theta term turned on. The basics of generalized symmetry and gauge theory are introduced in the first two chapters including sections on gauging global symmetry and topological monopole. The remaining chapters examine the Yang-Mills duo with Lie theory and discrete gauge background field.

Acknowledgements

I want to express my gratitude to my family for their financial and mental support throughout my academic journey. Without their full trust it would have never started.

I am very grateful to Professor Chris Hull for the amazing opportunity to work on this project. He has provided immense help to us through lectures, organized sessions, and discussions.

I want to extend my special thanks to Max. His unlimited support during both differential geometry course and this project is truly irreplaceable. His passion and work ethic has been aspiring to all of us. I hope we will get to work together again.

I want to thank my batchmates Javier and Thomas for their company throughout the year and help with problems/exams. This year would be so much more stressful without them.

I also want to thank the group members of Generalized Symmetry 2023, Zizang and Adi, for all the fruitful discussions and heartfelt encouragements during a very meaningful few months.

Lastly, I want to thank the entire theory group at Imperial for inducting me into the world of theoretical physics.

Palestine, Jerusalem, and the Temple severally and concurrently represent the image of the universe and the Center of the World. This multiplicity of centers and this reiteration of the image of the world on smaller and smaller scales constitute one of the specific characteristics of traditional societies.

To us, it seems an inescapable conclusion that the religious man sought to live as near as possible to the Center of the World.

— Mircea Eliade, *The Sacred and the Profane*

Contents

Chapter 1

Introduction

According to [1], higher-form gauge theories were first discussed by [29] in the 70s. Since then they have been a stable in physicists' discourse. They showed up in various fields of theoretical physics such as Chern-Simons theory, rational CFTs, lattice gauge theories, supergravity and string theory. The list of literature goes on ([30], [31], etc.). Most of the earlier studies focused on a local gauged version of the higher-form fields. A wave of discussion started by the seminal work [1] shifted attention to the global symmetry of higher-dimensional objects and its consequence. In recent literature, generalized symmetry has further expanded the language of physics by bringing in concepts from category theory and higher-group to explain non-invertible symmetries.

This dissertation will not go into the territory of non-invertible symmetries. Instead, we will focus on the interplay of generalized symmetry and gauge theory, or more specifically, Yang-Mills theory. The paper loosely centers around the line spectrum of $SU(N)$ and $PSU(N)$ Yang-Mills theories in 4d with theta term turned on. We will derive some of their properties through different mathematical frameworks.

In chapter 2, we will start with Noether's theorem and ordinary (0-form) global symmetry. From there we will re-envision the unitary action of global symmetry as topological operator. We will use the topological nature of the operator to bend it into shape that helps us understand its property, and we will extend the notion we have developed with 0-form symmetry to p-form symmetry with arbitrary degree of p. We will finish with a concrete example of 1-form symmetries rooted in Maxwell theory.

In chapter 3, we begins with a crash course on principal fibre bundle, the mathematical construction behind gauge theory. We will derive the gauge transformation of connection and explain what "gauge" means on a mathematical stand point. We will then use the fibre bundle picture to construct various topologically non-trivial objects in gauge theory, including Dirac monopole and $SU(2)$ winding number. Alongside we will introduce the role of theta term in Yang-Mills theory. The chapter finishes with a generalization of the process of gauging global symmetry to local symmetry.

In chapter 4 we will mainly focus on the line operators of Yang-Mills theory. We will start by deriving the Wilson Line operator with a fibre bundle setup, followed by a discussion of line screening in 1-form symmetry. Seeking to extend the discussion to magnetic lines, the remaining of the chapter focuses on two specific group: $SU(N)$ and

$PSU(N)$. We will show that it is possible to obtain a classification for magnetic charges and line spectrum in the two theories via abstract derivations with Lie theory. The chapter finishes with a concrete example of $SU(2)$ Yang-Mills and its centerless form $SO(3)$. An important result is the discretization of theta angle in $PSU(N)$ theory.

Finally, in chapter 5 we seek to reproduce some of the results in chapter 4 by gauging the center with discrete gauge fields. We will first need to introduce the mathematical background for discrete gauge fields: algebraic topology. After constructing a discrete gauge field from cohomology, we will discuss how it can be conveniently expressed in the form of $U(1)$ gauge field. The chapter finishes with an explicit calculation of $PSU(N)$ action and its theta term. The discretization of theta angle comes out as a result of the discrete period of $\mathbb{Z}_N$ field.

Chapter 2

Generalized Global Symmetry

In this chapter, we will develop a generalized formal language of global symmetry in Quantum Field Theory (QFT) starting with a review on ordinary global symmetry. We will then extend the formal language to include p-form symmetry and illustrate the topological nature of its operators. It is followed by a physical example of dual 1-form symmetries in Maxwell Theory. The chapter mostly traces the footstep of [1] that formulated and introduced the framework. The detailed treatments are heavily inspired by some of the recent lecture notes [10],[9] and [11] and two series of lecture videos [33], [32]. It should be noted that for the purpose of this paper, we will only consider group-like symmetries that are always invertible. The non-invertible side of the story is a burgeoning topic in theoretical discussion and provides interesting link between QFT and category theory. A good resource for this topic would be [8].

2.1 Continuous 0-Form Global Symmetry

We will start, in a typical fasion, with a review of ordinary symmetry we all know and love. Consider a d-dimensional QFT with an action and a symmetry group $G^{(0)}$. The notation $^{(0)}$ marks the degree of the symmetry. We will soon see what it means. Immediately following from Noether's theorem we can find the conservation of Noether current and a conserved charge defined by integrating time component of the current over a spatial slice:

$$\partial_\mu j^{\mu a} = 0 \qquad Q^a(t) = \int d^{d-1}\vec{x}\; j^{0a}(t, \vec{x}) \tag{2.1}$$

where a running from 1 to $dim(G^{(p)})$ is the group index. A classical result is that $Q^a(t)$ are also generators of the symmetry group, a proof can be found in [6]. We will now quantize the system so that the generator becomes an operator on the Hilbert space. It acts on a local operator $\mathcal{O}(x)$ through a commutator bracket:

$$[\theta^a Q^a(t), \mathcal{O}(t, \vec{x})] = \delta_\theta \mathcal{O}(t, \vec{x})$$

In other words, it changes $\mathcal{O}$ by an infinitesimal amount $\delta\mathcal{O}$. Such expression is obviously problematic when we later discuss a discrete group action. In some sense we can say that the charge is not the most fundamental object associated with a group structure. Instead we should construct a unitary operator out of the generator:

$$U_\theta(t) = \exp\left(i\theta^a Q^a\right) = \exp\left(i\theta^a \int d^{d-1}\vec{x}\, j^{0a}(t,\vec{x})\right) \tag{2.2}$$

Here θ^a is the parameter marking an element of group $G^{(0)}$:

$$e^{i\theta^a T^a} = g \in G^{(0)}$$

so we can equivalently denote U_θ as U_g. Notice that $U_\theta(t)$ is manifestedly unitary and therefore invertible. It acts on a local operator in the familiar form:

$$U_\theta(t)\mathcal{O}(t,\vec{x})U_\theta^\dagger(t) = \mathcal{O}'(t,\vec{x}) \tag{2.3}$$

We say that the operator $\mathcal{O}$, which undergoes transformation of U, is charged under the symmetry. Following from [10], we will call U_θ a Symmetry Defect Operator (SDO) of the symmetry. An alternative view is that the operator $U_\theta(t)$ is essentially equivalent to inserting U_θ every point on a spatial slice of the space-time at t, Σ_t (Fig.2.1). We can see this more clearly by switching to a geometric notation using form technique. We can rewrite 2.1, 2.2 as follow:

$$d * J_1^a = 0 \qquad Q^a(\Sigma_t) = \int_{\Sigma_t} *J_1^a \tag{2.4}$$

$$U_\theta(\Sigma_t) = \exp\left(i\theta^a Q^a\right) = \exp\left(i\theta^a \int_{\Sigma_t} *J_1^a\right) \tag{2.5}$$

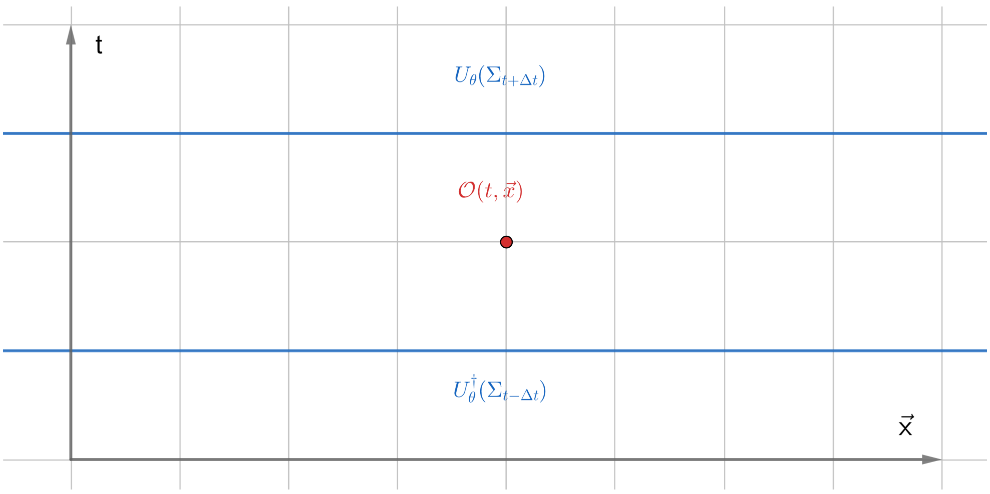

Figure 2.1: Symmetry actions as topological operators

J_1^a is now the 1-form associated to the current and the subscript denotes the dimension of the form. Differential Geometry has provided us with a natural way of integrating (d-1)-form $*J_1^a$ on the (d-1)-dim submanifold Σ_t. We now rewrite 2.3 as:

$$U_\theta(t)\mathcal{O}(t,\vec{x})U_\theta^\dagger(t) = \lim_{\Delta t \to 0} U_\theta(\Sigma_{t+\Delta t})\mathcal{O}(t,\vec{x})U_\theta^\dagger(\Sigma_{t-\Delta t}) \tag{2.6}$$

Some new insight would be gained if we see the above expression as **operators within a path integral**, $\langle U_\theta(t)\mathcal{O}(t,\vec{x})U_\theta^\dagger(t)\rangle$, instead of acting on a Hilbert space. Since path integrals are automatically time-ordered, it makes sense to insert U_θ on any (d-1)-dim surface instead of a spatial slice restricted to time t. For the same reason, the order of operators doesn't matter anymore. The revelation comes when we try to find an expression for U_θ on arbitrary surface. It follows that for a (d-1)-dim submanifold Σ_{d-1} deformed infinitesimally to Σ'_{d-1} by whipping through a d-dim submanifold $\tilde{\Sigma}_d$:

$$\begin{aligned}
U_\theta(\Sigma'_{d-1}) &= \exp\left(i\theta\int_{\Sigma'_{d-1}} *J_1^a\right)\\
&= \exp\left(i\theta\left(\int_{\Sigma_{d-1}} *J_1^a + \int_{\Sigma'_{d-1}-\Sigma_{d-1}} *J_1^a\right)\right)\\
&= U_\theta(\Sigma_{d-1})\exp\left(i\theta\int_{\partial\tilde{\Sigma}_d} *J_1^a\right)\\
&= U_\theta(\Sigma_{d-1})\exp\left(i\theta\int_{\tilde{\Sigma}_d} d*J_1^a\right)\\
&= U_\theta(\Sigma_{d-1})
\end{aligned}$$

We have used 2.4 and Stocke's theorem in the derivation, and the notation $\partial\tilde{\Sigma}_d$ refers to the boundary of $\tilde{\Sigma}_d$. The equation illustrates the **Topological Nature** of the SDO. $U_\theta(\Sigma)$ remains unchanged under a deformation of Σ that does not cross over any charged object. Such deformations are called topological deformations.

2.1.1 Linking Picture and Fusion Rule

We can now rewrite 2.3 as

$$U_\theta(\Sigma_{d-1})\mathcal{O}(x)U_\theta^\dagger(\Sigma'_{d-1}) = \mathcal{O}' \tag{2.7}$$

where $\mathcal{O}$ should be inserted at a point x between Σ_{d-1} and Σ'_{d-1}. Also,

$$U_\theta^\dagger(\Sigma'_{d-1}) = U_\theta^{-1}(\Sigma'_{d-1}) = U_\theta(-\Sigma'_{d-1})$$

and 2.7 can be reformulated as

$$U_\theta(\Sigma_{d-1})\mathcal{O}(x)U_\theta(-\Sigma'_{d-1}) = U_\theta(\Sigma_{d-1}-\Sigma'_{d-1})\mathcal{O}(x)$$

With a proper boundary condition at spatial infinity, the surface $\Sigma_{d-1}-\Sigma'_{d-1}$ can be connected at infinity to form a close surface without affecting the path integral. We now use the topological nature of the operator to shrink the surface to proximity of $\mathcal{O}$ and for the sake of simplicity, make it a sphere. The final form of 2.7 follows:

$$U_\theta(S_{d-1})\mathcal{O}(x) = \mathcal{O}'(x) \tag{2.8}$$

The action can be visualized by Fig.2.2, taken from [9]. We can equivalently think of 2.8 as passing $U_\theta(S_{d-1})$ through $\mathcal{O}(x)$ leaving $\mathcal{O}'(x)$ and $U_\theta(S'_{d-1})$ on an empty sphere.

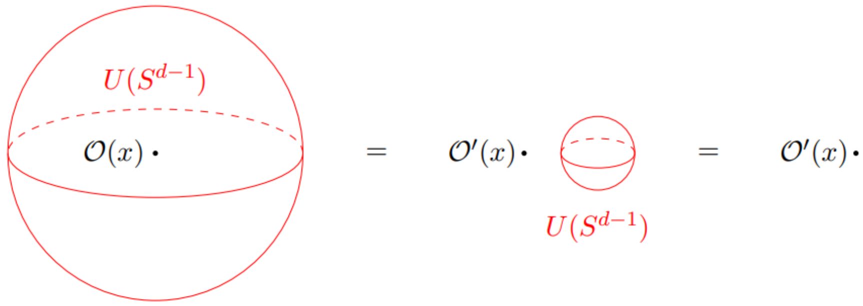

Figure 2.2: Action of symmetry as linking

The sphere must then shrink to a trivial operator. It is possible to find $\mathcal{O}'(x)$ explicitly. In group theory language, $\mathcal{O}$ is acted on by the group with left action $(\cdot)$ and therefore "in representation" of the group. We can denote it as $\mathcal{O}_R$ where R is some representation $R : G^{(0)} \to End(V)$ and this action is formalized as:

$$g \cdot \mathcal{O}_R \coloneqq R(g)\mathcal{O}_R$$

Recall that Ward identities can be presented as follow:

$$d * J^a_p = 0 \tag{2.9}$$

$$d * J^a_p \mathcal{O}_R(M) = R(T^a)\delta_{d-p+1}(x \in M)\mathcal{O}_R(M) \tag{2.10}$$

$$d * J^a_p \mathcal{O}_R(M^1)...\mathcal{O}_R(M^n) = 0\ (1 < n \in \mathbb{Z}) \tag{2.11}$$

where $M, M^1...$ are submanifolds that $\mathcal{O}_R$ lives on. $\delta_{d-p+1}(x \in M)$ is the (d-p+1)-form Poincaré dual of a delta function. These are a series of operator identities that hold true inside a path integral. The first one is a direct correspondence to the classical conservation of current and we have used it earlier without deliberation. For a derivation of Ward identities, see [15]. It is now straight forward to write

$$\begin{aligned}
U_\theta(S_{d-1})\mathcal{O}_R(x) &= \mathcal{O}'_R(x) \\
&= \exp\left(i\theta^a \int_{S_{d-1}} *J^a_1\right)\mathcal{O}_R(x) \\
&= \exp\left(i\theta^a \int_{D_d} d * J^a_1\right)\mathcal{O}_R(x) \\
&= \exp\left(i\theta^a \int_{D_d} R(T^a)\delta_d(x)\right)\mathcal{O}_R(x) \\
&= \exp\left(i\theta^a R(T^a)\right)\mathcal{O}_R(x) = R(g)\mathcal{O}_R
\end{aligned}$$

where D_d is the "inside" of the sphere S_{d-1}. We can see that an SDO linking to a charged operator simply performs its associated group action on the charged operator.

The linking picture is useful because it naturally suggests the kind of operator that can be charged under a certain symmetry. In precise mathematical vocabulary we say that the point x and the sphere S_{d-1} are two linking submanifolds. We will discuss the notion further when we go to higher-form symmetries later.

While SDO acts on charged operator via linking, there is also an algebra formed between two SDOs. Notice that SDOs are marked by elements of the group and therefore beg the existence of a map:

$$U_{g_1} \otimes U_{g_2} = U_{g_1 \times g_2} \tag{2.12}$$

where $g_1 \times g_2$ is the group multiplication. This map obviously obeys the group structure and is called the fusion rule. Notice that the fusion rule is manifested by acting U_{g_1} and U_{g_2} on a charged operator consecutively. See Fig.2.3 below, taken from [9].

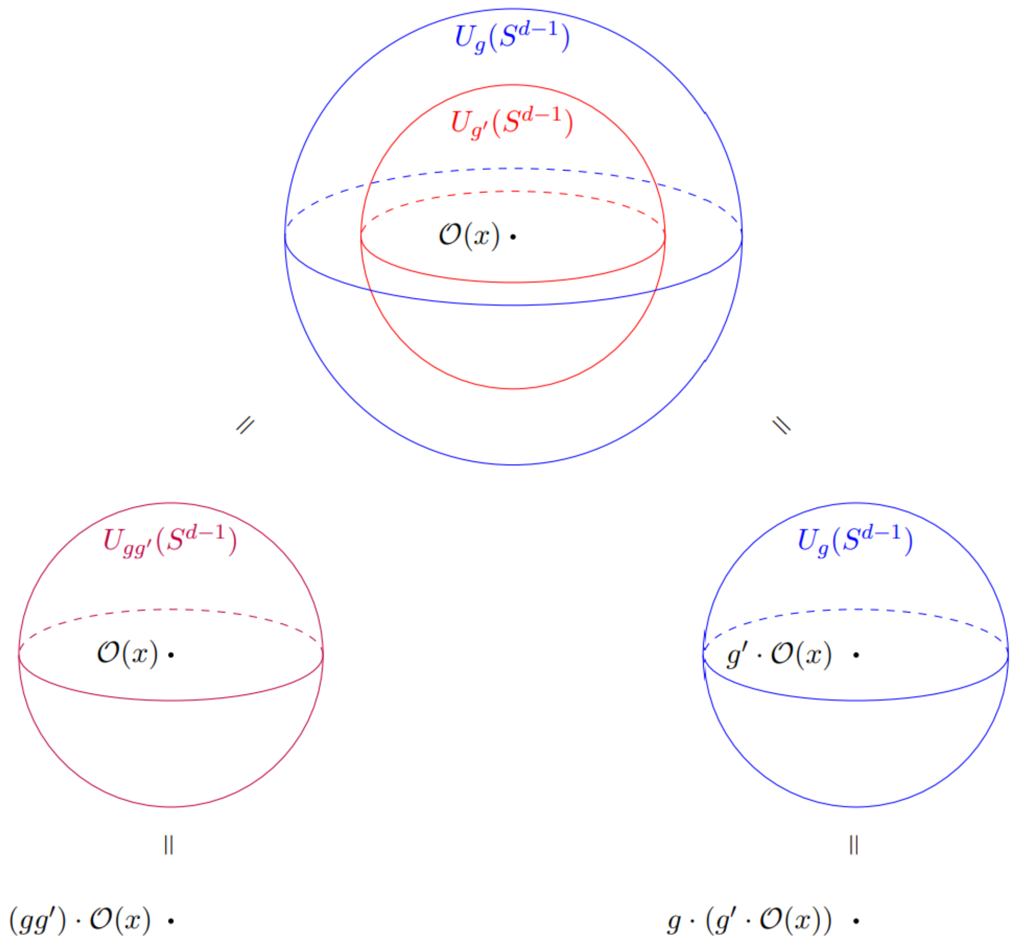

Figure 2.3: Caption:Two ways of thinking about fusion rule

$$\begin{aligned} U_{g_2}(S_{(g_2)d-1})U_{g_1}(S_{(g_1)d-1})\mathcal{O}_R(x) &= R(g_1)U_{g_2}(S_{(g_2)d-1})\mathcal{O}_R(x) \\ &= R(g_1)R(g_2)\mathcal{O}_R(x) \\ &= R(g_1 \times g_2)\mathcal{O}_R(x) = U_{g_1 \times g_2}(S_{d-1})\mathcal{O}_R(x) \end{aligned}$$

where in the last line we have used the homomorphic property of representation R. We can visually understand the fusion rule as U_{θ_1} and U_{θ_2} fusing together before acting on the charged operator. There are a few subtle points in our description. The charged operator is required to be in linear representation of SDO but the Hilbert Space is not. We allow the symmetry to be realized projectively on the states (think about spin representation). Also, SDOs are not required to act faithfully on the charged operator. ([32]).

The fusion rule seems like a harmless group action so far, but it is more general than that. In cases where there is categorical or higher-group structure in the theory, the fusion rule of two operator could involve non-invertible transformation, which clearly violates the group axioms. See [8] for more information.

2.1.2 Special Case: U(1) Ordinary Symmetry

A specific and arguably more useful case of this generalized language is when $G^{(0)} \cong U(1)$. The representation theory of $U(1)$ is relatively simple. We can denote an element of the group as

$$e^{i\alpha} \in U(1) \qquad \alpha \in [0, 2\pi)$$

and therefore SDO will be marked by a parameter α:

$$U_\alpha(\Sigma_{d-1}) = \exp\left(i\alpha Q\right) = \exp\left(i\alpha \int_{\Sigma_{d-1}} *J_1\right)$$

Also, the representation of $U(1)$, denoted as $\phi(g)$, is marked by some integer. This is a result following from Schur's lemma. More explicitly:

$$\phi(g) \coloneqq \phi_q(g) = e^{i\alpha q},\ q \in \mathbb{Z}$$

Thus the charged operators are now equivalently denoted by an integer q, known as the "charge" of the operator. The action of the symmetry now is

$$U_\alpha(S_{d-1})\mathcal{O}_q(x) = e^{i\alpha q}\mathcal{O}_q(x)$$

i.e. it simply dresses the operator with a phase $e^{i\alpha q}$. There is a more elegant way to classify these charge operators and it would benefit us later to introduce the mathematical notation of it.

We define the set of all homomorphic maps:

$$\Phi : G \to U(1)$$

These maps themselves form another Abelian group under the multiplication of $U(1)$ group. For two maps $\phi, \phi' \in \Phi$:

$$(\phi \times_\Phi \phi')(g) \coloneqq \phi(g) \times_{U(1)} \phi'(g)$$

This group is called the **Pontryagin dual group** of the original group G and denoted as $\widehat{G}$. When the group G is Abelian, we claim without proving that the Pontryagin dual of G marks the irreps of G. Therefore **the charged operators under an Abelian symmetry G is marked by elements of the Pontryagin dual group $\widehat{G}$.**

For example, for $G \cong U(1)$ we know that $\widehat{G} \cong \mathbb{Z}$ i.e. the group of integer under addition. Therefore the charged operators of a $U(1)$ symmetry are marked by an integer.

The Pontryagin dual further has a nice property that $\widehat{\widehat{G}} \cong G$. The notion of Pontryagin dual works not only for continuous group but also for discrete group. It helps simplify the situation once we complicate the picture with Abelian or non-Abelian Gauge Group. Notice that this method only works for Abelian group since the Pontryagin dual of non-Abelian group does ont correspond to the irreps.

2.2 p-Form Global Symmetry

Generalization to p-form symmetry is simple with the formal language we just introduced. Assuming $G^{(p)} \cong U(1)$ instead of a 1-form current J_1 we now have a (p+1)-form J_{p+1}. Everything else follows exactly like the 0-form symmetry case we have studied. We have the list of essential equations:

$$d * J_{p+1} = 0 \qquad U_\alpha(\Sigma_{d-p-1}) = \exp\left(i\alpha \int_{\Sigma_{d-p-1}} *J_{p+1}\right) \qquad U_\alpha(S_{d-p-1})\mathcal{O}_q(\Gamma_p) = \mathcal{O}'_q(\Gamma_p) \tag{2.13}$$

Notice that the charged operator is p-dimensional and hence the name p-form symmetry. More specifically, a p-dimensional operator is the lowest-dimensional object that can be acted on by $U_\alpha(S_{d-p-1})$. To see this heuristically, consider the case where $d = 3$, $p = d - p - 1 = 1$. In this case, $U_\alpha(S_1)$ acts trivially on a point operator because we can always use the topological nature of SDO to move the S_1 circle away some the point so that it contracts to identity. The lowest-dimensional operator that can be charged under $U_\alpha(S_1)$ in 3d is a line. We haven't discussed how operator with $dim > p$ can be charged under p-form symmetry. This is a topic related to higher-representation and beyond the scope of this paper. A good introduction can be found in [8]. **From now on we will consider $dim = p$ operator as the only relevant operator charged under p-form symmetry.** We have hinted on a more precise definition of this action via linking number. To see this, let us write out the action of the symmetry more explicitly:

$$U_\alpha(S_{d-p-1})\mathcal{O}_q(\Gamma_p) = \mathcal{O}'_q(\Gamma_p) = \exp\left(i\alpha q Link(S_{d-p-1}, \Gamma_p)\right)\mathcal{O}_q(\Gamma_p) \tag{2.14}$$

This is follows directly from 2.10 with a twist that a p-dimensional charge operator can wrap around the SDO for more than once, reflected as the linking number $Link(S_{d-p-1}, \Gamma_p)$.

The concept is based on a sense of direction on the d-dimensional oriented space-time manifold M_d. For two submanifolds U_p and V_q to possess a linking number, it is required that they **intersect transversally**, meaning that **1.** it is possible to introduce a (p+1)-dimensional submanifold W_{p+1} for which $U_p = \partial W_{p+1}$ is its boundary, **2.** there are points of intersection p_i of W_{p+1} and V_q and **3.** The tangent spaces of p_i separately on W_{p+1} and V_q give a subspace of the tangent space of M_d with $T_{p_i}W_{p+1} \bigoplus T_{p_i}V_q \subseteq T_{p_i}M_d$, which is to say that for each p_i there is a well defined orientation on M_d. The linking number is given by summing up the signs of orientations of p_i. In addition, the map is by definition symmetric:

$$Link(U_p, V_q) = \sum_i sign(p_i) \qquad Link(U_p, V_q) = Link(V_q, U_p)$$

and we claim without proving that it is independent of the choice of W_{p+1} on U_p. $sign(p_i)$ is defined to be $+1$ if the induced orientation at p_i goes along that of M_d and -1 if otherwise. This is not the most rigorous explanation of the concept but I hope it is enough for graphical intuition. See Fig.2.4 (taken from [10]) for a picture of how it works in 3d with $p = q = 1$.

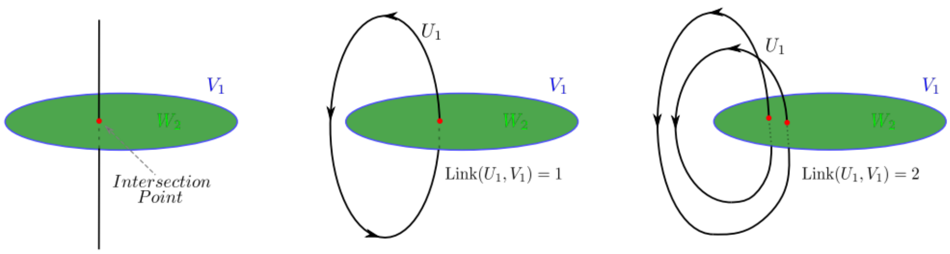

Figure 2.4: linking number: how it works

The most important takeaway is that for $T_{p_i}W_{p+1} \bigoplus T_{p_i}V_q \subseteq T_{p_i}M_d$ to satisfy the dimensions must match on two sides, $p+1+q \stackrel{!}{=} d$. We have now recovered the dimension restriction of a charged operator. Using dual form, there is a more useful and nonetheleast equivalent definition of the Linking number:

$$\begin{aligned} Link(U_p, V_q) &= \int_{M_d} \delta_{d-p-1}(x \in W_{p+1}) \wedge \delta_{d-q}(x \in V_q) \\ &= \int_{W_{p+1}} \delta_{d-q}(x \in V_q) = \int_{V_q} \delta_{d-p-1}(x \in W_{p+1}) \end{aligned} \tag{2.15}$$

So far we have only considered the case where the symmetry is a continous Abelian group $U(1)$. One might be tempted to extend the discussion to non-Abelian symmetry. However, as far as we are concerning about not-so-ordinary symmetry ($p > 0$) it won't be necessary. In fact, **any symmetry with degree $p > 0$ must be Abelian**.

To see this, consider the following situation in 3-dimension: for a 1-form symmetry, the SDO $U_g(S_1)$ marks an element of the group, lives on a 1-sphere (circle), and is topological. This is generally true even if we do not know what the group $G^{(1)}$ is. If we acts on a charged operator $\mathcal{O}(\Gamma_1)$ with two SDOs marked by two different element of the group consecutively, we claim that there is no proper definition of order because we can

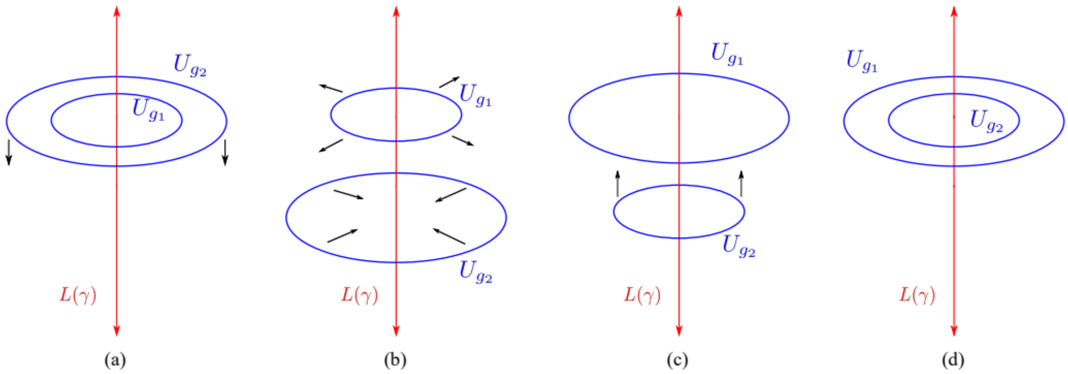

Figure 2.5: Deformations to swap the order

always use the topological nature to perform a series of deformations to swap the order (Fig.2.5 from [10]).

That implies that whatever the fusion rule may be, it must be Abelian, e.g.

$$U_{g_1} \otimes U_{g_2} = U_{g_2} \otimes U_{g_1}$$

2.3 Dual 1-Form Symmetry in Maxwell Theory

We are now ready to look at some physical theory with what we have established. The first example would be a pure gauge theory with gauge group $\mathcal{G} \cong U(1)$: Maxwell. Let us review what we knew about the theory. Without any interesting topology, its action can be written as:

$$\mathcal{S}_m = \frac{1}{2g^2} \int_{M_d} F_2 \wedge *F_2 = \frac{1}{4g^2} \int d^4x F_{\mu\nu} F^{\mu\nu} \tag{2.16}$$

F_2 is the field strength 2-form and with Abelian field A_1, $F_2 = dA_1$. We can extract two equations known as the Maxwell equations:

$$d * F_2 = 0 \qquad dF_2 = 0 \tag{2.17}$$

The first one is the equation of motion while the second one is the Bianchi identity following from the nilpotency of exterior derivative. They reassemble the format of the conservation of a 2-form current and a (d-2)-form current. For reason soon to be apparent, they are called electric and magnetic currents. What physical property are they describing and what are the charged objects under them? We can find some clues in the classical d=4 Maxwell theory on space-time M_4 in which both F and $*F$ are 2-form. Take a spatial slice, R_3, at a constant time. It is now possible to relate these currents with the physical Electric and Magnetic fields:

$$(*_4 F)_{ij} = \epsilon_{ijk} E^k \qquad F_{ij} = \epsilon_{ijk} B^k \tag{2.18}$$

where $i, j, k \in 1, 2, 3$ are the spatial indices. The notation $*_4$ reminds us that the Hodge dual was taken in M_4 instead of R_3. It suffices to look at $*_4F$ as just another 2-form. We can package them back into forms in Euclidean space:

$$*_4F_2 = *E_1 \qquad F_2 = *B_1 \tag{2.19}$$

Integrate the currents over a closed S_2, we find:

$$\begin{aligned}\int_{S_2} *_4F &= \int_{S_2} *E_1 = \int_{S_2} E_i dS^i \\ \int_{S_2} F \quad &= \int_{S_2} *B_1 = \int_{S_2} B_i dS^i\end{aligned}$$

The first one is the famous Gauss's Law or equivalently, the modified version of Equation of Motion in 2.17 when an electric source is added into the action:

$$\int_{S_2} E_i dS^i = cQ^e \tag{2.20}$$

Where Q^e is the amount of electric charge enclosed by S_2. c is a coupling constant. It equals to g^2 in our convention. We can renormalize the current to absorb this constant and make it look nicer:

$$J_2^e = \frac{F_2}{g^2} \qquad \int *J_2^e = Q^e \tag{2.21}$$

It is analogous that same procedures can be applied to the magnetic one:

$$J_2^m = \frac{*F_2}{2\pi} \qquad \int *J_2^m = Q^m \tag{2.22}$$

Q^m would be the charge of a magnetic monopole (if it ever exists). This rather familiar example gives out a few hints on what we should look for, that **1.** the electric/magnetic 1-form charges are the same as electric/magnetic charges under the gauge group $\mathcal{G} \cong U(1)$. Therefore both of the dual 1-form symmetries are $U(1)$ as well. They are often denoted as $U(1)_e \times U(1)_m$. Also **2.** The symmetries are non-trivial only when the existence of electric/magnetic point source (a point in space, which is a line in the full space-time) modify the equation.

2.3.1 Electric 1-Form Symmetry

We will now derive the symmetry action of $U(1)_e$ properly for $d = 4$. The main goal is to find an operator identity analogous to 2.14 in the form

$$U_\alpha^e(S_2) = \exp\left(i\alpha \int_{S_2} *J_2^e\right) \qquad U_\alpha^e(S_2) W_q(\Gamma_1) = e^{iq\alpha Link(S_2,\Gamma_1)} W_q(\Gamma_1) \tag{2.23}$$

where W_q is a proper operator charged under $U(1)_e$ marked with an integer q. We claim that W_q takes the following form

$$W_q(\Gamma_1) = \exp\left(iq\int_{\Gamma_1} A_1\right) \tag{2.24}$$

This is the Wilson line operator. We will discuss Wilson line in more details in the next chapter. To reveal a few interesting insights and also as a good review of how QFT works, here we provide two approaches to derive 2.23:

Approach 1:
Writing out the path integral explicitly:

$$\langle U^e_\alpha(S_2)W_q(\Gamma_1)\rangle = \int \mathcal{D}A\ \exp\left(\frac{i}{2g^2}\int_{M_4} F_2\wedge *F_2 + iq\int_{\Gamma_1} A_1 + i\alpha\int_{S_2} *J^e_2\right) \tag{2.25}$$

Using Poincaré dual, we can write

$$\int_{\Gamma_1} A_1 = \int_{M_4} \delta_3(x\in\Gamma_1)A_1$$

It's possible to define a new action

$$\mathcal{S}'_M = \int_{M_4} \frac{1}{2g^2}F_2\wedge *F_2 + \delta_3(x\in\Gamma_1)A_1$$

Varying the action with respect to A_1 gives equation of motion

$$d * F_2 = qg^2\delta_3(x\in\Gamma_1)$$

Now 2.25 is

$$\langle U^e_\alpha(S_2)W_q(\Gamma_1)\rangle = \int \mathcal{D}A\ \exp\left(i\mathcal{S}'_M + i\alpha\int_{S_2} *J^e_2\right) = \langle\langle U^e_\alpha(S_2)\rangle\rangle$$

where $\langle\langle\ \ \rangle\rangle$ is the new path integral under action $\mathcal{S}'_M$. By the merit of Schwinger-Dyson equation, equation of motion is satisfied as operator identity within the path integral (for a proof, see [5]). Thus

$$\begin{aligned}
\langle\langle U^e_\alpha(S_2)\rangle\rangle &= \langle\langle \exp\left(\frac{i\alpha}{g^2}\int_{S_2} *F_2\right)\rangle\rangle \\
&= \langle\langle \exp\left(\frac{i\alpha}{g^2}\int_{U_3} d * F_2\right)\rangle\rangle \\
&= \langle\langle e^{iq\alpha Link(S_2,\Gamma_1)}\rangle\rangle \\
&= \langle e^{iq\alpha Link(S_2,\Gamma_1)} W_q(\Gamma_1)\rangle
\end{aligned}$$

where U_3 is a submanifold with S_2 as its boundary and $\int_{U_3}\delta_3(x\in\Gamma_1) = Link(S_2,\Gamma_1)$. This approach is the closest to the 3d Gauss's Law example. It also shows that Wilson Line in the path integral is equivalent to adding an electric source to the action, like the one we required for Gauss's Law to work.

Approach 2:
The second way to derive the identity requires us to notice that

$$W_q(\Gamma_1) \overset{elec.}{\to} e^{iq\alpha Link(S_2,\Gamma_1)} W_q(\Gamma_1)$$

Under an infinitesimal shift of the dynamical gauge field

$$A_1 \overset{elec.}{\to} A_1' = A_1 + \lambda_1 \tag{2.26}$$

where λ_1 is a $U(1)$ parameter satisfying $\int_{\Gamma_1} \lambda_1 = \alpha Link(S_2, \Gamma_1)$. We can find an expression for λ from definition of Linking number, that

$$\lambda_1(\alpha) = \alpha \delta_1(x \in U_3)$$

and we take α to be small. We can then rewrite

$$\begin{aligned} U_\alpha^e(S_2) &= \exp\left(i\frac{\alpha}{g^2}\int_{U_3} d * F_2\right) \\ &= \exp\left(i\frac{\alpha}{g^2}\int_{M_4} \delta_1(x \in U_3) \wedge d * F_2\right) \\ &= \exp\left(-\frac{i}{g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *F_2\right) \end{aligned}$$

We have used integral by part assuming that nothing special happens at the boundary of M_4. Then

$$U_\alpha^e(S_2) \overset{elec.}{\to} \exp\left(-\frac{i}{g^2}\left(\int_{M_4} d\lambda_1(\alpha) \wedge *F_2 - \int_{M_4} d\lambda_1(\alpha) \wedge *d\lambda_1(\alpha)\right)\right) \tag{2.27}$$

in which the second term is a term at order α^2 and does not contribute. Lastly, we can find how the action shifts under the transformation:

$$\mathcal{S}_M = \frac{1}{2g^2}\int_{M_4} F_2 \wedge *F_2 \overset{elec.}{\to} \mathcal{S}_M + \frac{1}{g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *F_2 + \frac{1}{2g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *d\lambda_1(\alpha) \tag{2.28}$$

Again the last term is a higher order term. We have used the identity $\alpha \wedge *\beta = \beta \wedge *\alpha$ in the derivation. Now we can put everything together:

$$\begin{aligned} \langle U_\alpha^e(S_2) W_q(\Gamma_1)\rangle &= \int \mathcal{D}A \ \exp\left(i\mathcal{S}_M[A] - \frac{i}{g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *F_2[A] + i\alpha\int_{\Gamma_1} *J_2^e[A]\right) \\ &= \int \mathcal{D}A' \ \exp\left(i\mathcal{S}_M[A'] - \frac{i}{g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *F_2[A'] + i\alpha\int_{\Gamma_1} *J_2^e[A']\right) \\ &= \int \mathcal{D}A \ \exp\left(i\mathcal{S}_M[A] + iq\alpha Link(S_2,\Gamma_1) + i\alpha\int_{\Gamma_1} *J_2^e[A] + \mathcal{O}(\alpha^2)\right) \\ &= \langle e^{iq\alpha Link(S_2,\Gamma_1)} W_q(\Gamma_1)\rangle \end{aligned}$$

where from line 1 to line 2 we simply relabeled $A \to A'$ and from line 2 to line 3 the transformations are used to cancel a term. Notice that an important assumption was made that $\mathcal{D}A = \mathcal{D}A'$ in order to proceed from line 2 to line 3. This assumption was already made in Approach 1 during the derivation of S-D equation and therefore invisible to us at first glance.

Another important comment is that $\mathcal{S}_M$ is not invariant in 2.28. How is it still a global symmetry? In general, it is required that λ_1 is flat ($d\lambda = 0$) for the 1-form symmetry to be global. Check that 2.28 is now invariant. It makes sense that

$$U^e_\alpha = \exp\left(\frac{i}{g^2}\int_{M_4} \lambda_1(\alpha) \wedge d * F_2\right) = \exp\left(-\frac{i}{g^2}\int_{M_4} d\lambda_1(\alpha) \wedge *F_2\right) \tag{2.29}$$

is now consistently trivial (either because of equation of motion or flatness of λ) for any α until the presence of Wilson line alters the action. Also, λ is closed but not exact because otherwise

$$\lambda = d\epsilon \qquad U^e_\alpha = \exp\left(\frac{i}{g^2}\int_{M_4} d\epsilon(\alpha) \wedge d * F_2\right) = \exp\left(\frac{i}{g^2}\int_{M_4} d(\epsilon(\alpha) \wedge d * F_2)\right)$$

is a boundary term and does not act on anything. In other words, λ are cohomology classes. It makes sense considering that the form structure of λ follows from the Poincaré dual of manifold U_3 which is by definition cohomology class.

The third comment is that the presence of the SDO U^e_α allows $\lambda(\alpha)$ of a particular α to be not flat. Indeed we did not make the flatness assumption in the earlier derivation and the SDO canceled out the change of action in 2.28. If we integrate all configuration of SDO, any $\lambda(\alpha)$ is allowed to be non-flat without varying the path integral. This is equivalent to gauging the global 1-form symmetry with a background field, more in the next chapter.

2.3.2 Magnetic 1-Form Symmetry

We can now finish our discussion by extending everything we have studied to the magnetic 1-form symmetry $U(1)_m$. The charged operator is now the 't Hooft Line operator

$$T_m(\Gamma_1) = \exp\left(im\int_{\Gamma_1} \tilde{A}_1\right) \tag{2.30}$$

which depends on the dual gauge field defined by

$$\frac{*F_2}{g^2} = \frac{*dA_1}{g^2} \equiv \frac{d\tilde{A}_1}{2\pi} = \frac{\tilde{F}_2}{2\pi} \tag{2.31}$$

The essential identities are:

$$U^m_\alpha(S_2) = \exp\left(i\alpha\int_{S_2} *J^m_2\right) \qquad U^m_\alpha(S_2)T_m(\Gamma_1) = e^{im\alpha Link(S_2,\Gamma_1)}T_m(\Gamma_1) \tag{2.32}$$

Explicitly

$$*J_2^m = \frac{**F_2}{2\pi} \qquad **F_2 = -(-1)^{2(d-2)}F_2 = -F_2$$

The electric 1-form symmetry was derived from looking at the transformations of action in **Approach 2**. However, the magnetic 1-form symmetry is independent of Maxwell action in its usual form. The identity 2.31 might lead to confusion that A_1 and $\tilde{A}_1$ can be expressed in each other. In fact the relation is non-local. It is impossible to specify A_1 locally given $\tilde{A}_1$. Analogous to electric 1-form symmetry, $\tilde{A}_1$ is shifted to $\tilde{A}_1 + \tilde{\lambda}_1$ under the magnetic symmetry. The parameter $\tilde{\lambda}_1$ is required to be flat in the global symmetry, thus it is invisible in the relation $\frac{*dA_1}{g^2} = \frac{d\tilde{A}_1}{2\pi}$. For this reason both the Maxwell action and the SDO remained invariant under the transformation:

$$\mathcal{S}_M \overset{mag.}{\to} \mathcal{S}_M \qquad U_\alpha^m \overset{mag.}{\to} U_\alpha^m \qquad T_m(\Gamma_1) \overset{mag.}{\to} e^{im\alpha Link(S_2,\Gamma_1)}T_m(\Gamma_1) \tag{2.33}$$

This will still give us the desired result.

In **Approach 1** of electric 1-form symmetry we proved the identity by looking at the modified equation of motion. The magnetic 1-form symmetry has nothing to do with the equation of motion. It follows from the Bianchi identity $dF_2 = 0$ that exists separately from the Maxwell action. The 't Hooft Line modifies Bianchi identity by making the topology non-trivial. Explicitly the identity is now

$$dF_2 = 2\pi m\delta_3(x \in \Gamma_1) \tag{2.34}$$

where m is the charge of the magnetic monopole and 2π is a normalization factor due to Dirac quantization. This also justifies the normalization in J_2^m. More on this in the next chapter. Under this identity we can also get the desired result.

One might ask why the magnetic 1-form symmetry is so different. Why would the action prefers one of the dual symmetries to be in the action and exiles the other? Turns out they can be swapped according to [33]. Instead of assuming $dF_2 = 0$ is true, it can be enforced by a Lagrange multiplier in the action:

$$\mathcal{S}'_M = \int_{M_4} \frac{1}{2g^2}F_2 \wedge *F_2 + \frac{1}{2\pi}L_1 \wedge dF_2 = \int_{M_4} \frac{1}{2g^2}F_2 \wedge *F_2 - \frac{1}{2\pi}F_2 \wedge dL_1 \tag{2.35}$$

On the equation of motion of L_1, the identity $dF_2 = 0$ is enforced and we goes back to the original Maxwell action. If we instead find the equation of motion of F_2, it would be

$$\frac{1}{g^2} * F_2 - \frac{1}{2\pi}dL_1 = 0 \tag{2.36}$$

By 2.31 it is clear that on the equation of motion, L_1 is exactly what we defined as $\tilde{A}_1$. Insert it back to the action yields:

$$\mathcal{S}'_M = \int_{M_4} \frac{g^2}{8\pi^2}\tilde{F}_2 \wedge *\tilde{F}_2 - \frac{g^2}{4\pi^2}\tilde{F}_2 \wedge *\tilde{F}_2 = -\frac{1}{2\tilde{g}^2}\int_{M_4} \tilde{F}_2 \wedge *\tilde{F}_2 \tag{2.37}$$

where the new coupling is $\tilde{g} = \frac{2\pi}{g}$. This reformulation flips the electric and magnetic symmetry and is known as the "S-Duality" of Maxwell theory. One aside is that the duality is between a strongly coupled theory and a weakly coupled one because the new coupling is inversely proportional to the original one. For more information on this topic, see [12].

Chapter 3

Continuous Gauge Theory

We have fully introduced the mechanism of generalized symmetry and looked at d=4 Maxwell theory as an example, and we found that the 1-form symmetries present in the Maxwell theory inevitably mingle with gauge symmetry. Many argue that gauge symmetry is poorly named for it's not a physical symmetry but a redundancy we choose to include in our theory for its various advantages. We need more vocabulary of gauge theory to describe higher-form symmetries in them, and this chapter is devoted to this purpose. We will start with an introduction of the mathematical construction behind gauge theory. Then we will look at certain aspects of gauge theory, namely topological soliton, background fields, and gauged symmetry. All gauge groups discussed in this chapter will be continuous.

3.1 A Crash Course on Fibre Bundle

The Wu-Yang Dictionary was proposed in 1975 [17] as the Rosetta Stone to decode Gauge Theory in the language of Principal Fibre Bundle. Since then it has been clear that there is a coherent global picture behind the ambiguous local theory we used to describe almost everything in nature. This section provides a stripped down explanation of how gauge field is converted to mathematical language. We will be neglecting many nuances that are not central to our concern. This section is a summary of the relevant chapters in [2]. A complete and more rigorous treatment can be found in the book.

Definition 3.1. A **Fibre Bundle** (E, π, M, F, G), often denoted as $E \xrightarrow{\pi} M$, has the following elements:
(1) A differentiable manifold E called the total space.
(2) A differentiable manifold M called the base manifold.
(3) A differentiable manifold F called the typical fibre.
(4) A surjective map $\pi : E \to M$ called projection. The inverse map maps a point $x \in M$ to the fibre at the point $\pi^{-1}(x) = F_x \cong F$.
(5) A Lie group G that acts on F with a left action called the structure group.
(6) For a open cover U_i exists a set of diffeomorphisms called local trivializations: $\phi_i : U_i \times F \to \pi^{-1}(U_i)$. If this maps can be chosen to be identity map everywhere, it means

that the entire bundle is a cross product $M \times F$. This is called a trivial bundle. The existence of ϕ_i everywhere says that any fibre bundle is locally trivial. It is important that **the choice of local trivialization is not unique.** This is a foreshadowing of our physical theory.
(7) Consider diffeomorphism $\phi_{x,i} : F \to F_x$ defined as $\phi_{x,i} = \phi_i(x, u)$, $u \in F$. On the overlap $U_i \bigcap U_j \neq \emptyset$ there is a transition function $t_{ij}(x) = \phi_{x,i}^{-1} \circ \phi_{x,j}$. Transition functions are required to be in group G and we define it to act with a left action: $\phi_j(x, u) = \phi_i(x, t_{ij}(x)u)$.

A way to imagine a fibre bundle is to paste a copy of F on every point of the base manifold M. It is difficult to see how this space is different from a trivial bundle, i.e. cross product $M \times F$. Turned out it has to do with the topology of the base manifold. To see this, we need to define pull-back of a bundle: Suppose there is a map $f : N \to M$ between two differential manifold N and M. Given a fibre bundle $E \xrightarrow{\pi} M$ defined on top of M, we can pull back E onto N as $f^*E \xrightarrow{\pi'} N$. Specifically, for $u \in E$ with $\pi(u) = f(x), x \in N$, we have $f^*u = (x, u) \in N \times E$ and $\pi'(f^*u) = x$ is the projection of f^*E. See the following Fig.3.1 taken from [2]:

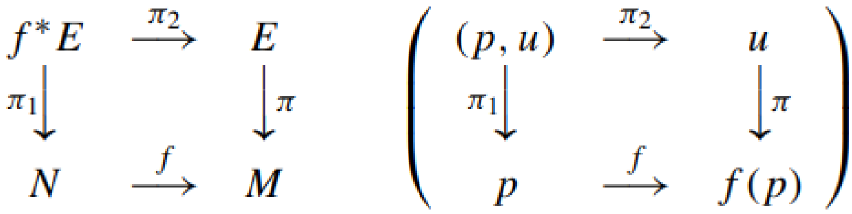

Figure 3.1: Pull-back of Fibre Bundle

With this pull-back defined it is apparent that f^*E and E shares the same typical fibre, i.e. $\pi^{-1}(f(x)) \cong \pi'^{-1}(x) \cong F$. Notice that if f is the identity map, $f^*E = id^*E$ is diffeomorphic, or equivalent, to E. We then state the following theorem without proof:

Definition 3.2. Two maps $f, g : N \to M$ are said to be **Homotopic** if $\exists$ smooth map $F : N \times [0, 1] \to M$ such that $F(x, 0) = f(x)$, $F(x, 1) = g(x)$.

Theorem 3.3. *Given $E \xrightarrow{\pi} M$ a bundle and $f, g : N \to M$ two homotopic maps , f^*E and g^*E are equivalent bundles on N.*

For a proof see (quote). If M is contractible to a point, it means that there exists a smooth map $F : M \times [0, 1] \to M$ and $F(\cdot, 0) = id$, $F(\cdot, 1) : M \to x_0$ for some fixed point x_0. It follows from Theorem 3.3 that $F(\cdot, 0)^*E = id^*E$ is equivalent to a trivial bundle $F(\cdot, 1)^*E$ (for a single point the only bundle $x_0 \times F$ is trivial). In other words, any base manifold contractible to a point can only obtain a trivial bundle. On the other hand, manifolds not contractible can have non-trivial bundle. For example, a trivial bundle on base manifold S_1 with fibre $F = [-1, 1]$ is a cylinder. A mobius strip is also a bundle on S_1 with the same typical fibre.

Another way to see how fibre bundles can be non-trivial is to ask what necessary information is needed to construct a fibre bundle. It turns out to uniquely define (E, π, M, F, G) we need M, F, G, a set of cover U_i and their transitional functions $t_{ij}(x) \in G$.
Step 1: Fetch the union of trivial bundles on all patch $X = \bigcup_i U_i \times F$.
Step 2: Define an equivalent relationship on the overlapped area where t_{ij} connect two patches: For $f \in F$ and $(x, f) \in X$, $(x, f) \sim (x, t_{ij}(x)f)$.
Step 3: Construct a bundle $E = X/\sim$ the elements of which are equivalent classes $(x, [f])$. Find projection $\pi(x, [f]) = x$ and trivializations $\phi_i(x, f) = (x, [f])$.
In short we mod out the part where fibres overlap each other. Clearly that only happens in non-trivial topology where more than one patch is necessary.

One additional concept that would be useful later is section. A section of $E \xrightarrow{\pi} M$ is a smooth map $s : M \to E$ which satisfies $\pi \circ s = id$, which is to say that $s(x)$ is an element in $\pi^{-1}(x) = F_x$ for every point x. We denote the sets of all sections on M as $\Gamma(M, F)$. It can also be defined locally on a cover U_i as $s_i : U_i \to E$.

Fibre bundle is a broad concept. It turns out that for gauge theory we only need a very specific subset of fibre bundles called Principal Fibre Bundle:

Definition 3.4. A **Principal Fibre Bundle**, denoted as $P(M, G)$ is a fibre bundle with base manifold M and a typical fibre F identical to the structure group G. In physics, G is called the gauge group

Now the local trivialization is $\phi_i : U_i \times G \to \pi^{-1}(U_i)$. Since the typical fibre is the structure group, there is not only a left action but naturally a right action of G on the fibre. Then for $u \in \pi^{-1}(U_i)$, $\pi(u) = x \in U_i$, and $\phi_i^{-1}(u) = (x, g_i)$, we can define the right action of $a \in G$ on the total space P to be $ua = \phi_i(x, g_i a)$. This action has the following interesting properties:
1. It is independent of the choice of local trivialization, i.e. for $x \in U_i \bigcap U_j$, $ua = \phi_i(x, g_i a) = \phi_i(x, t_{ij}(x) g_j a) = \phi_j(x, g_j a)$.
2. It provides a mapping to the entire bundle G_x by exploiting the group structure. In other words, take $u \in \pi^{-1}(x)$, $G_x = \pi^{-1}(x) = \{ua | a \in G\}$.
3. Given a local section $s'_i : U_i \to P$ as reference, we can define a local trivialization by $\phi'_i(x, e) := s'_i(x)$ where e is the identity element in G. Other elements of G_x is given by a right action: $\tilde{u} = ua = \phi'_i(x, a) = s'_i(x)a$. This is called the **canonical trivialization**.
4. We can choose a trivialization by choosing a section. We can also relate two choices of sections on U_i, $s'_i(x)$ and $s''_i(x)$, with a group element $s'_i(x) = s''_i(x)g(x)$, $g(x) \in G$. This follows simply from the group property. This is the redundant "gauge symmetry" in physics. In this sense choosing a section(trivialization) is equivalent to choosing a gauge in physics.
5. For $x \in U_i \bigcap U_j$,

$$s'_i(x) = \phi'_i(x, e) = \phi'_j(x, t_{ji}(x)e) = \phi'_j(x, e)t_{ji}(x) = s'_j(x)t_{ji}(x)$$

Notice that this looks very much like the transformation in **4**. This relation shows how gauge on different patches are stitched together. If the base manifold has non-contractible topology, it also contains that information. We will discuss this in a later section.

3.1.1 Connection, Covariant Derivative, Curvature

Given a principal bundle $P(M,G)$, it is possible to find a tangent space T_uP of the total space P at point $u \in P$. It is also possible to uniquely decompose $T_uP = V_uP \oplus H_uP$ as vertical subspace V_uP and horizontal subspace H_uP. This unique separation is called a **connection**. The connection allows any smooth vector field on P to be decomposed into two smooth vector fields: $X = X^H + X^V$. Let's take a look at their properties.

Vertical subspace V_uP contains vectors tangent to a curve within fibre G_u. Another way to show that is to push forward a vector $V \in V_uP$ onto the base manifold and find that it has no "shadow": $\pi_* V = 0 \in T_xM$. Furthermore, since V_uP points inside the fibre G_u, it is related to Lie algebra $\mathbf{g} \cong T_eG$. In fact, there is an isomorphism $\# : \mathbf{g} \to V_uP : A \mapsto A^\#$. $A^\#(u)$ is called the **fundamental vector field**. It is induced by the right action of G on the fibre:

$$A^\#(u) \equiv \frac{d}{d\epsilon}(u \exp{(\epsilon A)})|_{\epsilon=0} = uA \tag{3.1}$$

The group element is written as $exp(\epsilon A)$ without the coefficient of i as in chapter 2. This is simply an issue of convention. See section 3.1.3 for more detail. Also, assume from now on that the group G is a matrix group so that the multiplication in 3.1 makes sense. We then have the following definition:

Definition 3.5. Given $\mathbf{g} \cong Lie(G)$, the **Connection 1-Form** $\omega \in \mathbf{g} \otimes \Omega^1(P)$ is a projection of T_uP onto $V_uP \cong \mathbf{g}$ with $\omega(A^\#) = A$. The horizontal subspace is defined to be the kernel: $H_uP \equiv \{U \in T_uP | \omega(X) = 0\}$. Lastly, ω is under adjoint representation of the pull back of right action on $u \in P$. i.e. with $R_g u := ug$, $R_g^*(\omega) \equiv Ad_{g^{-1}}\omega = g^{-1}\omega g$.

Note that $\Omega^1(P)$ is the notation for the set of 1-forms on manifold P. Defining a connection 1-form field on P is equivalent to choosing a unique connection, which tells us how to fit the fibres together. So far there is no involvement of any local coordinates. It is helpful to find a local form of ω for calculation in physics. That motivates us to pull back ω on a cover U_i by choosing a section s_i:

$$\mathcal{A}_i \equiv s_i^*\omega \tag{3.2}$$

Notice a few things:

1. $\mathcal{A}_i$ is now an object in $\mathbf{g} \otimes \Omega^1(U_i)$, same as 1-form gauge field in physics. Alternatively, the local 1-form can be denoted as $\mathcal{A}_i \in \Omega^1_{\mathbf{g}}(U_i)$, or "1-form on U_i valued in $\mathbf{g}$". It can be proved that the identity 3.2 is two-way, which means that given any $\mathcal{A}'_i \in \mathbf{g} \otimes \Omega^1(U_i)$ and a local section s_i it is possible to construct an ω globally with the information. A full proof is in [2].

2. ω is actually a 1-form field. We can look at it at one point $u \in P$ and it naturally takes the argument from tangent space $U \in T_u P$ via contraction. The argument returns an element of $\mathbf{g}$. Similarly, $\mathcal{A}_i$ also take argument from $T_x U_i$ to $\mathbf{g}$:

$$\mathcal{A}_i : T_x U_i \to \mathbf{g} \tag{3.3}$$

$\mathcal{A}_i$ also takes a curve Γ in U_i as argument. We can integrate out the 1-form on a curve and also return a lie algebra element:

$$\int_\Gamma \mathcal{A}_i \in \mathbf{g} \qquad \mathcal{A}_i : C_1(U_i) \to \mathbf{g} \tag{3.4}$$

where $C_1(U_i)$ is the set of 1-chain on U_i. We can understand it simply a curve for now. One particularly interesting case is when the integral is over a loop. More discussions on this follows in later chapters.

3. Recall identity $< f^* w, v >=< w, f_* v >$ for $w \in \Omega^1(M)$, $v \in T_x M$. We can then write $\mathcal{A}_i = s_i^* \omega(X) = \omega(s_{i*} X)$.

We want to find how $\mathcal{A}_i$ and $\mathcal{A}_j$ on two patches U_i and U_j are related. Equivalently, we can try to write $s_{i*} X$ in terms of $s_j = s_i t_{ij}$. Work on a point $x \in U_i \bigcap U_j$ and find a curve $\gamma : [0,1] \to M$ for which $\gamma(0) = x$ and $\frac{d}{dt}\gamma(t)|_{t=0} = X \in T_x M$:

$$\begin{aligned} s_{i*} X &= \frac{d}{dt} s_i(\gamma(t))|_{t=0} \\ &= \frac{d}{dt}(s_j(\gamma(t)) t_{ji}(\gamma(t)))|_{t=0} \\ &= \left(\frac{d}{dt} s_j(\gamma(t)) \cdot t_{ji}(\gamma(t)) \right)|_{t=0} + \left(s_j(\gamma(t)) \cdot \frac{d}{dt} t_{ji}(\gamma(t)) \right)|_{t=0} \\ &= s_{j*} X \cdot t_{ji}(x) + s_i(x) t_{ji}^{-1}(x) \cdot \frac{d}{dt} t_{ji}(\gamma(t))|_{t=0} \end{aligned}$$

In the first term, assuming G is a matrix group we have $s_{j*} X \cdot t_{ji} = R_{t_{ji}*} s_{j*} X$. This is because the push forward acting on $U \in T_u P$ is defined as

$$R_{g*} U[f(R_g u)] = U[f \circ R_g](u) \tag{3.5}$$

for some generic function $f : u \to \mathbb{R}$. The LHS is

$$(R_{g*} U)^\mu \partial_\mu f(R_g u) \tag{3.6}$$

The RHS can be written as

$$U^\mu \partial_\mu (f \circ R_g)(u) = U^\mu \partial_\mu f(R_g u) R'_g(u) \tag{3.7}$$

For a matrix group, it makes sense to write

$$R'_g(u) = \frac{d}{du}(ug) = g \tag{3.8}$$

thus the RHS is now

$$U^\mu \partial_\mu f(R_g u) g = U^\mu g \partial_\mu f(R_g u) \tag{3.9}$$

Identifying with LHS we find

$$R_{g*} U = Ug \tag{3.10}$$

Back to the main story, in the second term,

$$t_{ji}^{-1}(x) \cdot \frac{d}{dt} t_{ji}(\gamma(t))|_{t=0} = \frac{d}{dt}(t_{ji}^{-1}(x) t_{ji}(\gamma(t)))|_{t=0} \in T_e G \cong \mathbf{g}$$

because at $t = 0$, $t_{ji}^{-1}(x) t_{ji}(\gamma(t)) = e$. Also, we have identity:

$$\frac{d}{dt} t_{ji}(\gamma(t))|_{t=0} \equiv X[t_{ij}] \equiv < dt_{ij}, X >= dt_{ij}(X)$$

Putting the information together, according to 3.1:

$$s_i(x) t_{ji}^{-1}(x) \cdot \frac{d}{dt} t_{ji}(\gamma(t))|_{t=0} = (t_{ji}^{-1}(x) \cdot \frac{d}{dt} t_{ji}(\gamma(t))|_{t=0})^\#([s_i(x)]) = (t_{ji}^{-1}(x) dt_{ij}(X))^\#([s_i(x)])$$

Input $s_{i*}X$ to ω, we reach the conclusion

$$\begin{aligned} s_i^* \omega(X) = \omega(s_{i*}X) &= \omega(R_{t_{ij}*} s_{j*} X) + t_{ij}^{-1}(x) dt_{ij}(X) \\ &= R_{t_{ij}}^* \omega(s_{j*} X) + t_{ij}^{-1}(x) dt_{ij}(X) \\ &= t_{ij}^{-1}(x) \mathcal{A}_j(X) t_{ij}(x) + t_{ij}^{-1}(x) dt_{ij}(X) \end{aligned}$$

This is true for all point. It reduces to the identity

$$\mathcal{A}_i = t_{ij}^{-1} \mathcal{A}_j t_{ij} + t_{ij}^{-1} dt_{ij} \tag{3.11}$$

Following the same procedure, it is easy to show that choosing different sections $s_i(x) = s_i'(x) g(x)$ yields the same transformation on local connection:

$$\mathcal{A}_i' = g^{-1} \mathcal{A}_i g + g^{-1} dg \tag{3.12}$$

This is the gauge transformation. Recall that exterior derivative is defined to be the map $d : \Omega^r(M) \to \Omega^{r+1}(M)$ on a manifold M. We denote $d_p : \Omega^r(P) \to \Omega^{r+1}(P)$ as the exterior derivative on P and generalize it to act on object $\phi \in \mathbf{g} \otimes \Omega^r(P)$ by requiring derivative to act on the r-form subspace. With the presence of a connection there exists a covariant derivative:

Definition 3.6. For $\phi \in \mathbf{g} \otimes \Omega^r(P)$ the **Covariant Derivative** $D : \mathbf{g} \otimes \Omega^r(P) \to \mathbf{g} \otimes \Omega^{r+1}(P)$ is defined as

$$D\phi(X_1, ..., X_{r+1}) \equiv d_p \phi(X_1^H, ..., X_{r+1}^H)$$

Another definition follows:

Definition 3.7. The **Global Curvature 2-Form** Ω is defined as:

$$\Omega \equiv D\omega \in \mathbf{g} \otimes \Omega^2(P)$$

The curvature satisfies the Cartan Structure equation:

$$\Omega = d_p\omega + \omega \wedge \omega \tag{3.13}$$

which has the local form:

$$\mathcal{F}_i = d\mathcal{A}_i + \mathcal{A}_i \wedge \mathcal{A}_i \tag{3.14}$$

For a derivation of 3.13 from definition, see [2].

In the global picture, it is straight forward to recognize what Ω means geometrically. Take $X, Y \in H_uP$:

$$\Omega(X,Y) \equiv D\omega(X,Y) \equiv d_p\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]) = -\omega([X,Y])$$

where $[\cdot,\cdot]$ is the lie bracket and we have used the coordinate-free definition of exterior derivative. Define a basis of vectors on M as $\{\partial_{x_1}, \partial_{x_2}, ..., \partial_{x_{d+1}}\}$ and set $\pi_*(X) = \epsilon\partial_{x_1}$, $\pi_*(X) = \delta\partial_{x_2}$ with infinitesimal parameters ϵ, δ. Recall identity $f_*[X,Y] = [f_*X, f_*Y]$, we find

$$\pi_*([X,Y]) = [\pi_*X, \pi_*Y] = \epsilon\delta[\partial_{x_1}, \partial_{x_2}] = 0$$

which states that $[X,Y] \in V_uP$ and $\Omega(X,Y) = -\omega([X,Y]) = -A \in \mathbf{g}$ where $A^\# = [X,Y]$ is the infinitesimal discrepancy of the fibre after parallel transporting it around a close parallelogram on M constructed with sides $\epsilon\partial_{x_1}$ and $\delta\partial_{x_2}$.

3.1.2 Associated Vector Bundle

In physics we would like to consider a gauge theory along with "matter particles" that are charged under the gauge group. For example, in scalar QED there would be a quark that is in a certain representation of the gauge group. To describe these particles it is necessary to introduce the concept of an associated bundle:

Definition 3.8. Given a principal fibre bundle $P(M,G)$ and a k-dimensional vector space V acted upon by G on the left by a k-dimensional representation R of G, an **Associated Vector Bundle** (E,π,M,G,V,P) is the product $E = P \times V$ defined up to an equivalent relation $(u,v) \sim (ug, R(g^{-1})v)$ (or equivalently $(ug,v) \sim (u, R(g)v)$)) where $u \in P$, $v \in V$ and $g \in G$.

The projection $\pi_E : E \to M$ is defined as $\pi_E(u,v) = \pi(u)$ where π is the projection in principal bundle $P(M,G)$. We can check that the projection is compatible with the equivalent condition: $\pi_E(ug, R(g^{-1})v) = \pi(ug) = \pi(u) = \pi_E(u,v)$. The new local trivialization is $\psi_i : U_i \times V \to \pi_E^{-1}(U_i)$. Perhaps more important to us, the new transition functions are in the representation of the vector space. It is simply $R(t_{ij})$.

Charged particles live in the associated vector bundle. Their phase is stored as their position in fibre. The equivalent relation seems arbitrary and unintuitive at first sight. It can be understood in two ways:

1. At every point $x \in M$ the relation quotients out the information of G fibre in P by identifying u with ug. It ensures that the fibre of E is diffeomorphic to the vector space V.

2. There is an induced left action of u on v written as $R(u)v$. Of course we would like $R(ug)R(g^{-1})v = R(u)v$, which is exactly the equivalence defined above.

3.1.3 Physical Notation of Gauge Field

This section is a heads-up on some rather annoying convention issue that might confuse the more meticulous readers.

Though the discussion of fibre bundle works for general gauge group, the remaining of this paper would focus almost exclusively on $\mathcal{G} = SU(N)$. In this case physicists usually wants the generators (basis) of the Lie algebra to be Hermitian instead of anti-Hermitian. The way to do that is to write Lie algebra element with a coefficient i.

For this reason the physical gauge field A_1 is often defined as

$$\mathcal{A}_1 = iA_1 \tag{3.15}$$

The physical gauge field has transformation

$$A_1 = -i\mathcal{A}_1 \to -i(g^{-1}\mathcal{A}_1 g + g^{-1}dg) = g^{-1}A_1 g - ig^{-1}dg = g^{-1}A_1 g + idg^{-1}g \tag{3.16}$$

It is also custom to write the group element as

$$g = e^{i\alpha} \tag{3.17}$$

where α is the Hermitian Lie algebra element. Transformation 3.16 can thus be written as

$$A_1 \to e^{-i\alpha}A_1 e^{i\alpha} + d\alpha \tag{3.18}$$

3.2 Aspects of Topology

The fact that local connection is not a well-defined 1-form on non-trivial topology allows us to patch together seemingly impossible configuration of gauge field. There are various solutions in gauge theory that are particle-like and unlike elementary particles, obtain topological structure within them. They are usually called "Solitons". This aspect of gauge theory has been intensively studied since Polyakov's paper in 1975 ([28]). Hereby a basic introduction of monopole and winding number is provided to give a heuristic intuition for other parts of the paper. The majority of this section is based on [12], [2], [15], and [14]. Interested readers should refer to these works as well as [7] for further reading.

3.2.1 Dirac Monopole as Non-trivial Bundle

In Chapter 2 we constructed 't Hooft Line as an analogy of Wilson Line with a dual gauge field $\tilde{A}_1$. There is an alternative construction from gauge field A_1. In 4d the 't Hooft Line is equivalent to inserting a 2-sphere S_2 with a Dirac monopole configuration on it around every point of the specified line. The Dirac monopole configuration is a principal bundle $P(S_2, U(1))$ on a sphere. The 2-sphere can be covered by two patches,

$$U_N = \{(\theta,\phi)|0 \leq \theta < \frac{\pi}{2} + \epsilon\} \qquad U_S = \{(\theta,\phi)|\frac{\pi}{2} - \epsilon < \theta \leq \pi\} \tag{3.19}$$

This patching effectively divides the 2-sphere into the Northern and Southern halves with a band of overlap around the equator with an infinitesimal width. Denote the global connection 1-form on the bundle as ω, there are two local 1-forms:

$$A_1^N = s_N^*\omega \qquad A_1^S = s_S^*\omega \tag{3.20}$$

obtained by choosing sections. We can skip this step and borrow the result from Wu and Yang ([17],[2]):

$$A_1^N = \frac{1}{2}m(1 - cos\theta)d\phi \qquad A_1^S = -\frac{1}{2}m(1 + cos\theta)d\phi \tag{3.21}$$

where m denotes the strength of the monopole. This configuration of gauge field will give the field strength of a magnetic monopole outside of S_2. There exists a transition function between them and it is an element of the structure group $U(1)$. The band of overlap can be taken as just the equator and the transition function only depend on ϕ. It can be written in the form:

$$t_{NS}(\phi) = \exp\left(i\sigma(\phi)\right) \tag{3.22}$$

Since the group is Abelian, the gauge transformation 3.18 breaks down to

$$A_1^N = A_1^S + d\sigma \qquad d\sigma = A_1^N - A_1^S = md\phi \tag{3.23}$$

Integrate both side:

$$\Delta\sigma = \int_0^{2\pi} md\phi = 2\pi m \tag{3.24}$$

Unlike the gauge field, transition function is required to be single-valued on each point of the equator S_1. This gives the condition that $m \in \mathbb{Z}$, i.e. the magnetic charge is quantized.

In a physical setting where charged matter is included, the bundle is instead the associated vector bundle of $P(S_2, U(1))$ labeled by representation. In the case of $U(1)$, irreps are labeled by integer $\widehat{G} \cong \mathbb{Z}$. In this bundle, the transition function is an element of the corresponding representation of the group. We write:

$$t_{NS}(\phi) = \exp\left(i\alpha\sigma(\phi)\right) \tag{3.25}$$

where α marked the representation charged matter is in, also known as the electric charge of the matter. Repeating all the calculation, quantization condition is now $m\alpha \in \mathbb{Z}$. This is the Dirac quantization condition.

There are a few useful comments:

1. Notice that the local connection on the northern hemisphere has a singularity on the south pole and vice versa. This construction does not work on $\mathbb{R}^3$ but instead requires $\mathbb{R}^3\backslash\{0\}$, which is homotopic to S_2. In other words, there is a singularity on the exact point of the monopole.

2. The existence of Dirac monopole on a principal bundle demands the structure group to be compact. In other words, it proves to us that the gauge group is indeed $U(1)$ instead of $\mathbb{R}$, which is indistinguishable from $U(1)$ locally. Dirac quantization condition also relies on the assumption of a compact group, thus any observation of a magnetic monopole gives us an explanation for quantization of electric charge. Unfortunately that has not happened yet.

3. The monopole exists in the gauge theory without the need for any charged matter. As we will see in the next section, Wilson Lines also exist in the pure gauge theory as a basic feature.

There is a more elegant way to classify magnetic monopole and for that purpose we need to introduce the concept of Homotopy group as taken from [3]:

Definition 3.9. In Definition 3.2 we have defined what it means for two maps to be homotopic. Equivalent classes can be constructed from it by denoting $[f] = \{g | g \textit{ homotopic to } f\}$. They are called homotopy classes. We can then define **Homotopy Group** $\Pi_n(X)$ as the set of all homotopy classes $[f]$ in continuous maps $f : S_n \to X$ with X some topological space and S_n the n-sphere parameterized by $[0, 1]^n$. We claim that for $n \geq 1$ the set $\Pi_n(X)$ forms a group under multiplication $[f] \times [g] = [f * g]$ where $*$ is the composition defined by gluing the maps together. For example, if $n = 1$:

$$f * g(s) = \begin{cases} f(2s) & 0 \leq s \leq \frac{1}{2} \\ g(2s-1) & \frac{1}{2} \leq s \leq 1 \end{cases}$$

Notice that the quantization condition requires transition function $t_{NS}(\phi)$ to be a map $t_{NS} : S_1 \to U(1) \cong S_1$ and therefore any possible t_{NS} belongs to an element in homotopy group $\Pi_1(S^1)$. We claim without proof that $\Pi_1(S_1) \cong \mathbb{Z}$. This is to say that one configuration of t_{NS} is in some homotopy class marked by an integer and can not be continuously deformed into configurations in other homotopy classes. This integer is just the number m we have seen above. This number is called the "winding number" or “topological charge” in various contexts. In ours, it is simply the magnetic charge of the monopole. It is called a “winding” number because it describes how many times the parameter ϕ is wrapped around $U(1)$. In the case where $n > 1$, the gauge transformation of A_1 associated with the transition function is called a **Large Gauge Transformation**.

3.2.2 The Theta Term

In Yang-Mills Theory (and Maxwell alike), there is a famous term that can be added to the action. Classically, it is a boundary term and has no influence on the classical equation of motion. In quantized Maxwell theory, the term also does not contribute to the partition function unless there are interesting boundary condition created by material or otherwise. In quantized non-Abelian Yang-Mills Theory, the term plays a different role and alters the ground state of the Hilbert space. It marks inequivalent quantizations of the classical theory. (See [15]). This is the so-called theta term. On action level in 4d Yang-Mills, the term writes us:

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2}\int_{M_4} tr(F_2 \wedge F_2) \tag{3.26}$$

where $tr()$ is the trace over the gauge group indices. The term is often described as “topological” because unlike the original Maxwell action, there is no Hodge dual involved and therefore it is independent of the metric on M_4. Note that in Maxwell, the coupling is instead $\frac{\theta}{4\pi^2}$ because of a normalization chosen for the trace of generators. It can be shown that the theta term is a total derivative: define the Chern-Simons form as

$$\kappa_3 = tr(A_1 \wedge dA_1 + \frac{2}{3}A_1 \wedge A_1 \wedge A_1) \tag{3.27}$$

we find that:

$$\begin{aligned} d\kappa_3 &= tr(dA_1 \wedge dA_1 + 2dA_1 \wedge A_1 \wedge A_1) \\ &= tr(dA_1 \wedge dA_1 + dA_1 \wedge A_1 \wedge A_1 + A_1 \wedge A_1 \wedge dA_1 + A_1 \wedge A_1 \wedge A_1 \wedge A_1) \\ &= tr(F_2 \wedge F_2) \end{aligned}$$

Notice that we have used the cyclic property of the trace and the anti-commutating property of 1-form to add a trivial term:

$$\begin{aligned} tr(A_1 \wedge A_1 \wedge A_1 \wedge A_1) &= A_1^a \wedge A_1^b \wedge A_1^c \wedge A_1^d \otimes tr(T_aT_bT_cT_d) \\ &= -A_1^d \wedge A_1^a \wedge A_1^b \wedge A_1^c \otimes tr(T_dT_aT_bT_c) \\ &\overset{relabel}{=} -A_1^a \wedge A_1^b \wedge A_1^c \wedge A_1^d \otimes tr(T_aT_bT_cT_d) \\ &= 0 \end{aligned}$$

Despite being a boundary term, the theta term has interesting physical implication even in Maxwell Theory. One example is the topological insulator, which refers to material that has a different value of θ in its interior. A good introduction could be found in

[15]. We will not be talking about theta term in Maxwell. Instead, we will examine how the theta term contributes to quantum Yang-Mills in the specific case of $\mathcal{G} \cong SU(2)$ in 4d.

Winding Number in SU(2) Yang-Mills Theory

In this section, we will work in Euclidean space-time R_4. Recall the theta term is

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2} \int_{R_4} d\kappa_3 = \frac{\theta}{8\pi^2} \int_{\partial R_4} \kappa_3 \tag{3.28}$$

There need to be some boundary of the space-time ∂R_4 for this term to contribute. It can be chosen as the 3-sphere at spatial infinity: $\partial R_4 \cong S_3^\infty$. The canonical procedures as in [15] follows that for the action to be finite,

$$A_1(x) \to idg(X)^{-1}g(x) \qquad as\ x \to \infty \tag{3.29}$$

There is an alternative construction given by [2] that is closer to the Dirac monopole discussion. Suppose that the space-time is allowed to be compactified by including the infinity and it takes the form of a 4-sphere S_4. Choose one pole on the sphere to be the origin and the opposite pole to be the point of infinity. The manifold can be covered by two patches:

$$U_0 = \{x|0 \leq \|x| < L + \epsilon\} \qquad U_\infty = \{x|L - \epsilon < |x| \leq \infty\} \tag{3.30}$$

where L is some 4-distance that characterizes the long distance behaviour of the theory. There are two local connections A_1^0 and A_1^∞. If we make the choice that $A_1^\infty = 0$ then according to 3.16 the gauge transformation at the overlap reads

$$A_1^0 = g^{-1}A_1^\infty g + idg^{-1}g = idg^{-1}g = -ig^{-1}dg \tag{3.31}$$

where g now maps the equator of S_4, which is a 3-sphere, into the group $SU(2)$:

$$g : S_3 \to SU(2) \cong S_3 \qquad [g] \in \Pi_3(S_3) \tag{3.32}$$

We claim without prove that $\Pi_3(S_3) \cong \mathbb{Z}$. Thus the gauge transformation is marked by the winding number of S_3 over S_3. In fact, this statement is true for any simple, compact Lie group $\mathcal{G}$ since it has been proven mathematically that $\Pi_3(\mathcal{G}) \cong \mathbb{Z}$. What is the consequence of this classification? Plugging the expression 3.31 back into the theta term yields

$$\mathcal{S}_\theta = \frac{i\theta}{12\pi^2} \int_{S_3} tr(g^{-1}dg \wedge g^{-1}dg \wedge g^{-1}dg) = i\theta n \tag{3.33}$$

where $n \in \mathbb{Z}$. We claim that the triple wedge term gives an integer because it is in the form of Haar Measure for S_3. The contribution of this term in the Euclidean path integral is $e^{i\theta n}$. The fact that n is an integer states that θ is 2π periodic in the eyes of the quantum theory.

Theta Term under Time Reversal

We can consider the what symmetries the theta term obeys. Like the Maxwell action, the theta term is gauge invariant and Lorentz invariant. Unlike the Maxwell action, theta term breaks the discrete time reversal symmetry with two exceptions. More explicitly, time reversal is the following transformation:

$$\mathbb{T}^{\mu}{}_{\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.34}$$

The theta term dressed with full indices is

$$\mathcal{S}_\theta = \frac{\theta}{16\pi^2} \int d^4x \;\; \epsilon_{\mu\nu\sigma\tau} F^{\mu\nu a} F^{\sigma\tau a} \tag{3.35}$$

On flat metric $\epsilon_{\mu\nu\sigma\tau}$ is not a proper tensor. It is designed to have the same components in all frames. Here it does not transform. Also, it only admits non-zero value when it has three space and one time indices. The field strength is a proper tensor and transforms as one:

$$F^{\mu\nu a} \to \mathbb{T}^{\mu}{}_{\alpha}\mathbb{T}^{\nu}{}_{\beta} F^{\alpha\beta a} = \begin{cases} F^{\mu\nu a} & \mu,\nu = 0 \; or \; \mu,\nu \in \{1,2,3\} \\ -F^{\mu\nu a} & otherwise \end{cases} \tag{3.36}$$

According to the property of ϵ, all non-zero terms has one F with two spatial indices and one F with one spatial and one time indices, which means that we always get a minus sign for transforming the product of them:

$$\mathcal{S}_\theta \to -\mathcal{S}_\theta \tag{3.37}$$

We can absorb the negative into the constant term and equally say that under time reversal,

$$\theta \to -\theta \tag{3.38}$$

This is clearly stating that the theta term is breaking time reversal symmetry. There are, however, two special values of θ. When $\theta = 0$, the symmetry is preserved because $0 = -0$. This is rather trivial because it is equivalent to erasing the theta term. Another special case arises from the periodic property of θ. Since $\theta \sim \theta + 2\pi$, obviously $\pi \sim -\pi$ and the term preserves time reversal symmetry in a quantum theory when $\theta = \pi$. The periodic property is not exact. In the next chapter we will see how this condition can be broken.

3.2.3 Witten Effect

As Witten discovered in [25], a magnetic monopole automatically gains an electric charge due to a non-zero theta term. There are multiple ways to show this. Tong coins a version based on a thought experiment of covering the magnetic monopole with topological insulator in [15]. As summarized in [12] Coleman has also given an explanation of Witten Effect in Maxwell Theory with electric and magnetic fields. We will reproduce Coleman's explanation below.

Recall the classical electric magnetic fields are

$$E^i = \partial^i A_0 \tag{3.39}$$

$$B^i = \epsilon^{ijk}\partial_j A_k + \frac{m}{2|r|^2}r^i \tag{3.40}$$

In a situation where there is a magnetic monopole of charge m sitting at $r = 0$ but no electric source. Written in the same language is the theta term

$$\mathcal{S}_\theta = \frac{\theta}{4\pi^2}\int d^3r \ \ E^i B_i \tag{3.41}$$

Up and down indices don't really matter in Euclidean space. Substitute the expression of electric field and magnetic field, we get

$$\begin{aligned}
\mathcal{S}_\theta &= \frac{\theta}{4\pi^2}\int d^3r \ \ E^i B_i \\
&= \frac{\theta}{4\pi^2}\int d^3r \ \ \partial^i A_0(\epsilon_{ijk}\partial^j A^k + \frac{m}{2|r|^2}r_i) \\
&= \frac{\theta}{4\pi^2}\int d^3r \ \ -\epsilon_{ijk}A_0\partial^i\partial^j A^k - \frac{m}{2}A_0\frac{1}{|r|^2}\partial^i r_i \\
&= \frac{m\theta}{8\pi^2}\int d^3r \ \ A_0\partial^i\frac{1}{r_i}
\end{aligned}$$

We find that

$$-\partial^i\frac{1}{r_i} = \nabla\cdot\frac{\vec{r}}{|r|^2} = \frac{1}{|r|^2}\frac{\partial}{\partial r}(\frac{r^2}{r^2}) = 0 \tag{3.42}$$

with an asterisk that $r \neq 0$. The derivative tends to infinity at $r = 0$ and it is actually a delta function

$$\nabla\cdot\frac{\vec{r}}{|r|^2} = 4\pi\delta(|r| = 0) \tag{3.43}$$

with the value coming from $\int_V \nabla\cdot\frac{\vec{r}}{|r|^2} = \int_{\partial V}\frac{\vec{r}}{|r|^2} = 4\pi$. Thus the theta term returns

$$\mathcal{S}_\theta = -\frac{m\theta}{2\pi}\int d^3r \ \ A_0\delta(|r| = 0) = -\frac{m\theta}{2\pi}A_0(|r| = 0) \tag{3.44}$$

which is an electric source at the origin with magnitude $\frac{m\theta}{2\pi}$. Admittedly, this only works in Maxwell theory. We will see a derivation for more general situation next chapter.

3.3 Background Fields and Gauged Symmetry

In Chapter 2, we have hinted on the insertion of SDOs as a way to "gauge" global symmetry of the theory, but what is "gauging" exactly? Let's return to a familiar story. In the Standard Model, gauge theory is necessary when a global 0-form symmetry wants to become a local symmetry. Recall that massless complex Higgs field in 4d has action:

$$\mathcal{S}_\phi = \int d^4x \ \ (\partial_\mu \phi)^* \partial^\mu \phi \tag{3.45}$$

This action is equiped with a global symmetry

$$\phi \to e^{i\theta}\phi \approx \phi + i\theta\phi$$

By an infinitesimal transformation, the symmetry has a Noether current

$$j_\mu = \phi \partial_\mu \phi^* - \phi^* \partial_\mu \phi$$

This is a U(1) 0-form symmetry. We can rewrite everything in form language for clarity:

$$\mathcal{S}_\phi = \int_{M_4} d\phi^* \wedge *d\phi \qquad J_1 = i(\phi d\phi^* - \phi^* d\phi) \qquad d * J_1 = 0 \tag{3.46}$$

When the global symmetry becomes local, Or in other words, when the parameter $\theta(x)$ becomes a function on M_4, the original action is no longer conserved under the symmetry. Instead, we have

$$\delta\mathcal{S} = \int d^4x \ \ i\partial_\mu\theta(x)((\partial^\mu\phi)^*\phi - \phi^*\partial^\mu\phi) = i\int_{M_4} d\theta \wedge *J_1 \tag{3.47}$$

We claim that a new action invariant under the local symmetry can be constructed by coupling the current to a background field. Explicitly, the action is modified by a term:

$$\mathcal{S}_A = ie\int_{M_4} A_1 \wedge *J_1 = ie\int d^4x \ \ A_\mu((\partial^\mu\phi)^*\phi - \phi^*\partial^\mu\phi) \tag{3.48}$$

The background field is non-dynamical and not summed over in path integral. It can be promoted to a dynamical field by adding a dynamical term to the action. The background field is required to have a gauge symmetry accompanying the local symmetry of ϕ:

$$A_1 \to A_1 - \frac{1}{e}d\theta(x) \qquad \phi \to e^{i\theta(x)}\phi \approx \phi + i\theta(x)\phi \tag{3.49}$$

The action is still not invariant under the transformation. We need to add a third term that only depends on the background field. We are allowed to add this as a counter term:

$$\mathcal{S}_{c.t.} = e^2\int_{M_4} |\phi|^2 A_1 \wedge *A_1 \tag{3.50}$$

The action is now

$$\begin{aligned}\mathcal{S}[A] &= \int_{M_4} d\phi^* \wedge *d\phi + ieA_1 \wedge *J_1 + e^2|\phi|^2 A_1 \wedge *A_1 \\ &= d\phi^* \wedge *d\phi + ie\phi A_1 \wedge *d\phi^* - ie\phi^* A_1 \wedge *d\phi + e^2|\phi|^2 A_1 \wedge *A_1 \\ &= (d\phi^* - ie\phi^* A_1) \wedge *(d\phi + ie\phi A_1) \\ &\equiv (D\phi)^* \wedge *(D\phi)\end{aligned}$$

We have recovered the covariant derivative in particle physics notation. This action is gauge invariant as we know it. The parameter e is a coupling we can choose. It is convenient for the purpose of our discussion to set it to 1. We can also normalize $|\phi| = 1$, under which $\phi \in \mathbb{C}$ takes the form $\phi = e^{i\Phi}$ and $\phi^* = e^{-i\Phi}$. Φ is cyclic, valued in $[0, 2\pi)$. The symmetry transformation is rewritten as

$$\Phi \to \Phi + \theta \tag{3.51}$$

where θ is also valued in $U(1)$. Take Φ as the fundamental field and we find a sleek reformulation of massless scalar field that is closer to the generalized language in chapter 2:

$$\mathcal{S}_\Phi = \int d^4x \;\; \partial_\mu e^{-i\Phi} \partial^\mu e^{i\Phi} = \int d^4x \;\; \partial_\mu \Phi \partial^\mu \Phi = \int_{M_4} d\Phi \wedge *d\Phi \tag{3.52}$$

The symmetry is global when the parameter θ is flat/closed, which makes it a constant. We can find the global current J_1 by varying the action as if θ is not flat and look at the linear part:

$$\begin{aligned}\delta_\theta \mathcal{S}_\Phi &= \delta \int_{M_4} d(\Phi + \theta) \wedge *d(\Phi + \theta) \\ &= \int_{M_4} d\delta\theta \wedge *d\Phi + d\Phi \wedge *d\delta\theta + \mathcal{O}(\theta^2) \\ &= \int_{M_4} 2\delta\theta \wedge d * d\Phi + \mathcal{O}(\theta^2) \\ &\overset{!}{=} -i \int_{M_4} \delta\theta \wedge d * J_1\end{aligned}$$

We recognize that

$$J_1 = 2id\Phi \tag{3.53}$$

When θ is not flat, the action is not invariant under the local/gauged symmetry. We need to couple it to a gauge field to find a new action that is invariant. The coupled action is then, following from 3.48 and 3.50:

$$\mathcal{S}[A] = \int_{M_4} d\Phi \wedge *d\Phi - 2A_1 \wedge *d\Phi + A_1 \wedge *A_1 = \int_{M_4} (d\Phi - A_1) \wedge *(d\Phi - A_1) \tag{3.54}$$

where A_1 is the background field that follows the transformation rule:

$$A_1 \to A_1 + d\theta(x) \tag{3.55}$$

Canonically this is not the end of the story. The background field is usually promoted to a dynamical field by adding the Maxwell term 2.16. In light of path integral, all configurations of A_1 are summed over to produce a partition function.

3.3.1 Generalized to p-Form Symmetry

We are now using a language that is suitable for higher form symmetry. The generalization is straight forward. Assuming the p-form symmetry group is

$$G^{(p)} \cong U(1)$$

without loss of generality. Recall that we are allowed to do so because symmetry with $p \geq 1$ is always Abelian. The global symmetry can be gauged by a (p+1)-form background field with a transformation

$$B_{p+1} \to B_{p+1} + d\Lambda_p \tag{3.56}$$

The background field couple to the action via a term

$$\mathcal{S}_B = i \int_{M_d} B_{p+1} \wedge *J_{p+1} \tag{3.57}$$

as well as other local counter terms dependent on the background field only. J_{p+1} here is of course the Noether current for the global p-form symmetry.

The formal definition of gauging given by [10] requires us to also promote the background field to a dynamical field by path integrating all gauge-inequivalent classes. The **Gauged Theory** of an original theory $\mathcal{T}$ is denoted as $\mathcal{T}/G^{(p)}$ where $G^{(p)}$ is the gauged global symmetry. It is defined by the partition function

$$\mathcal{Z}_{\mathcal{T}/G^{(p)}} \propto \int \mathcal{D}B_{p+1} \mathcal{Z}_{\mathcal{T}}[B_{p+1}] \tag{3.58}$$

This only makes sense when $Z_{\mathcal{T}}[B_{p+1}]$ is invariant under gauge transformation 3.56. We will find out in the next section that this is not always true.

3.3.2 Maxwell Theory Revisited

Let's see how this new tool can be used on the 1-form symmetries of 4d Maxwell Theory. Before we gauge the theory, recall that the electric 1-form symmetry acts on the dynamical gauge field by shifting

$$A_1 = A_1 + \lambda_1 \tag{3.59}$$

where λ_1 is required to be flat in order to preserve the action. The analogy to the Higgs model should be obvious by now. Also recall from 2.29 that the SDO of electric 1-form symmetry can be written as

$$U^e_\alpha = \exp\left(-\frac{i}{g^2}\int_{M_4} d\lambda_1(\alpha)\wedge *F_2\right) = \exp\left(-i\int_{M_4} d\lambda_1(\alpha)\wedge *J^e_2\right)$$

where $\lambda_1(\alpha)$ is a specific parameter

$$\lambda_1(\alpha) = \alpha\delta_1(x\in U_3)$$

that lives in U_3 which is the 3-dim submanifold the SDO is positioned. Immediately the form of the SDO looks like the coupling term 3.57 but instead of background field B_2 the current is coupled to $d\lambda_1$. This is actually a gauge transformation of the 3.57 term! As discussed earlier, inserting every possible SDO into the path integral allows λ to be non-flat in general. This is equivalent to summing over all possible configurations of B_2 and by gauge transformation 3.56 of the fields, spawning all the possible SDO into existence.

4d Maxwell Theory has a dual $U(1)_e\times U(1)_m$ 1-form symmetry. Let's try to couple the theory to both background fields B^e_2 and B^m_{4-2}. For the electric 1-form symmetry, the terms to add are

$$\mathcal{S}_e = i\int_{M_4} B^e_2\wedge *J^e_2 = -\frac{1}{g^2}\int_{M_4} B^e_2\wedge *F_2 \qquad \mathcal{S}_{c.t.} = \frac{1}{2g^2}\int_{M_4} B^e_2\wedge *B^e_2 \tag{3.60}$$

with appropriate coupling parameters. For the magnetic 1-form symmetry, the coupling term is

$$\mathcal{S}_m = i\int_{M_4} B^m_2\wedge *J^m_2 = \frac{i}{2\pi}\int_{M_4} B^m_2\wedge F_2 \tag{3.61}$$

There is no counter term because the magnetic symmetry does not act on the dynamical field A_1 and there is no need to compensate for that. The final action looks like

$$\mathcal{S}[B^e, B^m] = \int_{M_4}\frac{1}{2g^2}(F_2 - B^e_2)\wedge *(F_2 - B^e_2) + \frac{i}{2\pi}B^m_2\wedge F_2 \tag{3.62}$$

with two sets of gauge transformations:

$$B^e_2 \to B^e_2 + d\Lambda^e_1 \qquad A_1 \to A_1 + \Lambda^e_1 \tag{3.63}$$

$$B^m_2 \to B^m_2 + d\Lambda^m_1 \tag{3.64}$$

Assuming we have gauged both of the 1-form symmetries successfully, the action is expected to be invariant under these transformations. Putting 3.63 into action 3.62:

$$\begin{aligned}\mathcal{S} &\to \int_{M_4} \frac{1}{2g^2}(F_2 - B_2^e) \wedge *(F_2 - B_2^e) + \frac{i}{2\pi} B_2^m \wedge (F_2 + d\Lambda_1^e) \\ &= \mathcal{S} + \frac{i}{2\pi} \int_{M_4} B_2^m \wedge d\Lambda_1^e\end{aligned}$$

shows that gauging of $U(1)_e$ is not successful. For completeness, check 3.64 as well:

$$\begin{aligned}\mathcal{S} &\to \int_{M_4} \frac{1}{2g^2}(F_2 - B_2^e) \wedge *(F_2 - B_2^e) + \frac{i}{2\pi}(B_2^m + d\Lambda_1^m) \wedge F_2 \\ &= \mathcal{S} - \frac{i}{2\pi} \int_{M_4} \Lambda_1^m \wedge dF_2 = \mathcal{S}\end{aligned}$$

following from Bianchi Identity. The gauging of magnetic 1-form symmetry seems successful. It is possible to force 3.63 to be a symmetry by adding counter term, i.e.

$$\mathcal{S}_{c.t.} = -\frac{i}{2\pi} \int_{M_4} B_2^m \wedge B_2^e \qquad \mathcal{S}' = \int_{M_4} \frac{1}{2g^2}(F_2 - B_2^e) \wedge *(F_2 - B_2^e) + \frac{i}{2\pi} B_2^m \wedge (F_2 - B_2^e) \tag{3.65}$$

If we go back and check 3.64, we find

$$\begin{aligned}\mathcal{S}' &\to \int_{M_4} \frac{1}{2g^2}(F_2 - B_2^e) \wedge *(F_2 - B_2^e) + \frac{i}{2\pi}(B_2^m + d\Lambda_1^m) \wedge (F_2 - B_2^e) \\ &= \mathcal{S}' - \frac{i}{2\pi} \int_{M_4} d\lambda_1^m \wedge B_2^e\end{aligned}$$

The main takeaway is that gauging both of the 1-form symmetries simultaneously is impossible. The action and subsequently, the partition function is not invariant under gauge transformation. This is usually addressed as a mixed 't Hooft anomaly of the theory. 't Hooft anomaly is an RG-flow invariant quantity that can tell us about behaviour of the theory in strongly coupled region. The concept has abundant applications albeit we are not going to use it extensively in this paper. Interested reader should check [13] and [15] for more details

Chapter 4

Wilson Lines and 't Hooft Lines in Four Dimensions

Wilson Lines and 't Hooft Lines are introduced in chapter 2 as the charged objects under 1-form symmetries. At the same time they are also basic objects in gauge theory. They are probes that reveal the different spectrum of otherwise indistinguishable gauge groups. In some way they serve as connections between gauge theory and the language of higher-form symmetries. In this chapter, we will continue to follow mathematical rendition of gauge theory to derive some properties of line operators under different gauge groups. Then we will discuss the line spectrum of $SU(N)$ and $PSU(N)$ Yang-Mills through group theory. Note that the setting is in 4d for this chapter.

4.1 What is Wilson Line?

Where does Wilson Line come from? Experimentally, Wilson Line (or rather, loop) comes out naturally in a phenomenon called Aharonov–Bohm effect. An introduction can be found in [15] and [2]. [2] also provides an explanation emerging from fibre bundle picture. It is reproduced as follow. Let's return to the Principal Fibre Bundle $P(M, G)$. When a path is given on base manifold M, we can ask for the movement of fibre along the path. First define:

Definition 4.1. On a principal bundle $P(M, G)$ given a path $\gamma : [0, 1] \to M$, the curve $\tilde{\gamma} : [0, 1] \to P$ is the **Horizontal Lift** of γ if $\pi \circ \tilde{\gamma} = \gamma$ and the tangent vector of $\tilde{\gamma}(t)$ belongs to $H_{\tilde{\gamma}(t)}P$.

Given a curve on M, the choice of horizontal lift is far from unique. However, we can fix a horizontal lift by fixing its initial point. This is a result granted by the fundamental theorem of ordinary differential equation. We claim without proof that:

Theorem 4.2. *Given $\gamma : [0, 1] \to M$ and $u_0 \in \pi^{-1}(\gamma(0))$, there exists a unique horizontal lift $\tilde{\gamma}$ such that $\tilde{\gamma}(0) = u_0$.*

Define a reference section s_i on a patch U_i that contains γ. Without loss of generality, assume $s_i(\gamma(0)) = u_0$. We can then express $\tilde{\gamma}$ with respect to the reference section by writing $\tilde{\gamma}(t) = s_i(\gamma(t))g_i(\gamma(t))$ with $g_i(t) \in G$. Denote X as the tangent vector to γ at

$\gamma(t)$. Then $\tilde{X} = \tilde{\gamma}(t)_* X$ is the tangent vector to $\tilde{\gamma}$ at $\tilde{\gamma}(t)$. It follows, similar to the derivation of 3.11, that:

$$\begin{aligned}\tilde{X} &= \tilde{\gamma}(t)_* X \\ &= \frac{d}{dt}(s_i(\gamma(t))g_i(\gamma(t))) \\ &= g_i^{-1}(t)s_{i*}Xg_i(t) + (g_i^{-1}(t)dg_i(X))^{\#}\end{aligned}$$

Then using the property that $\tilde{X} \in H_{\tilde{\gamma}(t)}P$:

$$\begin{aligned}0 &= \omega(\tilde{X}) \\ &= g_i^{-1}(t)\omega(s_{i*}X)g_i(t) + g_i^{-1}(t)dg_i(X) \\ &= g_i^{-1}(t)\mathcal{A}_i(X)g_i(t) + g_i^{-1}(t)dg_i(X)\end{aligned}$$

A differential equation is recovered:

$$\frac{dg_i(t)}{dt} = -\mathcal{A}_i(X)g_i(t) = -\mathcal{A}_{i\mu}\frac{dx^\mu}{dt}g_i(t) \tag{4.1}$$

This equation describes how the horizontal lift $\tilde{\gamma}$ moves in the fibre with respect to the reference section s_i, which is really just a reference because it is trivialized to identity $(\gamma(t), e) \in U_i \times G$ if canonical trivialization is chosen. Attempting to solve this equation in a general (non-Abelian) case yields an infinite series:

$$\begin{aligned}g_i(\gamma(t)) &= -1 - \int_0^t dt' \mathcal{A}_{i\mu}\frac{dx^\mu}{dt'} - \int_0^t dt' \int_0^{t'} dt'' \mathcal{A}_{i\mu}\frac{dx^\mu}{dt'}\mathcal{A}_{i\nu}\frac{dx^\nu}{dt''}... \\ &= -\sum_{n=0}^{\infty}\int_0^t dt' \int_0^{t'} dt''... \int_0^{t^{(n-1)}} dt^{(n)} \mathcal{A}_{i\mu_1}\frac{dx^{\mu_1}}{dt'}\mathcal{A}_{i\mu_2}\frac{dx^{\mu_2}}{dt''}...\mathcal{A}_{i\mu_n}\frac{dx^{\mu_n}}{dt^{(n)}} \\ &\equiv \mathcal{P}\exp\left(-\int_0^t dt\mathcal{A}_{i\mu}\frac{dx^\mu}{dt}\right) \\ &= \mathcal{P}\exp\left(-\int_{\gamma(0)}^{\gamma(t)} dx^\mu \mathcal{A}_{i\mu}(\gamma(t))\right)\end{aligned}$$

The infinite series took the form of an exponential. The connection is however not commutative in general so we need to define a path order operator $\mathcal{P}$ to take care of the order. The operator is similar to the time order operator seen in conventional treatment of QFT. It puts the connection at smaller t in front of connection at larger t in every term of the expansion. Taking the physical convention 3.15, we have

$$g_i(t) = \mathcal{P}\exp\left(-i\int_\gamma A_1\right) \tag{4.2}$$

the non-Abelian Wilson Line operator. If the group is Abelian, $\mathcal{P}$ can be safely removed. Notice that we first introduced Wilson Line in Chapter 2 as the world line of an infinitely heavy electric source particle. We have derived the Wilson Line here without including any charged matter in the picture. It is simply the parallel transportation along a curve in the principal fibre bundle and exists as a basic feature. It is also clear that Wilson Line is an element of the group because from 3.4 we saw that $\int_\gamma \mathcal{A}_i$ is an element of the Lie algebra. We can alternatively denote a Wilson Line along Γ_1 parameterized by $t \in [0,1]$ that starts at point x_i and ends at point x as

$$W(x_i, x(t); \Gamma) = \mathcal{P} \exp\left(i \int_{\Gamma(0)=x_i}^{\Gamma(t)=x(t)} A_1\right) \tag{4.3}$$

It has property:

$$W(x_i, x_i; \Gamma) = \mathcal{P} \exp(0) = 1 \tag{4.4}$$

A slight remark: the solution to the ODE (4.2) translates to the new notation as

$$g_i(t) = \mathcal{P} \exp\left(-i \int_\gamma A_1\right) = \mathcal{P} \exp\left(i \int_{-\gamma} A_1\right) = W(x(t), x_i; \Gamma) \tag{4.5}$$

The swapping of order has no particularly deep reason. It only has to do with how the ODE was set up in the first place. We would like to know how Wilson Line transforms. Under the component form of 3.16:

$$A_\mu(x) \to A'_\mu(x) = g^{-1}(x) A_\mu(x) g(x) + i\partial_\mu g^{-1}(x) g(x) \tag{4.6}$$

We claim that the Wilson Line transforms as:

$$W(x_i, x; \Gamma) \to W'(x_i, x; \Gamma) = g^{-1}(x_i) W(x_i, x; \Gamma) g(x) \tag{4.7}$$

Here is a proof: Recall the ODE of parallel transport 4.1, rewritten here as

$$\frac{dx^\mu}{dt} \partial_\mu W(x, x_i; \Gamma) = -iA_\mu \frac{dx^\mu}{dt} W(x, x_i; \Gamma)$$
$$\frac{dx^\mu}{dt} (\partial_\mu + iA_\mu) W(x, x_i; \Gamma) = 0$$

where ∂_μ is a notation for $\frac{\partial}{\partial x^\mu}$. We need the gauge transformed W' and A' to still satisfy the ODE. Inserting 4.7 and 4.6:

$$\begin{aligned}
&\frac{dx^\mu}{dt}(\partial_\mu + iA'_\mu)W'(x,x_i;\Gamma) \\
&= \frac{dx^\mu}{dt}(\partial_\mu + ig^{-1}(x)A_\mu(x)g(x) - \partial_\mu g^{-1}(x)g(x))g^{-1}(x)W(x,x_i;\Gamma)g(x_i) \\
&= \frac{dx^\mu}{dt}\{\partial_\mu g^{-1}(x)W(x,x_i;\Gamma)g(x_i) + g^{-1}(x)\partial_\mu W(x,x_i;\Gamma)g(x_i) + ig^{-1}(x)A_\mu(x)W(x,x_i;\Gamma)g(x_i) \\
&\quad - \partial_\mu g^{-1}(x)W(x,x_i;\Gamma)g(x_i)\} \\
&= \frac{dx^\mu}{dt}(g^{-1}(x)\partial_\mu W(x,x_i;\Gamma)g(x_i) + ig^{-1}(x)A_\mu(x)W(x,x_i;\Gamma)g(x_i)) \\
&= g^{-1}(x)\frac{dx^\mu}{dt}(\partial_\mu + iA_\mu(x))W(x,x_i;\Gamma)g(x_i) \\
&= 0
\end{aligned}$$

Also the solution of W' obeys a unique boundary condition

$$W'(x,x_i;\Gamma)|_{t=0} = W'(x_i,x_i;\Gamma) = 1 \tag{4.8}$$

Check from 4.4 that

$$W'(x_i,x_i;\Gamma) = g^{-1}(x_i)W(x_i,x_i;\Gamma)g(x_i) = g^{-1}(x_i)g(x_i) = 1 \tag{4.9}$$

for arbitrary g. Thus from the fundamental theorem of Ordinary Differential Equation, our solution for W' must be unique. So far, we write Wilson Lines in term of connection valued in an abstract Lie algebra element. In physics, we need to find for it a representation. In a more formal language, A_1 is not the connection on principal bundle but the induced connection on associated bundle. Explicitly,

$$W_R(x_i,x_f;\Gamma) = \mathcal{P}\exp\left(i\int_{\Gamma(0)=x_i}^{\Gamma(1)=x_f}\rho(A_1)\right) \tag{4.10}$$

$$W_R(x_i,x_f;\Gamma) \to R(g(x_i))^{-1}W_R(x_i,x_f;\Gamma)R(g(x_f)) \tag{4.11}$$

where ρ denotes the corresponding Lie algebra representation to R.

4.1.1 Wilson Line and Screening

The Wilson Line is a **non-genuine operator** because it is not gauge invariant. As all physical properties are gauge invariant, we want to construct a genuine operator out of Wilson Line. One way to do that is to connect Wilson Lines with a local operator in the same representation. The action of connecting two Wilson Lines with a local operator is called **Screening**. Screening provides a shortcut to finding the 1-form symmetry of the theory. We will illustrate this with a few examples taken from [9].

Screening in Abelian Higgs Model

Recall that Abelian Higgs model with a charge q Higgs has gauge transformations

$$A_1 \to A_1 + d\theta \qquad \phi_q \to e^{iq\theta}\phi_q \tag{4.12}$$

Consider a Wilson Line in "q" representation that only has one ending point x ($\partial\Gamma = -x$) but extends infinitely on the other end. It transforms as

$$\begin{aligned} W_q(x,\infty;\Gamma) = \exp\left(iq\int_\infty^x A_1\right) &\to \exp\left(iq\int_\infty^x (A_1 + d\theta)\right) \\ &= e^{-iq\theta(x)} W_q(x,\infty;\Gamma) \end{aligned}$$

Inserting ϕ_q at point x cancels the additional term under gauge transformation. In other words, the combination

$$\phi_q(x)W_q(x,\infty;\Gamma) \tag{4.13}$$

is a genuine operator. In the language of screening, this combination can be understood as ϕ_q connecting W_q to a trivial Wilson Line. We say that W_q is **completely screened**. We can define the screening as an equivalent relation

$$W_q \sim 1 \tag{4.14}$$

Similarly by adding the n-th order local field

$$\phi_q^n(x) \to e^{inq\theta(x)}\phi_q^n(x) \tag{4.15}$$

we can construct

$$W_p(-\infty,x;\Gamma')\phi_q^n(x)W_r(x,\infty;\Gamma) \to e^{i(p+nq-r)\theta(x)}W_p(-\infty,x;\Gamma')\phi_q^n(x)W_r(x,\infty;\Gamma) \tag{4.16}$$

connecting two Wilson Lines. The screening in Abelian Higgs Model is, in general, given by the condition $r = p + nq$:

$$W_{p+nq} \sim W_p \qquad p,n,q \in \mathbb{Z} \tag{4.17}$$

Screening in SU(2) Yang-Mills

The Maxwell field strength is gauge invariant. For that reason it can not screen any Wilson Line. The field strength of Yang-Mills Theory, however, transforms in the adjoint representation:

$$F_2 \to g^{-1}F_2 g = ad(g)F_2 \tag{4.18}$$

That means the field strength provides complete screening for Wilson Line in the adjoint representation. Recall that for $SU(2)$, representations are marked by integers and half-integers. We can assign each representation a spin number

$$j \in \frac{\mathbb{Z}}{2} \tag{4.19}$$

By adding order of F_2 it is possible to screen any integer value of spin. The screening is described by

$$W_{j+n} \sim W_j \qquad n \in \mathbb{Z} \tag{4.20}$$

There are only two equivalent classes $[0]$ and $[\frac{1}{2}]$.

Consequence of Screening: in the Language of Generalized Symmetry

Recall in Chapter 2 we have established that the representations of Abelian group G is labeled by the Pontryagin dual group $\widehat{G}$. The Pontryagin dual group of finite Abelian group is easy to find. Taken from [9],

Theorem 4.3. *For any finite Abelian Group $G^{(p)}$, $\widehat{G^{(p)}} \cong G^{(p)}$*

Along with property $\widehat{\widehat{G}} \cong G$, the mechanism of screening allows us to immediately find the group of 1-form electric symmetry once we know the complete set of Wilson lines and screening relations. First we claim: **If two line operators are equivalent via screening, they have the same 1-form charge.**

The proof is mostly visual. 1-form charge of a line is given by an SDO wrapping around it. If the line is screened to another line by a local operator, we can use the topological nature of the SDO to move it to the other line and contracts there instead. The resulting statement is that both of the lines must give the same charge after the symmetry action, otherwise the action is no longer topological. The best way to see this is through a picture Fig.4.1 taken from [9]:

It is then clear that the correct elements in the Pontryagin Dual group are equivalent classes of Wilson Lines. Let's take the last two examples and evaluate their 1-form group:

In the Abelian Higgs case, the original group of representation is $\widehat{U(1)} \cong \mathbb{Z}$. The existence of charge q Higgs field results in a quotient group

$$\widehat{G^{(1)}} \cong \frac{\mathbb{Z}}{q\mathbb{Z}} \cong \mathbb{Z}_q \tag{4.21}$$

Immediately the 1-form symmetry group is

$$G^{(1)} \cong \widehat{\mathbb{Z}_q} \cong \mathbb{Z}_q \tag{4.22}$$

In $SU(2)$ Yang-Mills, the Wilson Line spectrum is $\frac{\mathbb{Z}}{2}$ quotient $\mathbb{Z}$, it follows that

$$G^{(1)} \cong \widehat{\left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)} \cong \widehat{\mathbb{Z}_2} \cong \mathbb{Z}_2 \tag{4.23}$$

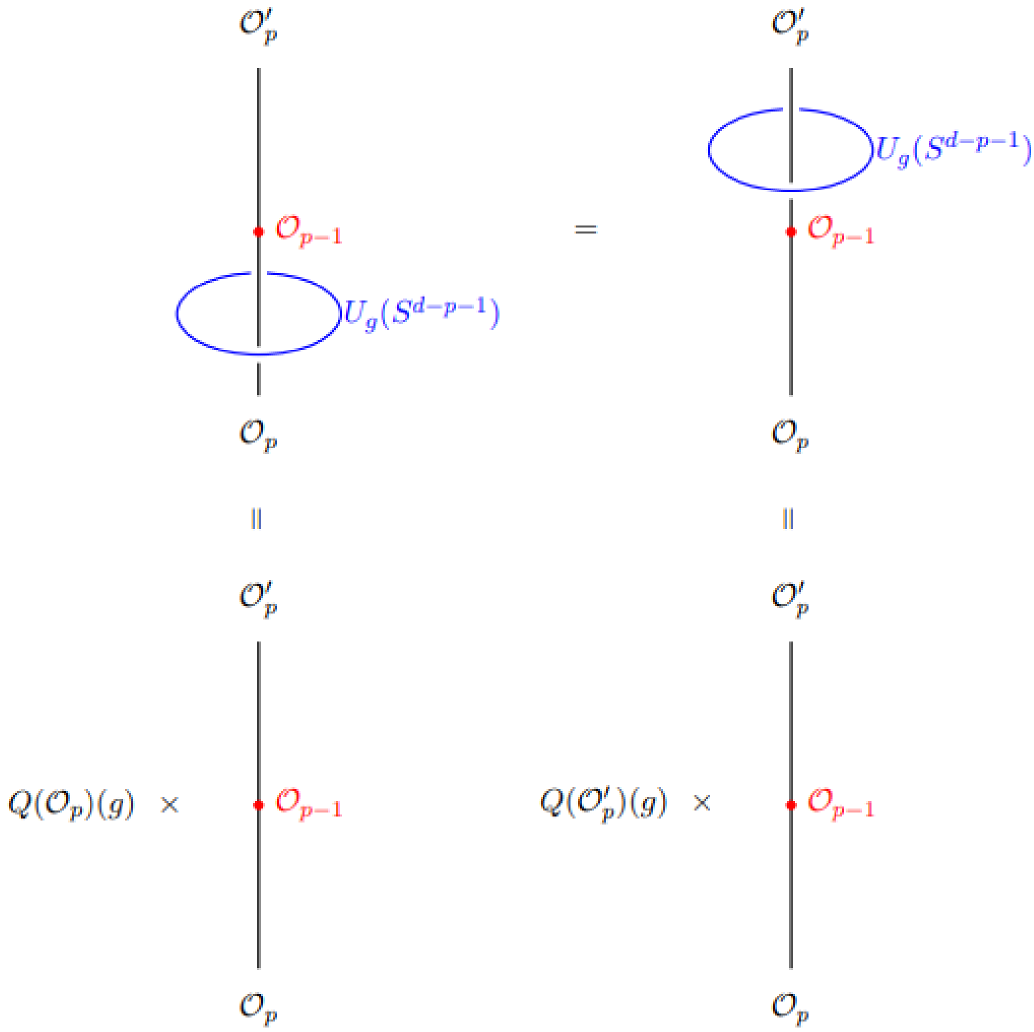

Figure 4.1: Screened lines have the same charge

Beyond Wilson Line

So far we have looked at Wilson Lines in higged $U(1)$ and $SU(2)$. They are simple because $U(1)$ and $SU(2)$ have simple sets of representations. Is there a more general method that applies to $SU(N)$ with $N > 2$?

Wilson Lines, when seen as the world line of a charged particle, is clearly labeled by (the particle's) representation of the gauge group. It is less clear how 't Hooft Lines are labeled. A slight spoiler: 't Hooft Lines are also in representation of a group but not the original one. To properly explain these questions, a reasonable starting point would be a review on representation theory of Lie groups.

4.2 A Review on Lie Theory

This section assumes the readers have some amount of knowledge on Lie Theory and runs relatively quickly. We refer the readers to [4] and [16] for the full mathematical construction.

We start with a complex, semi-simple Lie algebra $\mathbf{g}$. Recall that it admits a Cartan subalgebra $\mathbf{h} \subset \mathbf{g}$ that is maximal and Abelian. The Cartan provides a decomposition of the module. Pick a representation $\rho : \mathbf{g} \to End(V)$ where V is a vector space. Denote $\mathbf{h}^*$ as the dual space of the Cartan subalgebra: $\mathbf{h}^* : \mathbf{h} \to \mathbb{C}$. Recall that $\mathbf{w} \in \mathbf{h}^*$ is a weight of $\mathbf{g}$ if $\exists v \in V$ such that

$$\rho(H)(v) = \mathbf{w}(H)v \qquad \forall H \in \mathbf{h} \tag{4.24}$$

If V is the Lie algebra itself, the corresponding representation is the adjoint representation $Ad : \mathbf{g} \to End(\mathbf{g})$. The weights are called roots: $\alpha \in \mathbf{h}^*$ and for $G \in \mathbf{g}$,

$$Ad(H)(G) = \alpha(H)G \qquad \forall H \in \mathbf{h} \tag{4.25}$$

If the Lie algebra is expressed in matrices, these are essentially eigenvalue equations. The weight and root equations provides complete decompositions of the corresponding vector spaces V and $\mathbf{g}$.

$$V = \bigoplus_{\mathbf{w}} V_{\mathbf{w}} \tag{4.26}$$

$$\mathbf{g} = \mathbf{h} \oplus \bigoplus_{\alpha} \mathbf{g}_\alpha \tag{4.27}$$

where $V_{\mathbf{w}}$ and $\mathbf{g}_\alpha$ are subspaces with eigenvalues $\mathbf{w}$ and α. $\mathbf{h}$ has zero eigenvalue. Denote the set of all roots as Δ_α and the set of all weights as $\Delta_{\mathbf{w}}$, we can define the root lattice and the weight lattice:

$$\Lambda_\alpha(\mathbf{g}) = Span(\Delta_\alpha) = \{ \sum_{\alpha_{(i)} \in \Delta_\alpha} n_i \alpha_{(i)} | n_i \in \mathbb{Z} \} \tag{4.28}$$

$$\Lambda_{\mathbf{w}}(\mathbf{g}) = Span(\Delta_{\mathbf{w}}) = \{ \sum_{\mathbf{w}_{(i)} \in \Delta_{\mathbf{w}}} n_i \mathbf{w}_{(i)} | n_i \in \mathbb{Z} \} \tag{4.29}$$

They are also addition groups by construction. Moreover, any representation of $\mathbf{g}$ can be generated by subtracting roots from the a highest weight $\mathbf{w}_\rho$ of the representation. But if we ask what elements of the weight lattice uniquely define a representation, the answer is in fact an element of the weight lattice modulo the Weyl group $\lambda_\rho \in \Lambda_{\mathbf{w}}(\mathbf{g})/W(\mathbf{g})$. In other words, the label λ_ρ is a equivalent class $\lambda_\rho = [\mathbf{w}_\rho]$ defined by the equivalent relation $\mathbf{w} \sim s_\alpha(\mathbf{w})$ where s_α is the Weyl group action.

4.3 Line Spectrum of SU(N)

Most of the content in the following sections owes to brilliant explanation of the matter in [15] as well as [19] and [20]. Recall that in Chapter 3 we constructed a Dirac monopole for $U(1)$ gauge theory that provides a Dirac quantization condition. The key is a transition function

$$t_{NS}(\phi) = \exp\left(i\alpha\sigma(\phi)\right) \tag{4.30}$$

Similar treatment can be done on $\mathcal{G} \cong SU(N)$. In a vector bundle associated to principal bundle $P(m, \mathcal{G})$, the transition function is an element of a specified representation R of $SU(N)$. It can be written as

$$t_{NS}(\phi) = \exp\left(i\rho(\mathbf{s}(\phi))\right) \tag{4.31}$$

where $\mathbf{s}$ is an element of the abstract Lie algebra $\mathbf{g} = Lie(\mathcal{G})$ and ρ the corresponding Lie algebra representation. Note that this correspondence only works here because $SU(N)$ is simply-connected and has same representations as its algebra. This will come back to haunt us later. The gauge fields configuration is also similar but $\mathbf{m}$ is now Lie algebra valued and also in representation ρ:

$$A_1^N = \frac{1}{2}\rho(\mathbf{m})(1 - cos\theta)d\phi \qquad A_1^S = -\frac{1}{2}\rho(\mathbf{m})(1 + cos\theta)d\phi \tag{4.32}$$

We claim that a gauge transformation can choose $\mathbf{m}$ to be in the Cartan subalgebra of $\mathbf{g}$. For this reason $\mathbf{s}$ would also be in $\mathbf{h}$. Following from 3.18, the compatibility condition reads

$$A_1^N = t_{NS}^{-1} A_1^S t_{NS} + \rho(d\mathbf{s}) \tag{4.33}$$

Expand, we have

$$\begin{aligned}
\frac{1}{2}\rho(\mathbf{m})(1 - cos\theta)d\phi &= -\frac{1}{2}(1 + cos\theta)d\phi\; e^{-i\rho(\mathbf{s})}\rho(\mathbf{m})e^{i\rho(\mathbf{s})} + \rho(d\mathbf{s}) \\
\rho(\mathbf{m})(1 - cos\theta)d\phi &= -(1 + cos\theta)d\phi\; (\; \rho(\mathbf{m}) + [\rho(\mathbf{m}), \rho(\mathbf{s})]\;) + 2\rho(d\mathbf{s}) \\
\rho(\mathbf{m})d\phi &= \rho(d\mathbf{s})
\end{aligned}$$

where we have use the homomorphic identity and that both $\mathbf{m}$ and $\mathbf{s}$ are in the Abelian Cartan subalgebra:

$$[\rho(\mathbf{m}), \rho(\mathbf{s})] = \rho([\mathbf{m}, \mathbf{s}]) = 0 \tag{4.34}$$

Integrate both sides:

$$\int \rho(d\mathbf{s}) = \rho(\Delta\mathbf{s}) = \int_0^{2\pi} \rho(\mathbf{m})d\phi = 2\pi\rho(\mathbf{m}) \tag{4.35}$$

The resulting single-valued condition is

$$e^{i2\pi\rho(\mathbf{m})}v = v \tag{4.36}$$

here $v \in V$ is a vector in the representation ρ. Since $\mathbf{m} \in \mathbf{h}$, from 4.24:

$$e^{i2\pi\rho(\mathbf{m})}v = e^{i2\pi\mathbf{w}(\mathbf{m})}v \tag{4.37}$$

and

$$\mathbf{w}(\mathbf{m}) = w^i m_i = \mathbf{w} \cdot \mathbf{m} \overset{!}{\in} \mathbb{Z} \tag{4.38}$$

This is familiar because there is a similar relation between root and weight:

$$\frac{2\alpha \cdot \mathbf{w}}{\alpha^2} \in \mathbb{Z} \tag{4.39}$$

This follows from the Cartan-Weyl basis which states that the algebra can be decomposed into smaller $SU(2)$ algebra. On top of that we can define co-root to be

$$\alpha^\vee := \frac{2\alpha}{\alpha^2} \tag{4.40}$$

and this is the quantity that labels magnetic charge in $SU(N)$ theory. Like root and weight, co-root also spans a lattice

$$\Lambda_{\alpha^\vee}(\mathbf{g}) = \{ \sum_{\alpha_i^\vee \in \Delta_{\alpha^\vee}} n_i \alpha_i^\vee | n_i \in \mathbb{Z} \} \tag{4.41}$$

The condition 4.38 looks like Dirac quantization condition. It is named GNO quantization after the paper [26] of Goddard, Nuyts, and Olive. They also found that there exists a Lie algebra $\mathbf{g}^\vee$ such that

$$\Lambda_{\alpha^\vee}(\mathbf{g}) = \Lambda_\alpha(\mathbf{g}^\vee) \tag{4.42}$$

This algebra is called the GNO dual of the original algebra. To interpret the GNO dual graphically, the root lattice of $\mathbf{g}$ and $\mathbf{g}^\vee$ can be interchanged by swapping long and short roots (Fig.4.2). The Weyl group of GNO dual algebra is the same as the Weyl group of original algebra

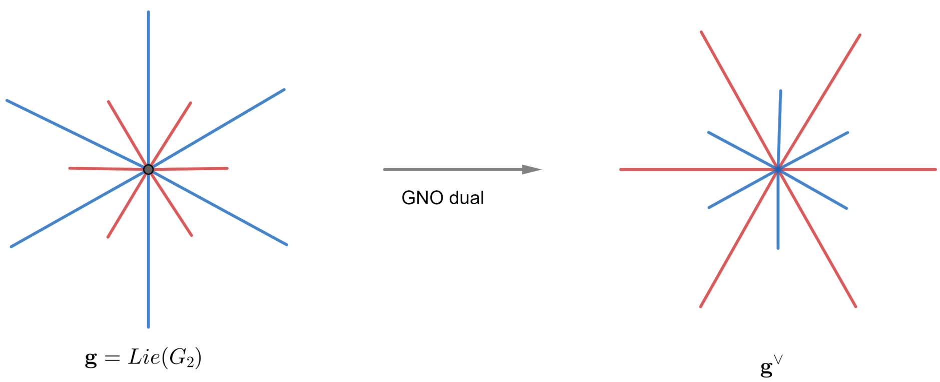

Figure 4.2: GNO dualization swaps long and short roots

$$W(\mathbf{g}^\vee) = W(\mathbf{g}) \tag{4.43}$$

This is not difficult to spot out because swapping long and short roots does not change the symmetry of the lattice. For $\mathbf{g} = su(N)$, all roots have the same length. This feature is called "simply-laced". Since the algebra of $SU(N)$ is simply-laced, $\mathbf{g}^\vee = \mathbf{g}$ and $\Lambda_{\alpha^\vee}(\mathbf{g}) = \Lambda_\alpha(\mathbf{g})$

We have found that all charged line operators in $SU(N)$ lies in the group [1]

$$\frac{\Lambda_{\mathbf{w}}(\mathbf{g})}{W(\mathbf{g})} \times \frac{\Lambda_\alpha(\mathbf{g}^\vee)}{W(\mathbf{g}^\vee)} \tag{4.44}$$

The element of the space looks like

$$[(\lambda_e, \lambda_m)] \qquad (\lambda_e, \lambda_m) \sim (s\lambda_e, s\lambda_m) \tag{4.45}$$

There are further restriction on the spectrum. Recall that in $SU(2)$ the existence of the field strength completely screen the adjoint Wilson Line. Analogically, the adjoint local operators always exist for both electric and magnetic group in $SU(N)$. They can be represented by $\Lambda_\alpha(\mathbf{g})/W(\mathbf{g})$ and $\Lambda_\alpha(\mathbf{g}^\vee)/W(\mathbf{g}^\vee)$. They should be quotient out from the main group, thus

$$\frac{\Lambda_{\mathbf{w}}(\mathbf{g})/W(\mathbf{g})}{\Lambda_\alpha(\mathbf{g})/W(\mathbf{g})} \times \frac{\Lambda_\alpha(\mathbf{g}^\vee)/W(\mathbf{g}^\vee)}{\Lambda_\alpha(\mathbf{g}^\vee)/W(\mathbf{g}^\vee)} = \frac{\Lambda_{\mathbf{w}}(\mathbf{g})}{\Lambda_\alpha(\mathbf{g})} \times 1 \tag{4.46}$$

We claim without proof that $\Lambda_{\mathbf{w}}(\mathbf{g})/\Lambda_\alpha(\mathbf{g}) \cong C(SU(N)) \cong \mathbb{Z}_N$ where $C(\cdot)$ means the center of the group. The resulting set of lines are

$$(z_e, 0) \in \mathbb{Z}_N^e \times 1 \qquad z_e \in \{0, 1, 2, ..., N-1\} \tag{4.47}$$

There is an additional constraint due to the requirement for two different lines to be mutually local. It states that for two line operators (n, m) and (n', m'),

$$nm' - mn' = 0 \; mod \; N \tag{4.48}$$

This is a generalized version of a condition on two dyons (particles with both electric and magnetic charges) called Dirac-Zwanziger-Schwinger quantization. It is not very relevant here because the magnetic charge group is trivial in $SU(N)$ and the condition is always satisfied.

4.4 SU(N) versus PSU(N)

We would like to generalize what we have done on $SU(N)$ in the last section to a related group $PSU(N) \cong SU(N)/\mathbb{Z}_N$. The choice of $PSU(N)$ has a historic reason. In older material it is sometimes stated that the correct group of Yang-Mills theory is actually $PSU(N)$ instead of $SU(N)$. The two gauge groups share the same Lie algebra $su(N)$ and thus the same action (if discrete gauge field is out of the picture). The gauge boson is also blind to the difference because it is in the adjoint representation, which is again shared by

[1] More accurately, the group is actually $\frac{\Lambda_{\mathbf{w}}(\mathbf{g}) \times \Lambda_\alpha(\mathbf{g}^\vee)}{W(\mathbf{g})}$. i.e. the final group of lines are not just simple product of pure electric and magnetic lines. We will not specify the distinction here. Interested readers can find an explanation in [20].

the two groups. In fact, all correlation functions of local operators are the same for two theory, as stated in [15]. That doesn't mean the two theories are completely identical. When put onto topologically non-trivial manifold, the two have distinct behaviours. We will not go into this topic. Instead, we will see how line operators can be used as tools to probe the difference between them. $SU(N)/\mathbb{Z}_N$ is not the only group of this kind. It is possible to discuss line spectrum of $SU(N)/\mathbb{Z}_k$ with k a divisor of N. It is also possible to go beyond $SU(N)$ into $SO(N)$ and $Sp(N)$ gauge theories. These are all studied in [19].

A defining characteristic of $PSU(N)$ is that it only admits a portion of the representations of $su(N)$. Going back to 4.31, we claimed that ρ is the corresponding representation to the group representation R. In the case of $SU(N)/\mathbb{Z}_N$, only a part of the representations of $su(N)$ is allowed to be selected, resulting in a modification of 4.37 to

$$e^{i2\pi\rho(\mathbf{m})}v = e^{i2\pi\alpha(\mathbf{m})}v \tag{4.49}$$

where α is the root. Why? Recall that we claim the following statement is true

$$\Lambda_{\mathbf{w}}(\mathbf{g})/\Lambda_{\alpha}(\mathbf{g}) \cong C(SU(N)) \cong \mathbb{Z}_N \tag{4.50}$$

Conversely,

$$\Lambda_{\mathbf{w}}(\mathbf{g})/\mathbb{Z}_N \cong \Lambda_{\alpha}(\mathbf{g}) \tag{4.51}$$

which is stating that the weight lattice of $SU(N)/\mathbb{Z}_N$ is simply the root lattice of $SU(N)$. In the associated bundle of $PSU(N)$, only adjoint representation is present, thus

$$\rho(\mathbf{m})v = ad(\mathbf{m})v = \alpha(\mathbf{m})v \tag{4.52}$$

The resulting GNO quantization condition is $\alpha \cdot \mathbf{m} \in \mathbb{Z}$, which identifies the magnetic charge as a "co-weight." This is usually named magnetic weight of $\mathbf{g}$. Equivalently, it is also the weight of $\mathbf{g}^\vee$.

$$\Lambda_{\mathbf{mw}}(\mathbf{g}) = \Lambda_{\mathbf{w}}(\mathbf{g}^\vee) \tag{4.53}$$

Repeating what we have done in the last section, we find that the line operators of $PSU(N)$ are in group

$$\frac{\Lambda_{\alpha}(\mathbf{g})}{W(\mathbf{g})} \times \frac{\Lambda_{\mathbf{w}}(\mathbf{g}^\vee)}{W(\mathbf{g}^\vee)} \tag{4.54}$$

After screening, they are elements

$$(0, z_m) \in 1 \times \mathbb{Z}_N^m \qquad z_m \in \{0, 1, 2, ..., N-1\} \tag{4.55}$$

We can put the two sets of different line spectrum in the same group

$$\mathbb{Z}_N^e \times \mathbb{Z}_N^m \tag{4.56}$$

The lines in $SU(N)$ occupy $(z_e, 0)$ and the lines in $PSU(N)_+$ occupy $(0, z_m)$. We are adding a subscript $+$ on $PSU(N)_+$ because this is just one solution to the locality constraint of $PSU(N)$. There is another theory $PSU(N)_-$ with a different spectrum still hidden from us. We will explore it in the next section.

We see that the two gauge groups are the opposite of each other in the sense that $SU(N)$ has the most electric charges while $PSU(N)_+$ flipped the spectrum to obtain the most magnetic charges. The GNO quantization condition set a balance in which more electric charges (irreps) leads to more restriction on magnetic charges. If the group of interest is $SU(N)/\mathbb{Z}_k$ where k is some divisor of N, we will find a spectrum that is somewhere inbetween $SU(N)$ and $PSU(N)_+$.

4.4.1 't Hooft Magnetic Flux

Recall that in chapter 3 we classified Dirac magnetic monopole with the fundamental group $\Pi_1(U(1))$. In [20], this is referred to as a classification of 't Hooft operators by its 't Hooft Magnetic Flux. An explanation is provided in [20] to bridge the two classifications. For a simple, compact Lie Group G with Lie algebra $\mathbf{g}$, we define $\tilde{G}$ to be the universal cover of G and G_0 to be the centerless group with algebra $\mathbf{g}$:

$$G_0 \cong \tilde{G}/C(\tilde{G}) \tag{4.57}$$

In the case of $\mathbf{g} = su(N)$, for example, $\tilde{G}$ is $SU(N)$ and G_0 is $PSU(N)$. We define the exponential map:

$$exp : \mathbf{g} \to G \qquad X \mapsto exp(iX) \tag{4.58}$$

The Dirac/GNO quantization condition states that the magnetic charge $\mathbf{m} \in \mathbf{g}$ is in the kernel of this map. The kernel is actually a lattice in $\mathbf{g}$, we denote it as

$$\Lambda_{ker}(G) \in \mathbf{g} \tag{4.59}$$

Also, we know that

$$\Lambda_{ker}(G) \subseteq \Lambda_{\mathbf{mw}}(\mathbf{g}) \qquad \Lambda_{\alpha^\vee}(\mathbf{g}) \subseteq \Lambda_{ker}(G) \tag{4.60}$$

from our previous analysis. Additionally, [20] clarifies that

$$C(G) \cong \Lambda_{\mathbf{mw}}(\mathbf{g})/\Lambda_{ker}(G) \qquad \Pi_1(G) \cong \Lambda_{ker}(G)/\Lambda_{\alpha^\vee}(\mathbf{g}) \cong \Lambda_{ker}(G)/\Lambda_{\alpha}(\mathbf{g}^\vee) \tag{4.61}$$

If we think of $\Lambda_{ker}(G)$ as the weight lattice of some **Group** $G^\vee$ that has Lie algebra $\mathbf{g}^\vee$, the fundamental group is actually representing the center of it

$$\Pi_1(G) = C(G^\vee) \tag{4.62}$$

This is called the **GNO Dual Group**. Notice that the choice of GNO dual group for a particular $\mathbf{g}^\vee$ is not unique. We can now check the two special cases. If G has a trivial center,

$$\Lambda_{ker}(G) = \Lambda_{\mathbf{mw}}(\mathbf{g}) \qquad \Pi_1(G) \cong \Lambda_{\mathbf{mw}}(\mathbf{g})/\Lambda_{\alpha^\vee}(\mathbf{g}) \cong C(\tilde{G^\vee}) \tag{4.63}$$

On the other hand, if $G \cong \tilde{G}$,

$$C(\tilde{G}) \cong \Lambda_{\mathbf{mw}}(\mathbf{g})/\Lambda_{ker}(\tilde{G}) \qquad \Lambda_{ker}(\tilde{G}) = \Lambda_{\mathbf{mw}}(\mathbf{g})/C(\tilde{G}) \tag{4.64}$$

and we can find the fundamental group to be

$$\Pi_1(\tilde{G}) \cong \frac{\Lambda_{\mathbf{mw}}(\mathbf{g})/C(\tilde{G})}{\Lambda_{\alpha^\vee}(\mathbf{g})} \cong \frac{C(\tilde{G^\vee})}{C(\tilde{G})} \tag{4.65}$$

In our special case of $\mathbf{g} = su(N)$, we have

$$\tilde{G^\vee} \cong \tilde{G} \cong SU(N) \tag{4.66}$$

We will recover

$$\Pi_1(SU(N)) \cong 1 \qquad \Pi_1(PSU(N)) \cong \mathbb{Z}_N \tag{4.67}$$

This is exactly what we found in the previous analysis. We have thus shown that we can recover the fundamental group classification via group theory manipulation.

4.5 Witten Effect and Discrete Theta Angle

In the last chapter, we have introduced Witten Effect and how it affects Maxwell Theory. Witten Effect extends to $SU(N)$ and $PSU(N)$ Yang-Mills, and it has interesting interplay with the line spectrum of these theories that leads to profound result. Before a deeper dive, we would like to present a proof for Witten Effect for a general simple, compact gauge group $\mathcal{G}$ in four dimensions. Recall that the theta term is

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2}\int_{M_4} tr(F_2 \wedge F_2) = \frac{\theta}{8\pi^2}\int_{\partial M_4} \kappa_3 \tag{4.68}$$

where κ_3 is the Chern-Simons form

$$\kappa_3 = tr(A_1 \wedge dA_1 + \frac{2}{3}A_1 \wedge A_1 \wedge A_1). \tag{4.69}$$

In general, A_1 is valued in Lie algebra $\mathbf{g}$ and is non-commutative. However, if we only consider a magnetic monopole configuration where A_1 is valued in $\mathbf{m} \in \mathbf{g}$, we have argued before that it is possible to choose $\mathbf{m}$ to be in the Cartan subalgebra such that A_1 commute. We can therefore simplify the expression to

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2}\int_{\partial M_4} tr(dA_1 \wedge A_1) \tag{4.70}$$

We would like to evaluate this term on the "boundary" of a 't Hooft Line. To do that, chose the t'Hooft Line to wrap on a curve Γ_1. Choose M_3 to be the 3-dimensional cut of M_4, transverse to Γ_1. We can then write $\partial M_4 = \partial M_3 \times \Gamma_1$. Also, we can write $F_2 = dA_1$ with commutative A_1, now the theta term is

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2} tr\left(\int_{\partial M_3}\int_{\Gamma_1} F_2 \wedge A_1\right) \tag{4.71}$$

Recall that the magnetic monopole is defined with the identity

$$\begin{aligned} dF_2 &= 2\pi\mathbf{m}\delta_3(x\in\Gamma_1) \\ \int_{M_3} dF_2 &= 2\pi\mathbf{m}\int_{M_3}\delta_3(x\in\Gamma_1) \\ \int_{\partial M_3} F_2 &= 2\pi\mathbf{m}\ Link(\partial M_3, \Gamma_1) \end{aligned}$$

Without loss of generality, we can take the boundary ∂M_3 to be a sphere S_2 and the linking number to be 1. Thus

$$\begin{aligned} \mathcal{S}_\theta &= \frac{\theta}{8\pi^2} tr\left(\int_{S_2} F_2 \wedge \int_{\Gamma_1} A_1\right) \\ &= tr\left(\frac{\theta\mathbf{m}}{4\pi}\int_{\Gamma_1} A_1\right) \\ &= \frac{\theta}{2\pi}\int_{\Gamma_1}\mathbf{m}(A_1) \end{aligned}$$

Which is in the form of a Wilson Line with charge $\frac{\theta\mathbf{m}}{2\pi}$! Adding this term to the action is equivalent to adding an electric source to the system. This is indeed the conclusion of Witten Effect.

Aside: 't Hooft-Polyakov Monopole

The above explanation gives a proper intuition to Witten Effect. It is however not the most rigorous proof available. We have been using Dirac Monopole setting and Wu-Yang forms as our main language to describe soliton in this paper. The major flaw of this language is that the gauge field configuration necessarily admits a singularity at a point in space-time. That means we don't really know what is happening at Γ_1 even though we propose that there is a Wilson Line wrapping on it.

Also, the explicit form of the gauge field 1-form A_1 is unclear. The Wu-Yang forms have basis $d\phi$, which is not well-defined outside of the 3-dimensional cut M_3 and it is unclear what it means to integrate them on the curve Γ_1.

A better formulation of magnetic monopole is the 't Hooft-Polyakov monopole. 't Hooft and Polyakov ([27],[28]) found that there is a monopole solution for a $U(1)$ orbit embedded in an $SU(2)$ gauge group. By using a Higgs field to break $SU(2)$ down to $U(1)$, it is possible to pick out the infinitesimal gauge transformations related to the

$U(1)$ orbit and write them in terms of the Higgs field. Other useful information can be deduced with Noether's theorem on top of that.

For a complete construction of 't Hooft-Polyakov Monopole and how it can be used to prove Witten Effect, see [12].

4.5.1 Discrete Theta and A Tale of Two SO(3) Spectrum

In this section we will look at how Witten Effect affects spectrum of a theory. The major inspirations of this section is again [15] and [19]. Let's look at a simple example of $SU(2)$ and $PSU(2) \cong SO(3)$. The center of $SU(2)$ is $\mathbb{Z}_2$. Thus we can immediately write out the spectrum of the two groups:

$$\mathbb{Z}_2 \times \mathbb{Z}_2 \ni (z_e, z_m) = \begin{cases} (0,0),\ (1,0) \ \ \mathcal{G} \cong SU(2) \\ (0,0),\ (0,1) \ \ \mathcal{G} \cong SO(3) \end{cases} \tag{4.72}$$

Assuming the action of the theory already has a theta term. If we try to use the cyclicity of the theta angle with Witten Effect in place, we find that upon the cyclic transformation:

$$\theta \to \theta + 2\pi \qquad \frac{\theta \mathbf{m}}{2\pi} \to \frac{\theta \mathbf{m}}{2\pi} + \mathbf{m} \qquad (z_e, z_m) \to (z_e + z_m, z_m) \tag{4.73}$$

If $\mathcal{G} \cong SU(2)$, we find that

$$(0,0) \to (0,0) \qquad (1,0) \to (1,0) \tag{4.74}$$

i.e. the cyclicity of the theta parameter is still a symmetry of the system. However, if we use it on the $SO(3)$ theory:

$$(0,0) \to (0,0) \qquad (0,1) \to (1,1) \tag{4.75}$$

which means by dialing the θ parameter by 2π, we have arrived at a new theory that is still an $SO(3)$ Yang-Mills theory but has different lines. In [19] this is summarized as:

$$\begin{cases} SO(3)_+ : \ \ (0,0),\ (0,1) \\ SO(3)_- : \ \ (0,0),\ (1,1) \end{cases} \tag{4.76}$$

What's special about $SO(3)_-$ is that neither fundamental Wilson Line (1,0) nor fundamental 't Hooft Line (0,1) exist. Instead there is an allowed dyon line (1,1). There is a nice way to show this graphically, presented in Fig.4.3 taken from [19].

We can further label the two theories with their theta angles. They are related as following

$$SO(3)_+^{\theta} = SO(3)_-^{\theta+2\pi} \tag{4.77}$$

More generally, $PSU(N)$ theory can be related by a shift of the θ parameter

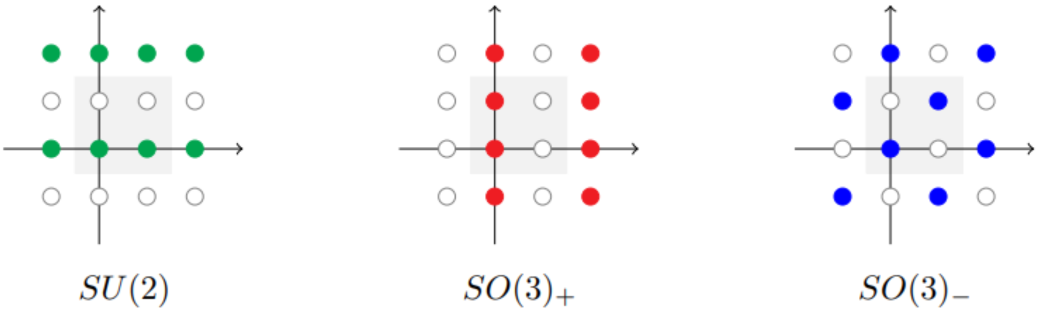

Figure 4.3: Spectrum of theory with g=su(n)

$$PSU(N)^{\theta}_{+} = PSU(N)^{\theta+N\pi}_{-} \tag{4.78}$$

The corresponding between $PSU(2)$ and $SO(3)$, famously, is a happy accident. It was stated in [19] that the shifting rule is not true in general for $SO(N)$ gauge theories. In fact, $SO(N)_+$ and $SO(N)_-$ are not related by a shift of theta angle when $N \geq 5$.

Going back to the $SO(3)$ example, it is clear that even though the theta angle is not 2π periodic, it is 4π periodic:

$$(0,1) \overset{2\pi}{\rightarrow} (1,1) \overset{2\pi}{\rightarrow} (2,1) = (0,1) \tag{4.79}$$

In general, $PSU(N)$ theories have a theta angle that is $2\pi N$ periodic. We can redefine the theta angle

$$\theta \coloneqq 2\pi p \tag{4.80}$$

Now p takes values in

$$p = \{0, 1, ..., N-1\} \tag{4.81}$$

p is N-periodic. This is called discrete theta angle. By dialing the angle we can reach different phases of $PSU(N)$ theory.

We have shown in the last chapter that the periodic property of θ leads to time-reversal symmetry of the theta term at $\theta = \pi$. How does the new discrete theta angle change the picture? We have arrived at the discrete theta angle in a very formal way. In the next chapter, we will try to show it explicitly. In fact, there is a mixed 't Hooft anomaly between the center symmetry and the time-reversal symmetry. To see how this statement unrolls, we will need to find a way to gauge the center with a background gauge field. This time, it needs to be discrete.

Chapter 5

Discrete Gauge Theory

In the final chapter, we would like to discuss how a gauge field with discrete gauge group can be formulated. Hopefully the previous chapters have provided enough motivations for a discrete gauge field: higher-form symmetries are always Abelian. So they must be discrete unless they are just a simple $U(1)$ group. Center symmetry of $SU(N)$ is discrete, so we need a discrete background field to gauge it. Gauging with background fields gives us direct access to the language of 't Hooft anomaly and allow us to track the anomaly along RG-flow... There is one thing left—find it.

5.1 A Crash Course on Algebraic Topology

The language of connection 1-form is not quite compatible with a discrete gauge group. The continuous gauge field is valued in Lie algebra, while for discrete group it is not clear what an algebra is. In a pioneering paper [24] now canonized as the reference for discrete gauge theory, Dijkgraaf and Witten found that the only degree of freedom for a gauge theory with discrete (finite) group G is the topology of the G-bundle over the base manifold. The language to describe it is, naturally, algebraic topology. This section aims to introduce the intuition and mathematical language necessary to understand discrete gauge theory. It will not provide a rigorous mathematical construction of algebraic topology, nor is it possible in the matter of a few pages. For a complete treatment, we recommend the readers to check on the major sources of this section [2], [3] and [24].

5.1.1 Motivation: de Rham's Theorem

Recall that in 3.4, we wrote the connection 1-form as a map

$$\int_\Gamma \mathcal{A}_i \in \mathbf{g} \qquad \mathcal{A}_i : C_1(U_i) \to \mathbf{g} \tag{5.1}$$

where $C_1(U_i)$ is a 1-chain. We can still think of it as a curve for now. This provides us with an opening to approach a discrete gauge theory. Connection 1-form is a space-time 1-form valued in Lie algebra $\mathbf{g}$:

$$\mathcal{A}_i \in \Omega^1_{\mathbf{g}}(U_i) \tag{5.2}$$

Let's look at a simpler object: a normal space-time r-form (valued in $\mathbb{R}$):

$$\omega_r \in \Omega^r(M) \tag{5.3}$$

where we assume M is a compact manifold without loss of generality. For a r-chain $C_r(M)$ (think of it as an r-dim sub-manifold for now), we can define an inner product

$$(\cdot,\cdot): C_r(M) \times \Omega^r(M) \to \mathbb{R} \tag{5.4}$$

$$c, \omega \mapsto \int_c \omega \tag{5.5}$$

Check that the inner product is bi-linear:

$$(c_1 + c_2, \omega) = \int_{c_1+c_2} \omega = \int_{c_1} \omega + \int_{c_2} \omega = (c_1, \omega) + (c_2, \omega) \tag{5.6}$$

$$(c, \omega_1 + \omega_2) = \int_c \omega_1 + \omega_2 = \int_c \omega_1 + \int_c \omega_2 = (c, \omega_1) + (c, \omega_2) \tag{5.7}$$

Take d as the exterior derivative operator on ω and ∂ as the boundary operator on c, the Stoke's Theorem can be written as

$$(c, d\omega) = (\partial c, \omega) \tag{5.8}$$

In this sense the boundary operator is the adjoint operator to exterior derivative. Recall that we can construct a de Rham Cohomology structure on top of forms. The following chain of maps is called a **de Rham Complex** $\Omega^*(M)$

$$\to \Omega^{r-1}(M) \xrightarrow{d_r} \Omega^r(M) \xrightarrow{d_{r+1}} \Omega^{r+1}(M) \to \tag{5.9}$$

Recall that we define the set of close r-forms $Z^r(M)$ and the set of exact r-forms $B^r(M)$ to be

$$Z^r(M) = ker(d_{r+1}) \tag{5.10}$$

$$B^r(M) = im(d_r) \tag{5.11}$$

It is easy to check that they are also groups under form addition. Since $B^r(M) \subseteq Z^r(M)$, a de Rham Cohomology group can be defined to capture this relationship. Define that

$$H^r(M) = Z^r(M)/B^r(M) \tag{5.12}$$

this can be understood as the group of equivalent classes of forms that are closed but not exact. The elements of $H^r(M)$ are

$$H^r(M) = \{[\omega] | \omega \in Z^r(M),\ \omega' \in B^r(M),\ \omega \sim \omega + \omega'\} \tag{5.13}$$

The above construction can be repeated on r-chains. There is a similar **Chain Complex** $C(M)$:

$$\leftarrow C_{r-1}(M) \overset{\partial_r}{\leftarrow} C_r(M) \overset{\partial_{r+1}}{\leftarrow} C_{r+1}(M) \leftarrow \tag{5.14}$$

The r-chains that has no boundary are called "cycles". They are the kernel of ∂_r and form a group

$$Z_r(M) = ker(\partial_r) \tag{5.15}$$

The r-chains that are boundaries of (r+1)-chain are the image of ∂_{r+1}. They also form a group

$$B_r(M) = im(\partial_{r+1}) \tag{5.16}$$

Similarly, a boundary has no boundary and $B_r(M) \subseteq Z_r(M)$. Thus homology group can be constructed as

$$H_r(M) = Z_r(M)/B_r(M) \tag{5.17}$$

Notice that the inner product defined earlier extends to an inner product between cohomology and homology groups. For $c \in Z_r(M)$, $\omega \in Z^r(M)$:

$$\Lambda : H_r(M) \times H^r(M) \to \mathbb{R} \qquad [c], [\omega] \mapsto (c, \omega) = \int_c \omega \tag{5.18}$$

We can check that this is well-defined and compatible with the equivalent classes. For $c' \in C_{r+1}(M)$ and $\omega' \in \Omega^{r-1}(M)$:

$$(c + \partial c', \omega) = (c, \omega) + (c', d\omega) = (c, \omega) \tag{5.19}$$

$$(c, \omega + d\omega') = (c, \omega) + (\partial c, \omega') = (c, \omega) \tag{5.20}$$

This paved the way to **de Rham's Theorem**

Theorem 5.1. *If M is compact, $H^r(M)$ and $H_r(M)$ are finite-dimensional. Also, the map*

$$\Lambda : H_r(M) \times H^r(M) \to \mathbb{R}$$

is bilinear and non-degenerate, thus $H^r(M)$ is the dual vector space of $H_r(M)$

To summarize, we have found that an r-form $\omega \in \Omega^r(M)$ together with the inner product provides a map

$$Hom(C_r(M), \mathbb{R}) \ni (\cdot, \omega) : C_r(M) \to \mathbb{R} \tag{5.21}$$

This is called a cochain. We have also found that the group of closed but not exact forms is dual to the group of chains that are cycles but not boundaries. This is why the name "cohomology" is given to the form structure at the first place. Now we would like to find a discrete version of differential form to describe a discrete gauge theory. The first step requires us to look closer to how the chains are defined.

The notation we have been using for chain is actually an abbreviation:

$$C_r(M) = C_r(M;\mathbb{R}) \tag{5.22}$$

which means r-chains with coefficient in $\mathbb{R}$. This motivates us to find chains with coefficient in discrete groups such as $\mathbb{Z}$ or $\mathbb{Z}_N$ such that the corresponding cochains

$$Hom(C_r(M,\mathbb{Z});\mathbb{Z}) : C_r(M;\mathbb{Z}) \to \mathbb{Z} \tag{5.23}$$

is possibly what we are looking for.

5.1.2 Simplicial Homology

There are a few ways to construct homology theory. Among them Simplicial Homology is more straight forward and serves our purpose of understanding. To motivate the seemingly bizarre concept of a chain group, let's look at an example similar to the one given by [3]. Consider a space shown here

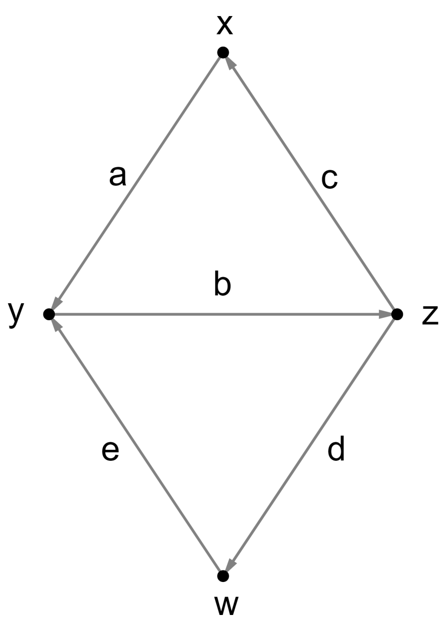

Figure 5.1: Example of a space

The space consists of four endpoints x, y, z, w and five oriented edges a, b, c, d, e. Suppose we want to calculate the homotopy group of this space, we need to pin down a base point and find a loop that starts and ends at it. For example, a loop with base point x that goes through the upper triangle counter-clockwise can be expressed as abc. The order of the multiplication matters because if the order is changed to bca, the base point is changed to y. Take a look at an even more complicated loop $ae^{-1}d^{-1}c$ where the inverse signals orientation. The loop starts and ends at x. By permuting the order cyclically to $e^{-1}d^{-1}ca$, the loop now bases on y.

It is possible to construct a new theory that does not care about base point by Abelianizing the loop. The loop $ae^{-1}d^{-1}c$ is now noted by $a-e-d+c$ utilizing addition instead of multiplication to stress the Abelian quality. The order does not matter anymore. The loop is now called a cycle. To continue, $a-e-d+c$ can be seen as a chain of addition with coefficients 1 and -1. We can further generalize this object by allowing it to have arbitrary integer coefficients

$$ka + lb + mc + nd + oe \qquad k, l, m, n, o \in \mathbb{Z} \tag{5.24}$$

Denote the space in fig.5.1 as Δ, we have just defined the set of all 1-chains of Δ:

$$C_1(\Delta;\mathbb{Z}) = \{\sum_i^{a_i \in l_1} k_i a_i | k_i \in \mathbb{Z}\} \tag{5.25}$$

where l_1 is the set of all edges in Δ. The set is obviously a group under the Abelianized addition, and the edges act like basis. What is the condition for a 1-chain to be a 1-cycle? We can define a boundary operator ∂_1 on the edges:

$$\begin{aligned}
\partial a &= y - x \\
\partial b &= z - y \\
\partial c &= x - z \\
\partial d &= w - z \\
\partial e &= y - w
\end{aligned}$$

by looking at end points connected by the edges. Acting the operator on a 1-chain gets

$$\partial(ka + lb + mc + nd + oe) = (m-k)x + (k+o-l)y + (l-m-n)z + (n-o)w \tag{5.26}$$

A cycle has no boundary, thus the condition for a 1-chain $ka + lb + mc + nd + oe$ to be 1-cycle is $(m-k) = (k+o-l) = (l-m-n) = (n-o) = 0$. Check that the previous example $a+c-d-e$ indeed satisfies the condition. The boundary map maps our arbitrary 1-chain to the group of 0-chains, which is defined as

$$C_0(\Delta;\mathbb{Z}) = \{\sum_i^{x_i \in l_0} k_i x_i | k_i \in \mathbb{Z}\} \tag{5.27}$$

where l_0 is the set of end points x, y, z, w. Since there is no object in Δ with dimension higher than 1, all of the higher dimensional chain groups are trivial. Not surprisingly, all negative dimensional chain groups are also set to be trivial.

We can increase the available dimension by filling up the triangles. The surfaces is now a two-dimensional object named A and B (fig. 5.2).

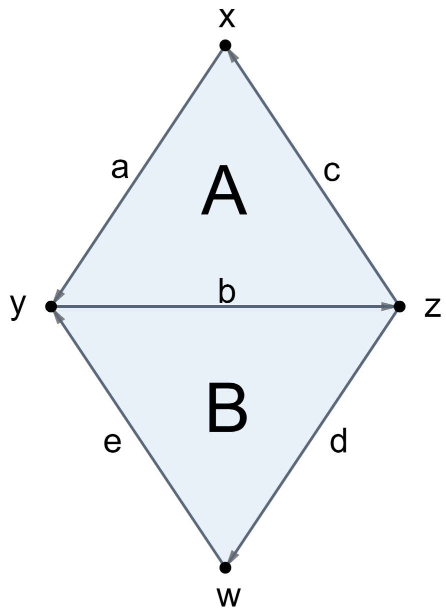

Figure 5.2: Second example of a space

The 2-chain group is

$$C_2(\Delta; \mathbb{Z}) = \{ \sum_i^{A_i \in l_2} k_i A_i | k_i \in \mathbb{Z}\} \tag{5.28}$$

A generic element of the group looks like $\alpha A + \beta B$, $\alpha, \beta \in \mathbb{Z}$. The boundary operator ∂_2 can be defined with the orientation of A, B. i.e., if both of them are facing up, we can set the boundary to obey “right-hand rule”

$$\partial A = a + b + c$$
$$\partial B = -b - d - e$$

Acting it on a 2-chain gives

$$\partial(\alpha A + \beta B) = \alpha a + (\alpha - \beta)b + \alpha c - \beta d - \beta e \tag{5.29}$$

There is no non-trivial solution for a 2-cycle, which makes sense given the shape of the graph.

Formal Definitions

The two sets of space in fig.5.1, 5.2 are called simplicial complexes in a more appropriate language. They are made of oriented r-simplexes. A formal definition:

Definition 5.2. An **Oriented r-Simplex** can be labeled by an ordered set of r+1 points $\sigma_r = (p_0, p_1, ..., p_r)$ where the points are geometrically independent. In other words, there is no (r-1)-dimensional hyperplane that contains all r+1 points. r-Simplex is the space enclosed by the r+1 vertices. It can be expressed as a subset of $\mathbb{R}^m$, $m \geq r$:

$$\sigma^r = \{x \in \mathbb{R}^m | x = \sum_{i=0}^{r} c_i p_i, c_i \geq 0, \sum_{i=0}^{r} c_i = 1\} \tag{5.30}$$

To illustrate the point, take fig.5.2 as an example. The 0-simplexes are (x),(y),(z),(w). The 1-simplexes are $a = (xy)$, $b = (yz)$, etc. The 2-simplexes are $A = (xyz)$, $B = (yzw)$. If there is a 3-simplex, it would be in the shape of tetrahedron. Within an r-simplex we can define a face:

Definition 5.3. For $q \in \mathbb{Z}$, $0 \leq q \leq r$, choose $q + 1$ points $p_{i_0}, ..., p_{i_q}$ out of the $r + 1$ vertices and they define a q-simplex $\sigma_q = (p_{i_0}, ..., p_{i_q})$. This is called a **q-Face** of σ_r

For example, edge a is a 1-face of 2-simplex A. A simplicial complex is constructed by assembling a set of simplexes. There is a manual on what to do and what not to do:

Definition 5.4. Let Δ be a set of finite number of simplexes in $\mathbb{R}^m$. Δ is a **Simplicial Complex** if
1. An arbitrary face of a simplex in Δ belongs in Δ
2. If σ, σ' are two simplexes in Δ, their intersection is either empty or a commong face.
If Δ is a simplicial complex, it can be assembled into a **Polyhedron** $|\Delta|$ in $\mathbb{R}^m$ by pasting the shared faces of the simplexes. $|\Delta|$ is a subset of $\mathbb{R}^m$ with the same dimension as Δ.

The rules prevent situation where two edges overlap in the middle or a surface missing an edge. It should be an easy exercise to check that both fig.5.1 and fig.5.2 are simplicial complexes. With all the basic notion in place we can build up chain groups and chain complex in the way we demonstrated earlier. Another piece of useful information is a formal definition of the boundary operator:

Definition 5.5. The **Boundary Operator** is a map

$$\partial_r : C_r(\Delta) \to C_{r-1}(\Delta) \tag{5.31}$$

algebraically, it is defined as

$$\partial_r \sigma_r = \sum_{i=0}^{r} (-1)^i (p_0, p_1, ..., \hat{p}_i, ..., p_r) \tag{5.32}$$

where $\hat{p}_i$ means to omit the point.

Notice that this is very similar to the coordinate free definition of exterior derivative and rightfully so. We have reviewed the process of assembling simplexes into complex and finding chains and homology of the complex. The homology can then be carried into spaces that are not complexes by the notion of triangulation:

Definition 5.6. For a topological space X. If there exists a simplicial complex Δ and a homeomorphism $f : |\Delta| \to X$, X is said to be **Triangulable** and the pair (Δ, f) is the **Triangulation** of X.

The choice of triangulation is not unique, but the Homology groups are independent of triangulations. Even better, they are topological invariant.

Theorem 5.7. *Homological groups are topological invariant. For two spaces X and Y homeomorphic to each other, take (Δ, f) and (Θ, g) to be triangulations of X, Y correspondingly,*

$$H_r(\Delta) \cong H_r(\Theta) \tag{5.33}$$

In the particular case where (Δ, f) and (Θ, g) are two triangulations of X, the property still holds. Thus it makes sense to simply denote

$$H_r(X) \equiv H_r(\Delta) \tag{5.34}$$

where (Δ, f) is an arbitrary triangulation of X.

There are spaces that are not triangulable. Singular homology is the correct tool to work with them. It is more general but less calculable than simplicial homology. We will not delve deeper into other constructions of homology such as singular homology or group homology. Interested readers can read about them in [3] and [24]. For now let's be satisfied with the promise that **all differentiable manifolds are triangulable**.

5.2 A Discrete Gauge Field

For a $\mathbb{Z}$ gauge theory on manifold M, a discrete r-form gauge fields is an r-cochain

$$\omega_r \in C^r(M; \mathbb{Z}) \tag{5.35}$$

$$\omega_r(\cdot) := (\cdot, \omega_r) \in Hom(C_r(M; \mathbb{Z}), \mathbb{Z}) : C_r(M; \mathbb{Z}) \to \mathbb{Z} \tag{5.36}$$

where $(\cdot, \cdot)$ is the inner product of the dual structure. Hopefully we have provided enough intuition to understand what $C_r(M; \mathbb{Z})$ represents. The r-cochain maps assign an element of $\mathbb{Z}$ to every r-simplex in M. We can exploit the analogy of de Rham's Theorem and think of them as r-forms with integer value:

$$\omega_r \in \Omega^r_{\mathbb{Z}}(M) \tag{5.37}$$

In this section we will explore some of the features of discrete gauge field. In order to do calculation with discrete gauge fields, we need a language that is less abstract and closer to what we are familiar with. The main sources of this section are [3], [15], [10], [9], and [24]. We will start with a brief introduction of the operations equipped by the cochain groups.

5.2.1 Coboundary, Cup Product, Integration

A coboundary operator $\delta_r : C^{r-1}(M) \to C^r(M)$ is defined as the adjoint operator to the boundary operator with respect to the inner product. For $c_r \in C_r(M)$:

$$(c_r, \delta_{r+1}\omega_r) = (\partial_r c_r, \omega_r) \tag{5.38}$$

Writing c_r as an ordered set of points $(p_0, ..., p_r)$, we can define the operator as

$$\delta_{r+1}\omega_r(p_0, ..., p_{r+1}) = \sum_{i=0}^{r}(-1)^i \omega_r(c_r(p_0, p_1, ..., \hat{p}_i, ..., p_{r+1})) \tag{5.39}$$

where $\hat{p}_i$ again means omitting the point. This is the discrete version of exterior derivative. Like exterior derivative, it is nilpotent

$$\delta_{r+1}\delta_r = 0 \tag{5.40}$$

We can define cocycle $Z^r(M)$ and coboundary $B^r(M)$ from the operator:

$$Z^r(M) = ker(\delta_{r+1}) \tag{5.41}$$

$$B^r(M) = im(\delta_r) \tag{5.42}$$

There is also a discrete version of wedge product. It is a map called cup product:

$$\cup_\eta : C^r(M; G) \times C^q(M; H) \to C^{r+q}(M; K) \tag{5.43}$$

defined upon a bi-homomorphism

$$\eta : G \times H \to K \tag{5.44}$$

For $\omega_r \in C^r(M; G)$ and $\zeta_q \in C^q(M; H)$, the cup product of them acting on a (r+q)-chain $(p_0, ..., p_{r+q})$ is defined as

$$\omega_r \cup_\eta \zeta_q(p_0, ..., p_{r+q}) = \eta(\omega_r(p_0, ..., p_r), \zeta_q(p_r, ..., p_{r+q})) \in K \tag{5.45}$$

In a special case where $G \cong H \cong \mathbb{Z}$, there is a natural bi-homomorphism $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ which is the multiplication of integers (decending from the ring structure of $\mathbb{Z}$, one might say). In this case the cup product is just

$$\omega_r \cup_\eta \zeta_q(p_0, ..., p_{r+q}) = \omega_r(p_0, ..., p_r) \cdot \zeta_q(p_r, ..., p_{r+q}) \tag{5.46}$$

where $\cdot$ is the integer multiplication. The cup product is quite special because chain groups does not have a cup product. It comes from cochains' property as maps to the coefficient and is defined by a multiplication of the coefficients. Like exterior derivative and wedge product, coboundary operator and cup product has the same graded product rule

$$\delta_{r+q+1}(\omega_r \cup \zeta_q) = \delta_{r+1}\omega_r \cup \zeta_q + (-1)^r \omega_r \cup \delta_{q+1} b_1 \tag{5.47}$$

Differential forms have a natural notion of integration. We can similarly define a discrete version of integration

$$\int_{M_r} \omega_r = \prod_{c_r \in \Delta} \omega_r(c_r) \in G \tag{5.48}$$

where (Δ, f) is a triangulation of M_r and c_r are r-simplexes in Δ. It is in this sense that ω_r can be thought of as an r-form. We can use the integration as a short-hand notation for the discrete product.

5.2.2 Local Flatness and Integral Lift

Suppose we have a discrete gauge field on a contractible manfold

$$\omega_r \in C^r(M; G) \tag{5.49}$$

where G is an unspecified discrete group. We can utilize the integration defined above to find the holonomy of ω on the boundary of a disk one dimension higher

$$\oint_{\partial D_{r+1}} \omega_r = \int_{D_{r+1}} \delta\omega_r \in G \tag{5.50}$$

Since M is contractible, we can continuously shrink ∂D_{r+1} until it is a point p_0

$$\oint_{\partial D_{r+1}} \omega_r \overset{continuous}{\longrightarrow} \int_{p_0} \omega_r \tag{5.51}$$

A point does not contain any r-chain. The integration is therefore zero

$$\int_{p_0} \omega_r = 0 \tag{5.52}$$

We have arrived at a dilemma

$$0 \overset{continuous}{\longrightarrow} \oint_{\partial D_{r+1}} \omega_r$$
$$0 \overset{discrete}{\longrightarrow} G$$

The only solution is to set the holonomy of ω_r to be constantly zero

$$\int_{D_{r+1}} \delta\omega_r = \oint_{\partial D_{r+1}} \omega_r = 0 \tag{5.53}$$

for any D_{r+1}, which equivalently states that $\delta\omega_r = 0$. Notice that this statement is not always true globally, namely that there could be non-contractible loops on topologically non-trivial manifolds. Locally, a discrete gauge field is always flat. Additionally, holonomy of the field is invariant upon adding a coboundary

$$\oint \omega_r + \delta\sigma_{r+1} = \oint \omega_r \tag{5.54}$$

Thus **locally discrete gauge fields are not just cochains but cohomology classes**

$$\omega_r \in H^r(M, G) \tag{5.55}$$

As a consequence, they are topological invariant and all of their information is stored in the holonomy, which is sometimes called "period".

The above derivation works for general discrete group, but for the remaining of the chapter we will look at a particularly useful example $G \cong \mathbb{Z}_N$. In this case the gauge field is usually normalized:

$$a_r = \frac{2\pi}{N}\omega_r \qquad a_r \in H^r(M, \frac{2\pi}{N}\mathbb{Z}_N) \tag{5.56}$$

With the specific grounp $\mathbb{Z}_N$ there is a clever way to write the discrete gauge field in the form of $U(1)$ gauge field because $\mathbb{Z}_N$ can be embedded into $U(1)$. We start with the holonomy as a defining character of the field. We have

$$\oint_\Gamma a_r \in \frac{2\pi}{N}\mathbb{Z}_N \tag{5.57}$$

We can pick an **Integra Lift** of a_r, which is a $U(1)$ gauge field $\widehat{a}_r \in \Omega^r(M)$ such that the following condition is met

$$\exp\left(i \oint_\Gamma a_r\right) = \exp\left(i \oint_\Gamma \widehat{a}_r\right) \tag{5.58}$$

This condition ensures that the integral lift can be used to replace the original discrete field in calculation. Notice that the choice of integral lift is not unique. Due to the cyclicity of exponential, integral lift is defined up to modulo:

$$\oint_\Gamma a_r = \oint_\Gamma \widehat{a}_r \; mod(2\pi) \tag{5.59}$$

What this is really doing is matching the period of

$$\oint_\Gamma a_r \in \frac{2\pi}{N}\mathbb{Z}_N = \frac{2\pi}{N}(\mathbb{Z}\; mod(N)) = \frac{2\pi}{N}\mathbb{Z}\; mod(2\pi) \tag{5.60}$$

with the period of exponential $e^{i\oint \widehat{a}}$ $(U(1))$. We will see another example when we talk about the Pontryagin Squuare. Two choices of integral lift $\widehat{a}_r$ and $\widehat{a}'_r$ are related by a gauge transformation

$$\widehat{a}_r \rightsquigarrow \widehat{a}'_r = \widehat{a}_r + \lambda_r \qquad \oint \lambda_r \in 2\pi\mathbb{Z} \tag{5.61}$$

We recognize it to be the large gauge transformation defined in 3.24. Following the same logic from 5.50 to 5.56, **locally λ_r are also cohomology classes**.

$$\lambda_r \in H^r(M, 2\pi\mathbb{Z}) \tag{5.62}$$

Conveniently, $d\widehat{a}_r$ is not zero in general. The ambiguity in choosing integral lift has saved it from flatness. This allow us to find a field strength for the integral lift

$$\widehat{f}_{r+1} = d\widehat{a}_r \in \Omega^2(M) \tag{5.63}$$

This field strength is locally independent of the choice of integral lift because the gauge shift λ is flat and therefore invisible to $\widehat{f}$. Globally this is again not necessarily true. The discrete field strength is defined to be the integral lift field strength $mod(2\pi)$

$$f_{r+1} = \widehat{f}_{r+1}\ mod(2\pi) \in H^{r+1}(M, \frac{2\pi}{N}\mathbb{Z}_N) \tag{5.64}$$

It is possible to derive f directly from the discrete gauge field a using Bockstein homomorphism. We will not introduce it here since it involves a lot of setup in Group Cohomology. Interested reader can find a comprehensive introduction in [3]. We have now finished introducing all the necessary tools. It is time to look at the physics once again.

5.3 Gauging the Center

In the grand finale, we will revisit $SU(N)$ and $PSU(N)$ Yang-Mills Theories with theta term turned on. Discrete gauge theory allows us to gauge the center $\mathbb{Z}_N$ 1-form symmetry directly and derive some of the familiar results from the last chapter. We will be using the integral lift convention for the most time. A lot of work presented in this section are first realized in [21]. This section is written with the help of [21], [1], [15], [10], [22], and [18].

5.3.1 A PSU(N) Action

We will start by looking at the center symmetry $\mathbb{Z}_N$ of $SU(N)$ gauge theory. Recall the action for $SU(N)$ Yang-Mills is

$$\mathcal{S}_{SU(N)} = \frac{1}{g^2}\int_M tr(F_2 \wedge *F_2) \tag{5.65}$$

We have omitted the theta term. It will be discussed separately in a later section. The action of the $\mathbb{Z}_N$ center symmetry shifts the dynamical gauge field by a discrete gauge field a_1. It is indeed a 1-form symmetry because it acts on the dynamical field and thus the Wilson Line. The discrete gauge field can be replaced by its integral lift $\widehat{a}_1$ which is a $U(1)$ gauge field. There remains one problem: $U(1)$ is not a subgroup of $SU(N)$ and it is not quite clear how the shifting happen. The way around it is to relax the dynamical field to be a $U(N)$ gauge field and impose the trace condition with a Lagrange multiplier in the action:

$$\mathcal{S}_{SU(N)} = \int_M \frac{1}{g^2} tr(\tilde{F}_2 \wedge *\tilde{F}_2) + \frac{i}{2\pi} Z_2 \wedge tr(\tilde{F}_2) \tag{5.66}$$

where $\tilde{F}_2 = D\tilde{A}_1$ is a $U(N)$ field strength and Z_2 is a 2-form Lagrange multiplier. Now it makes more sense to shift the $U(N)$ gauge field $\tilde{A}_1$ with a $U(1)$ integral lift $\widehat{a}_1$ because $U(1) \subset U(N)$. The shift is

$$\tilde{A}_1 \to \tilde{A}_1 + \widehat{a}_1 \mathbb{1}_N \qquad \tilde{F}_2 \to \tilde{F}_2 + \widehat{f}_2 \mathbb{1}_N \tag{5.67}$$

where $\mathbb{1}_N$ is the N-dimensional identity matrix. Locally, $\widehat{f}_2$ is invariant under different choices of integral lift. Explicitly, we denote the freedom of this choice by a curly arrow:

$$\widehat{f}_2 = d\widehat{a}_1 \rightsquigarrow d(\widehat{a}_1 + \lambda_1) = d\widehat{a}_1 \tag{5.68}$$

since the large gauge transformations are cohomology classes. An anology can be found in the electric 1-form symmetry of Maxwell theory 2.26, where A_1 is also shifted by cohomology classes. Globally on the manifold, the large gauge transformation is not flat. $\widehat{f}_2$ is allowed to transform

$$\widehat{f}_2 \rightsquigarrow \widehat{f}_2 + d\lambda_1 \qquad \tilde{F}_2 + \widehat{f}_2 \rightsquigarrow \tilde{F}_2 + \widehat{f}_2 + d\lambda_1 \mathbb{1}_N \tag{5.69}$$

This is analogous to gauging a global symmetry by making the gauge parameter non-flat, though the physical reason for the non-flatness might be different. Another interesting aspect is that the freedom of choice for integral lift acts like a "gauge freedom of a gauge freedom". Nonetheless the way to keep the symmetry is to couple the theory to a background field that transforms as

$$\widehat{b}_2 \to \widehat{b}_2 + d\lambda_1 \tag{5.70}$$

The way to couple it is similar to when we were gauging the electric 1-form symmetry. The end result is

$$\mathcal{S}_{PSU(N)} = \int_M \frac{1}{g^2} tr\left((\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N) \wedge *(\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N)\right) + \frac{i}{2\pi} Z_2 \wedge tr(\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N) \tag{5.71}$$

This is now a $PSU(N)$ action since we have successfully gauged the center symmetry. We have explained how the gauging works in an approach credited from [10], which is relatively straight-forward. There is another approach that is arguably more interesting but a bit obscure. It was used in [21], [1] and [15]. Let's take a look at it.

A Different View

The key to our second interpretation is the Lagrange multiplier term

$$\mathcal{S}_Z = \frac{i}{2\pi} \int_M Z_2 \wedge tr(\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N) \tag{5.72}$$

The trick is to recover $\tilde{F}_2$ as an $SU(N)$ gauge field F_2 and do the shift

$$A_1 \to A_1 + \widehat{a}_1 \mathbb{1}_N \qquad F_2 \to F_2 + d\widehat{a}_1 \mathbb{1}_N \tag{5.73}$$

without worrying about what the action $A_1 + \widehat{a}_1 \mathbb{1}_N$ means for now. Now the term 5.72 is reduced to

$$\begin{aligned} \mathcal{S}_Z &= \frac{i}{2\pi} \int_M Z_2 \wedge tr(F_2 + d\widehat{a}_1 \mathbb{1}_N - \widehat{b}_2 \mathbb{1}_N) \\ &= \frac{i}{2\pi} \int_M Z_2 \wedge (tr(F_2) + N d\widehat{a}_1 - N\widehat{b}_2) \\ &= \frac{iN}{2\pi} \int_M Z_2 \wedge (d\widehat{a}_1 - \widehat{b}_2) \end{aligned}$$

because the $SU(N)$ gauge field strength is traceless, the $U(1)$ fields are not traced over, and the identity matrix has trace N. This action is one format of a Topological Quantum Field Theory (TQFT) called **BF Theory**. We will not be able to give a proper introduction to BF Theory in this paper. In brief, BF Theory is a pure $\mathbb{Z}_N$ discrete gauge theory that also has a $\mathbb{Z}_N$ 1-form symmetry. For more information on BF Theory, check [10]. A specific instruction on how this format of BF Theory is derived can be found in [15]. In this picture the dynamical field is

$$\tilde{A}_1 = A_1 + \widehat{a}_1 \mathbb{1}_N \tag{5.74}$$

We have reused the notation $\tilde{A}_1$ because it is secretly a $U(N)$ field after gauging. This is realized by the relationship

$$U(N) \cong \frac{U(1) \times SU(N)}{\mathbb{Z}_N} \tag{5.75}$$

Sometimes this is written as a semi-direct product

$$U(N) \cong U(1) \rtimes SU(N) \tag{5.76}$$

This is a structure known as **Higher Group**. A good source for material related to higher symmetry would be [8]. An explanation à la Tong in [15] states that the obstruction for the $SU(N)$ gauge field A_1 to become an $PSU(N)$ gauge field cancels with the obstruction for a $\mathbb{Z}_N$ gauge field a_1 to become a $U(1)$ gauge field. We are therefore left with a well-defined $U(N)$ gauge field. The final action is thus

$$\mathcal{S}'_{PSU(N)} = \int_M \frac{1}{g^2} tr(\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N) \wedge *(\tilde{F}_2 - \widehat{b}_2 \mathbb{1}_N) + \frac{iN}{2\pi} \int_M Z_2 \wedge (d\widehat{a}_1 - \widehat{b}_2) \tag{5.77}$$

In a more general language, this is an example of stacking a QFT with a TQFT and gauging the diagonal (in this case the $\mathbb{Z}_N$) simultaneously. A higher group structure comes out naturally, and the result is a new theory with a different spectrum of operators. This is exactly the conclusion in [21].

5.3.2 The Theta Term of PSU(N) Action

We are now ready to look at the theta term. In this section, we will assume that manifold M is compactified so that the theta term is non-trivial. In the above $PSU(N)$ action, a theta term would look like

$$\begin{aligned}\mathcal{S}_\theta &= \frac{\theta}{8\pi^2}\int_M tr\left((\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N)\wedge(\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N)\right)\\ &= \frac{\theta}{8\pi^2}\int_M tr(\tilde{F}_2\wedge\tilde{F}_2) + N\widehat{b}_2\wedge\widehat{b}_2 - 2tr(\tilde{F}_2)\wedge\widehat{b}_2\\ &= \frac{\theta}{8\pi^2}\int_M tr(\tilde{F}_2\wedge\tilde{F}_2) - N\widehat{b}_2\wedge\widehat{b}_2 - 2tr(\tilde{F}_2)\wedge\widehat{b}_2 + 2tr(\widehat{b}_2\mathbb{1}_N)\wedge\widehat{b}_2\\ &= \frac{\theta}{8\pi^2}\int_M tr(\tilde{F}_2\wedge\tilde{F}_2) - N\widehat{b}_2\wedge\widehat{b}_2 - 2tr(\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N)\wedge\widehat{b}_2\end{aligned}$$

Using the trace condition

$$tr(\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N) = 0 \tag{5.78}$$

The last term vanishes and we are left with

$$\mathcal{S}_\theta = \int_M \frac{\theta}{8\pi^2} tr(\tilde{F}_2\wedge\tilde{F}_2) - \frac{\theta N}{8\pi^2}\widehat{b}_2\wedge\widehat{b}_2 \tag{5.79}$$

The first term is the winding number of the $U(N)$ theory. Following the same train of thought as in subsection 3.2.2, it can be written as an integer $n\in\mathbb{Z}$

$$\mathcal{S}_\theta = \theta n - \frac{\theta N}{8\pi^2}\int_M \widehat{b}_2\wedge\widehat{b}_2 \tag{5.80}$$

If we try to take $\theta\to\theta+2\pi$, we find that the exponential of the theta term is no longer conserved

$$\delta\mathcal{S}_\theta = -\frac{N}{4\pi}\int_M \widehat{b}_2\wedge\widehat{b}_2 + 2\pi n \tag{5.81}$$

$$e^{i\mathcal{S}_\theta}\to e^{i\mathcal{S}_\theta}\,\exp\left(-\frac{iN}{4\pi}\int_M \widehat{b}_2\wedge\widehat{b}_2\right) \tag{5.82}$$

Notice that $\frac{iN}{4\pi}\int_M \widehat{b}_2\wedge\widehat{b}_2$ is a counter term only depending on the background field. We have the freedom to add p copies of the term to the action and modifying the theta term to be

$$\mathcal{S}_\theta = \frac{\theta}{8\pi^2}\int_M tr\left((\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N)\wedge(\tilde{F}_2 - \widehat{b}_2\mathbb{1}_N)\right) + \frac{pN}{4\pi}\int_M \widehat{b}_2\wedge\widehat{b}_2 \tag{5.83}$$

Now taking $\theta\to\theta+2\pi$ is equivalent to

$$p\to p-1 \tag{5.84}$$

Pontryagin Square and Quantized Discrete Angle

To understand what the $\widehat{b}_2 \wedge \widehat{b}_2$ term means, let's unlift the field back to cohomology classes:

$$b_2 \in H^2(M; \frac{2\pi}{N}\mathbb{Z}_N) \tag{5.85}$$

It's often more convenient to undo the physical normalization as well and consider

$$\frac{N}{2\pi} b_2 = \omega_2 \in H^2(M; \mathbb{Z}_N) \tag{5.86}$$

A naive guess is that the term would look like a simple cup product $b_2 \cup b_2$. This is not the right answer. The term is in fact a **Pontryagin Square**

$$\int_M \widehat{b}_2 \wedge \widehat{b}_2 = \frac{4\pi^2}{N^2} \int_{M_4} \widehat{\omega}_2 \wedge \widehat{\omega}_2 \xrightarrow{unlift} \frac{4\pi^2}{N^2} \int_M \mathbf{P}(\omega_2) \tag{5.87}$$

where the Pontryagin Square is defined to be

$$\mathbf{P}(\omega_2 \in H^2(M; \mathbb{Z}_N)) \begin{cases} \in H^4(M; \mathbb{Z}_N) \;\; if\ N\ odd \\ \in H^4(M; \mathbb{Z}_{2N}) \;\; if\ N\ even \end{cases} \tag{5.88}$$

Heuristically, this comes from the freedom in choosing integral lift. Here is a proof for even N: Take the transformation

$$\widehat{\omega}_2 = \frac{N}{2\pi}\widehat{b}_2 \rightsquigarrow \widehat{\omega}_2 + \frac{N d\lambda_1}{2\pi} := \widehat{\omega}_2 + N d\lambda_1' \tag{5.89}$$

where we have defined $\lambda_1' = \lambda_1/2\pi$. We can find that $d\lambda'$ is

$$d\lambda_1' \in H^2(M, \mathbb{Z}) \tag{5.90}$$

Under this shift:

$$\widehat{\omega}_2 \wedge \widehat{\omega}_2 \rightsquigarrow \widehat{\omega}_2 \wedge \widehat{\omega}_2 + 2N d\lambda_1' \wedge \widehat{\omega}_2 + N^2 d\lambda_1' \wedge d\lambda_1' \tag{5.91}$$

We know that both $d\lambda_1'$ and $\widehat{\omega}_2$ have integer periods. If N is even, $\frac{N}{2} \in \mathbb{Z}$ and we can write the transformation inside a loop integral as

$$\oint \widehat{\omega}_2 \wedge \widehat{\omega}_2 \rightsquigarrow \oint \widehat{\omega}_2 \wedge \widehat{\omega}_2 + 2N(d\lambda_1' \wedge \widehat{\omega}_2 + \frac{N}{2} d\lambda_1' \wedge d\lambda_1') = \oint \widehat{\omega}_2 \wedge \widehat{\omega}_2 + 2N\mathbb{Z} \tag{5.92}$$

Thus

$$\oint \mathbf{P}(\omega_2) = \oint \widehat{\omega}_2 \wedge \widehat{\omega}_2 + 2N\mathbb{Z} \in \mathbb{Z}\ mod(2N) \cong \mathbb{Z}_{2N} \tag{5.93}$$

Additionally, in a path integral we can get rid of the $mod(2N)$ and take

$$\int_M \frac{\mathbf{P}(\omega_2)}{2} \in \mathbb{Z} \tag{5.94}$$

because the period of $\frac{\mathbf{P}(\omega_2)}{2}$ and the period of $e^{i\mathcal{S}}$ ($U(1)$) matches. Recall that this is how the integral lift works. The factor $\frac{1}{2}$ comes from the $2N$ cyclicity.

We would also want the entire theta term 5.83 to be invariant under the $\rightsquigarrow$ shift. The first term in 5.83 is manifestly invariant with the canceling between $\tilde{F}$ and $\widehat{b}$. Let's focus on the second term:

$$\begin{aligned}
\exp\left(\frac{ipN}{4\pi}\int_M \widehat{b}_2\wedge\widehat{b}_2\right) = \exp\left(\frac{ipN}{4\pi}\frac{4\pi^2}{N^2}\int_M \widehat{\omega}_2\wedge\widehat{\omega}_2\right) &= \exp\left(\frac{i2\pi p}{2N}\int_M \widehat{\omega}_2\wedge\widehat{\omega}_2\right)\\
&\rightsquigarrow \exp\left(\frac{i2\pi p}{2N}\int_M \widehat{\omega}_2\wedge\widehat{\omega}_2 + i2\pi p\mathbb{Z}\right)\\
&\overset{!}{=} \exp\left(\frac{i2\pi p}{2N}\int_M \widehat{\omega}_2\wedge\widehat{\omega}_2\right)
\end{aligned}$$

This simply limits p to be integer as well. We have prove the Pontryagin square for even N. There is also a work around for odd N that requires a spin manifold. Additionally, one can even construct Pontryagin square without using integral lift. For a more detailed exposition on Pontryagin Square, see [23].

In summary, we can write the theta term as

$$\mathcal{S}_\theta[p] = \frac{\theta}{8\pi^2}\int_M tr\left((\tilde{F}_2-\widehat{b}_2\mathbb{1}_N)\wedge(\tilde{F}_2-\widehat{b}_2\mathbb{1}_N)\right) + \frac{2\pi p}{N}\int_M \frac{\mathbf{P}(\omega_2)}{2} \tag{5.95}$$

p is an integer. The factor of N in denominator makes it periodic in N. It takes values

$$p = \{0, 1, ..., N-1\} \tag{5.96}$$

When we dial $\theta \to \theta + 2\pi$, $\mathcal{S}_\theta[p] \to \mathcal{S}_\theta[p-1]$. We have recovered the result from chapter 4. Moreover, we can connect the partition function of $SU(N)$ and $PSU(N)$ in a way similar to 3.58:

$$\mathcal{Z}_{PSU(N)} = \sum_{b_2\in H^2(M;\mathbb{Z}_N)} \mathcal{Z}_{SU(N)}[b_2,\theta] \tag{5.97}$$

$$\mathcal{Z}_{SU(N)}[b_2,\theta+2\pi n] = \mathcal{Z}_{SU(N)}[b_2,\theta]\exp\left(-\frac{2\pi n}{N}\int_M\frac{\mathbf{P}(\omega_2)}{2}\right) \tag{5.98}$$

Dialing θ adds a phase to the partition function. By dialing in different θ, we will arrive at different theories after gauging. This reproduces the Witten Effect and $SO(3)_\pm$ example. Notice that in $SU(N)$ this phase is the only possible phase (i.e. possible counter term) to add. In other gauge theories, there are possible phases that can not be reached by changing θ.

Mixed 't Hooft Anomaly

In $PSU(N)$, the theory has a new θ term. Plug 5.80 into 5.83 yields

$$\begin{aligned}\mathcal{S}_\theta &= \theta n - \frac{\theta N}{8\pi^2}\int_M \widehat{b}_2 \wedge \widehat{b}_2 + \frac{pN}{4\pi}\int_M \widehat{b}_2 \wedge \widehat{b}_2 \\ &= \theta n + \frac{(2\pi p - \theta)N}{8\pi^2}\int_M \widehat{b}_2 \wedge \widehat{b}_2 \\ &= \theta n + \frac{2\pi p - \theta}{N}\int_M \frac{\mathbf{P}(\omega_2)}{2}\end{aligned}$$

We denote the old θ as $\theta_{SU(N)}$ and find the new θ to be:

$$\theta_{PSU(N)} = -\theta_{SU(N)} + 2\pi p \tag{5.99}$$

We find that the new θ is $2\pi N$ periodic:

$$\theta_{PSU(N)} + 2\pi N = -\theta_{SU(N)} + 2\pi(p+N) \sim -\theta_{SU(N)} + 2\pi p = \theta_{PSU(N)} \tag{5.100}$$

Using the N-periodic property of p. Set $\theta_{SU(N)} = \pi$, the theta angle that preserves time reversal symmetry in $SU(N)$ theory.

$$\theta_{PSU(N)} = (2p-1)\pi \tag{5.101}$$

The action of Time reversal is

$$\theta_{PSU(N)} \xrightarrow{T} -\theta_{PSU(N)} \tag{5.102}$$

If N is odd, we find that for the specific value of p,

$$p = \frac{N+1}{2} \qquad \theta_{PSU(N)} = N\pi \sim -N\pi = -\theta_{PSU(N)} \tag{5.103}$$

At even N, $(N+1)/2 \notin \mathbb{Z}$ and is not an available value for p. Time reversal symmetry is completely broken. We can find explicitly the anomaly polynomial

$$\delta_T \mathcal{S}_{PSU(N)}[p] = \exp\left(-\frac{(2p-1)N}{4\pi^2} i \int_M \widehat{b}_2 \wedge \widehat{b}_2\right) = \exp\left(-\frac{2p-1}{N} i \int_M \mathbf{P}(\omega_2)\right) \tag{5.104}$$

Chapter 6

Conclusion and Further Works

This dissertation is an attempt to scratch the surface of gauge theory in modern literature in light of recently flourishing study of generalized symmetry. The main goal has been to approach the subject matter in a constructive, bottom-up manner so as to connect a natural flow from beginning graduate level courses in QFT to formalism used in frontier of theoretical physics.

In summary, in the first chapter we starts from the ordinary Noether's theorem and develops an abstract framework of generalized symmetry. The second chapter is based on fibre bundle picture through which we have introduced behaviour of gauge theory on non-trivial topology. We had also spent a few pages on gauging higher form symmetry with background field. Looking back retrospectively, this part would have fitted better in the first chapter. In the third chapter, line spectrum of $SU(N)$ and $PSU(N)$ becomes the central focus. We have illustrated more clearly what the lines stand for and how they were studied in terms of representation theory. In the last chapter, we turned to discrete gauge fields and relate the two gauge groups with $\mathbb{Z}_N$ 1-form symmetry.

Admittedly, much was sacrificed for variety. A lot of effort has been made during this project to incorporate and motivate three distinctive lines of mathematical theory: fibre bundle, Lie theory, and algebraic topology. They were all intensively used in modern literature of gauge theory. While they were very enjoyable to study, it was also painstaking to track the discrepancy in notations between different authors, especially between mathematicians and physicists. Care was taken during the writing process to keep a consistent, unambiguous system of notation throughout the paper.

Time has always been the biggest difficulty and limitation in this project. A few sections were relatively rushed and not polished to the level of desire. There were many important aspects of gauge theory that did not make their way into this paper. Here is a list of material I would try to include if there is more time to work with:

1. A proper introduction to BF theory. BF theory emerges naturally from the discussion of center symmetry in non-Abelian Yang-Mills as well as Abelian Higgs model. A section on BF theory would contribute to the understanding of the nature of discrete gauge field.

2. A proper introduction to 't Hooft Anomaly, Anomaly inflow and SPT phases. These

are concepts closely related to applications in physics. In particular we can discuss the low energy behavior of $SU(N)$ and $PSU(N)$ Yang-Mills.

3. A section on classifying space. Classifying space is heavily featured in [24]. According to [11], it provides a third picture for gauging of global symmetry apart from insertion of SDOs and coupling of background field, both of which were discussed in this paper.

4. Wilson loop and confinement. This is another topic with a lot of applications. Wilson loop is an important indicator of physical properties. [15] includes a very good exposition on the topic.

5. Appendixes. Originally there was a plan to include an appendix for differential form convention used in the paper. I would also like to include an appendix explaining the concept of Poincaré dual form. Unfortunately there isn't enough time to finish them.

Last of all, non-invertible symmetries remain an unexplored topic in this dissertation and in my current study. It would be great to include categorical symmetry and higher-group as a part of this paper and complete it as a set of note. Many advancements are happening in the non-invertible side of generalized symmetry, and more than half of the pages in [8] and [9] are dedicated to non-invertible symmetries. I would recommend any interested readers to check out these notes and my colleague's dissertation on 2-group structure.

Bibliography

[1] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, *Generalized Global Symmetries*, JHEP 02 (2015) 172, arXiv:1412.5148 [hep-th].

[2] M. Nakahara, *Geometry, topology and physics*, 2003.

[3] A. Hatcher, *Algebraic Topology*, 2001.

[4] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, 1972.

[5] M. Peskin and D. Shroeder, *An Introduction to Quantum Field Theory*, 1995.

[6] M. Schwartz, *Quantum Field Theory and the Standard Model*, 2014.

[7] N. S. Manton and P. Sutcliffe, *Topological Solitons*, 2004.

[8] S. Schafer-Nameki, *ICTP Lectures on (Non-)Invertible Generalized Symmetries*, JHEP 02, 172 (2015), arXiv:2305.18296 [hep-th].

[9] L. Bhardwaj, L. Bottini, L. Fraser-Taliente, L. Gladden, D. S. W. Gould, A. Platschorre, and H. Tillim, *Lectures on Generalized Symmetries*, arXiv:2307.07547 [hep-th].

[10] T. D. Brennan and S. Hong, *Introduction to Generalized Global Symmetries in QFT and Particle Physics*, arXiv:2306.00912 [hep-th].

[11] R. Luo, Q. Wang, Y. Wang, *Lecture Notes on Generalized Symmetries and Applications*, arXiv:2307.09215 [hep-th].

[12] J. A. Harvey, *Magnetic monopoles, duality, and supersymmetry*, arXiv:9603086 [hep-th].

[13] J. A. Harvey, *TASI 2003 Lectures on Anomalies*, arXiv:0509097 [hep-th].

[14] D.Tong, *TASI Lectures on Solitons*, damtp.cam.ac.uk.

[15] D.Tong, *Lectures on Gauge Theory*, damtp.cam.ac.uk.

[16] A. Lukas and H. Tillim, *Groups and Representations*, physics.ox.ac.uk

[17] C.N. Yang, and T.T. Wu, *Concept of nonintegrable phase factors and global formulation of gauge fields*, Phys. Rev. D 12 (1975) 3845.

[18] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, *Theta, Time Reversal, and Temperature*, JHEP, 91 (2017), arXiv:1703.00501 [hep-th].

[19] O. Aharony, N. Seiberg, Y. Tachikawa, *Reading between the lines of four-dimensional gauge theories*, JHEP, 115 (2013), arXiv:1305.0318 [hep-th].

[20] A. Kapustin, *Wilson-'t Hooft operators in four-dimensional gauge theories and S-duality*, Phys.Rev. D74 (2006) 025005, arXiv:0501015 [hep-th].

[21] A. Kapustin, N. Seiberg, *Coupling a QFT to a TQFT and Duality*, arXiv:1401.0740 [hep-th].

[22] P. Hsin, H. Lam, *Discrete Theta Angles, Symmetries and Anomalies*, SciPost Phys. 10, 032 (2021), arXiv:2007.05915 [hep-th].

[23] A. Kapustin, R. Thorngren, *Topological Field Theory on a Lattice, Discrete Theta-Angles and Confinement*, Adv. Theor. Math. Phys. 18 no. 5, (2014) 1233–1247, arXiv:1308.2926 [hep-th].

[24] R. Dijkgraaf and E. Witten, *Topological Gauge Theories and Group Cohomology*, Communications in Mathematical Physics volume 129, 393–429 (1990).

[25] E. Witten, *Dyons of charge $e\theta/2\pi$*, Phys. Lett. 86B (1979) 283.

[26] P. Goddard, J. Nuyts and D. I. Olive, *Gauge Theories and Magnetic Charge*, Nucl.Phys. B 125, 1 (1977).

[27] G. 't Hooft, *Magnetic monopoles in unified gauge theories*, Nucl. Phys. B79 (1974) 276.

[28] A.M. Polyakov, *Particle Spectrum in Quantum Field Theory*, JETP Lett. 20 (1974) 194.

[29] M. Kalb and P. Ramond, *Classical direct interstring action*, Phys. Rev. D 9, 2273 (1974).

[30] J. Villain, *Theory of one-dimensional and two-dimensional magnets with an easy magnetization plane. 2. The Planar, classical, two-dimensional magnet*, J. Phys. (France) 36, 581 (1975).

[31] C. Teitelboim, *Gauge Invariance for Extended Objects*, Phys. Lett. B 167, 63 (1986).

[32] Z. Komargodski, *Lectures: Advanced Quantum Field Theory*, Link to video.

[33] T. Dumitrescu, *Lectures: Generalized Symmetries in Quantum Field Theory* Link to video.