Universidade Estadual de Londrina

GABRIEL CAPELINI MAGALHÃES

TOPOLOGICAL ASPECTS AND GENERALIZED SYMMETRIES IN QUANTUM FIELD THEORY

Londrina
2026

GABRIEL CAPELINI MAGALHÃES

TOPOLOGICAL ASPECTS AND GENERALIZED SYMMETRIES IN QUANTUM FIELD THEORY

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Estadual de Londrina para obtenção do título de Mestre em Física.

Orientador: Prof. Dr. Carlos André Hernaski.

Londrina
2026

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

M188t Magalhães, Gabriel Capelini.
Topological Aspects and Generalized Symmetries in Quantum Field Theory / Gabriel Capelini Magalhães. - Londrina, 2026.
104 f. : il.

Orientador: Carlos André Hernaski.
Dissertação (Mestrado em Física) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Física, 2026.
Inclui bibliografia.

1. Teoria Quântica de Campos - Tese. 2. Simetrias Generalizadas - Tese. 3. Topologia - Tese. I. Hernaski, Carlos André. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Física. III. Título.

CDU 53

GABRIEL CAPELINI MAGALHÃES

TOPOLOGICAL ASPECTS AND GENERALIZED SYMMETRIES IN QUANTUM FIELD THEORY

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Estadual de Londrina para obtenção do título de Mestre em Física.

BANCA EXAMINADORA

Orientador: Prof. Dr. Carlos André Hernaski
Universidade Estadual de Londrina - UEL

Prof. Dr. Pedro Rogério Sergi Gomes
Universidade Estadual de Londrina - UEL

Prof. Dr. Lucas Tavares Cardoso
Universidade Federal de Santa Maria - UFSM

Londrina, 24 de fevereiro de 2026.

Este trabalho é dedicado ao meu primo Filipe Matheus Magalhães, que se foi, mas deixou como herança o seu amor por lecionar e pelas ciências exatas. R.I.P.

.

AGRADECIMENTOS

Gostaria de agradecer aos meus pais, Lázaro e Divina, por sempre terem acreditado em mim por fazer o possível e o impossível para que eu realizasse a pós-graduação. Agradeço também ao meu irmão Emmanuel por todo apoio durante esse período.

Agradeço ao meu orientador Carlos Hernaski, que esteve sempre disponível para me ajudar na realização deste trabalho, e aos professores Pedro Gomes e Paula Bienzobas por se disponibilizarem para me ajudar com dúvidas tanto sobre este trabalho quanto sobre disciplinas que cursei com os mesmos, que foram fundamentais para minha formação como cientista.

Agradeço aos meus amigos Gustavo Yoshitome, João Lucas Arguello, Giovanna Pimenta e Bruno Araujo pelas discussões científicas que foram fundamentais para entender pontos importantes neste trabalho e também pelas sessões de café que tivemos. Agradeço também aos meus amigos Luiz Felipe Demétrio, Gustavo Barreto e Orion Villanueva pelos bons momentos que passamos juntos em eventos científicos.

Aos meus demais amigos que não foram citados aqui, meu sincero agradecimento pois vocês com certeza me ajudaram a passar por esse período.

Por fim, agradeço a CAPES pelo apoio financeiro durante todo o período de realização deste trabalho e ao programa de pós-graduação em Física da Universidade Estadual de Londrina pela concessão do espaço físico.

"I'm scared to get close

I hate being alone

I long for that feeling

to not feel at all

the higher I get

the lower I'll sink

I can't drown my demons

they know how to swim."

— *Can You Feel My Heart*,

Bring Me the Horizon

RESUMO

MAGALHÃES, G. C. **Aspectos topológicos e simetrias generalizadas em Teoria Quântica de Campos.** 2026. 104 f. Dissertação (Mestrado em Física) - Centro de Ciências Exatas, Universidade Estadual de Londrina, Londrina, 2026

Neste trabalho, discutimos algumas teorias de campos topológicas, isto é, teorias de campos em que os graus de liberdade dependem exclusivamente da topologia da variedade na qual a teoria está definida. Apresentamos as teorias mais comuns que em que essa ideia emerge, sendo elas a teoria de Chern-Simons e a teoria BF. Dentro dessas teorias, estudamos seu espectro de energia e como a degenerescência depende da topologia. Por fim, introduzimos o simetrias de higher-forms e discutimos como essas simetrias se manifestam nas teorias mencionadas.

Palavras-chave: teoria de campos topológica; teorias de calibre; simetrias generalizadas.

ABSTRACT

MAGALHÃES, G. C. **Topological aspects and generalized symmetries in Quantum Field Theory.** 2026. 104 f. Master's Thesis (Master in Physical Science) - Centro de Ciências Exatas, Universidade Estadual de Londrina, Londrina, 2026.

In this work, we discuss some topological field thories, that is, field theories whose degrees of freedom depend exclusively on the topology of the manifold in which the theory is defined. We present the most common theories in which these ideas naturally emerge, namely Chern-Simons theory and BF theory. Within these theories, we study its energy spectrum and how the degeneracy depends totally on the topology, Moreover, we introduce the concept of higher-form symmetries and discuss how these symmetries manifest in the mentioned theories.

Key-words: topological field theories; gauge theories; generalized symmetries.

LISTA DE FIGURAS

SUMÁRIO

INTRODUCTION

There is a class of field theories whose defining feature is their dependence on the topology of the underlying spacetime manifold. In such theories, certain degrees of freedom may change when the topology of the manifold on which the theory is defined is modified. These theories are known as *topological field theories*. This class of field theories plays a central role in the study of *topological phase transitions*.

Topological phase transitions are phase transitions that are sensitive to changes in topology rather than to local order parameters. The first experimental realization of such a transition was observed in the *Quantum Hall Effect* [1]. Initially, these transitions did not seem to fit within the Landau paradigm, since the different phases are not characterized by the spontaneous breaking of an ordinary global symmetry. Nevertheless, more recent developments have shown that these phase transitions can be understood in terms of the spontaneous breaking of *generalized symmetries*, which extend the conventional notion of symmetry [2–7].

Generalized symmetries form a broad class that includes higher-form symmetries, whose charged operators are extended objects of higher dimension; non-invertible symmetries, which do not admit an inverse; and subsystem symmetries, which are realized only on lower-dimensional subsets of the system.

In this work, we introduce the notion of topological field theories by studying Chern–Simons theory, which provides an effective field-theoretic description of the quantum Hall effect, and BF theory, which serves as an example of a discrete gauge theory. We then explore higher-form symmetries, how they arise in these theories, and how they allow us to extract important physical information within a unified framework.

Before addressing these topics, we first develop the necessary theoretical background. In Chapter 1, we introduce gauge fields and gauge invariance, and generalize these ideas to non-Abelian $SU(N)$ symmetry groups, leading to Yang–Mills theory. In Chapter 2, we reinterpret these concepts using the language of fiber bundles. Chapter 3 is devoted to spontaneous symmetry breaking and its consequences, including the emergence of massless Goldstone bosons and the Higgs mechanism. In Chapter 4, we study topological field theories and apply the tools developed in the previous chapters to Chern–Simons theory. Finally, in Chapter 5, we introduce higher-form symmetries, analyze their realization in the theories discussed earlier, and investigate phenomena such as their spontaneous breaking and their role in the description of discrete gauge theories.

1 GAUGE THEORY

Gauge theory is a fundamental concept in the study of field theory. It relies on the simple idea of performing transformations that act independently at each point in spacetime. It is understood as a redundancy in our theory. However, it is a necessary redundancy, since all fundamental interactions in the Standard Model are described by gauge fields.

In this chapter, we begin by reviewing the Euler–Lagrange equations and Noether's theorem, as these will be used to derive several results. We then study the simplest theory that can be coupled to a gauge field: a complex scalar field possessing a $U(1)$ gauge symmetry. We will clarify the difference between a physical symmetry and a gauge symmetry and examine their geometric meaning. Finally, we generalize the idea to more general groups.

1.1 Euler-Lagrange equations and Noether theorem

In this section, we derive the Euler–Lagrange equations and Noether's theorem, as they will be needed in the following discussion. This is a standard presentation and can be found in several quantum field theory textbooks. Here, we follow the discussion in [8–10].

The Lagrangian is a well-known function from classical mechanics constructed from the symmetries of spacetime. It is related to a functional called the action, given by

$$S = \int dt \; L(\boldsymbol{x}, \dot{\boldsymbol{x}}), \tag{1.1}$$

which satisfies the principle that it must be stationary, that is,

$$\delta S = 0. \tag{1.2}$$

In the case of fields, the action, in order to maintain its functional form, is written as

$$S = \int d^4x \; \mathcal{L}(\phi, \partial_\mu \phi), \tag{1.3}$$

where

$$L = \int d^3x \; \mathcal{L}(\phi, \partial_\mu \phi), \tag{1.4}$$

is the Lagrangian and $\mathcal{L}$ is known as the *Lagrangian density*, which now depends on the field ϕ.

Varying the field,

$$\phi(x) \to \phi'(x) = \phi(x) + \delta\phi(x), \tag{1.5}$$

we obtain from (1.2),

$$\delta S = \int d^4x \, \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) \right] = 0. \tag{1.6}$$

Integrating by parts,

$$\delta S = \int d^4x \, \left\{ \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right] \delta\phi + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right] \right\} = 0. \tag{1.7}$$

The last term is what we call a *boundary term* because, by Stokes' theorem, this integral depends only on spatial and temporal infinity. We will generally assume that it vanishes[1].

Therefore, for arbitrary variations $\delta\phi$, in order to satisfy (1.7), we must have

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = 0. \tag{1.8}$$

These are the Euler–Lagrange equations for fields. Note that here we considered a scalar field, but for more general fields the structure of the equation remains the same, differing only in the presence of free indices.

For the Klein–Gordon Lagrangian,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2, \tag{1.9}$$

using (1.8) we obtain

$$(\partial_\mu \partial^\mu + m^2)\phi = 0. \tag{1.10}$$

This is the Klein–Gordon equation, which describes the dynamics of a scalar field.

Suppose now that under some transformation of the fields $\phi \to \phi + \delta\phi$, the Lagrangian changes at most by a total derivative $\partial_\mu K^\mu$, for some vector field K^μ. Hence,

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi_a} \delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta(\partial_\mu \phi_a) = \partial_\mu K^\mu. \tag{1.11}$$

The index a appears because the field may possess internal structure (this will become clear in the next section). Integrating by parts[2],

$$\partial_\mu K^\mu = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a \right) + \left[\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \right] \delta\phi_a. \tag{1.12}$$

The last term on the right-hand side is precisely the Euler–Lagrange equation. Thus, if the equations of motion are satisfied, we obtain

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a - K^\mu \right) = 0. \tag{1.13}$$

[1] There are situations in which boundary terms do not vanish and are physically relevant, but we will not discuss them here.

[2] Recall that these expressions are inside the action functional.

We conclude that there exists a *conserved current*,

$$j^\mu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a - K^\mu. \tag{1.14}$$

Integrating (1.13) over a 3-volume V,

$$\int_V d^3x \partial_\mu j^\mu = \int_V d^3x \; \partial_t j^0 + \int_V d^3x \; \partial_i j^i = 0. \tag{1.15}$$

The second term vanishes by Stokes' theorem in the limit $V \to \infty$, assuming there is no flux at infinity. We then conclude that there exists a *conserved charge*

$$\frac{d}{dt} Q = 0, \tag{1.16}$$

where

$$Q = \int_V d^3x \; j^0. \tag{1.17}$$

This is known as *Noether's theorem*, which states that for every continuous symmetry there exists an associated conserved charge. The expansion in infinitesimal variations is what restricts the theorem to continuous symmetries.

The unitary operator implementing this symmetry is defined as

$$U \equiv \exp\left(i\alpha \int_V d^3x \; j^0 \right), \tag{1.18}$$

where α is a real parameter.

1.2 Gauge Transformations

Suppose now we have a complex scalar field with two real components,

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{1.19}$$

$$\phi^* = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2). \tag{1.20}$$

For the action to remain real, we have

$$\mathcal{L} = \partial_\mu \phi \partial^\mu \phi^* - m^2 \phi^* \phi. \tag{1.21}$$

Regarding ϕ and ϕ^* as independent, then

$$(\Box + m^2)\phi = 0, \tag{1.22}$$

$$(\Box + m^2)\phi^* = 0. \tag{1.23}$$

Note that this Lagrangian is invariant under the transformations

$$\phi \to e^{-i\Lambda}\phi, \quad \phi^* \to e^{i\Lambda}\phi, \tag{1.24}$$

this is called a *global transformation* on the field ϕ as it acts the same way in every point of spacetime. Infinitesimally

$$\begin{gathered}\phi'(x) = (1 - i\Lambda)\phi \Rightarrow \phi(x') - \phi(x) = -i\Lambda\phi \\ \delta\phi = -i\Lambda\phi.\end{gathered} \tag{1.25}$$

And for ϕ^*,

$$\delta\phi^* = i\Lambda\phi^*, \tag{1.26}$$

where Λ is a real constant. Moreover, deriving both variations above, we obtain

$$\delta(\partial_\mu\phi) = -i\Lambda\partial_\mu\phi, \quad \delta(\partial_\mu\phi^*) = i\Lambda\partial_\mu\phi^*. \tag{1.27}$$

This symmetry does not add a boundary term $\partial_\mu K^\mu$. Hence the conserved current will be

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}(-i\phi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}(i\phi^*). \tag{1.28}$$

Note that now we have an internal structure, that is why we summed over the indices "a" before. Using the (1.21), we get

$$j^\mu = i(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*). \tag{1.29}$$

This current is indeed conserved

$$\begin{aligned}\partial_\mu j^\mu &= i(\partial_\mu\phi^*\partial^\mu\phi + \phi^*\Box\phi - \partial_\mu\phi\partial^\mu\phi^* - \phi\Box\phi^*) = i(\phi^*\Box\phi - \phi\Box\phi^*) \\ &= i(-\phi^* m^2\phi + \phi m^2\phi^*) = 0.\end{aligned} \tag{1.30}$$

The conserved charge then is

$$Q = \int d^3x\, j^0 = \int d^3x\, i(\phi^*\dot{\phi} - \phi\dot{\phi^*}), \tag{1.31}$$

where the dot means time derivatives. Let us get a geometric form for this transformation (1.24). Substituting (1.19) in (1.21) we have

$$\begin{aligned}\mathcal{L} &= \partial_\mu\left(\frac{\phi_1 + i\phi_2}{\sqrt{2}}\right)\partial_\mu\left(\frac{\phi_1 - i\phi_2}{\sqrt{2}}\right) - \frac{m^2}{2}(\phi_1^2 + \phi_2^2) \\ &= \left[\frac{1}{\sqrt{2}}\partial_\mu\phi_1 + \frac{i}{\sqrt{2}\partial_\mu\phi_2}\right]\left[\frac{1}{\sqrt{2}}\partial_\mu\phi_1 + \frac{-i}{\sqrt{2}\partial_\mu\phi_2}\right] - \frac{m^2}{2}(\phi_1^2 + \phi_2^2) \\ &= \frac{1}{2}\left[\partial_\mu\phi_1\partial^\mu\phi_1 + \partial_\mu\phi_2\partial^\mu\phi_2 - m^2(\phi_1^2 + \phi_2^2)\right].\end{aligned} \tag{1.32}$$

Applying the transformation

$$\phi_1' + i\phi_2' = e^{-i\Lambda}(\phi_1 + i\phi_2), \tag{1.33}$$

$$\phi_1' - i\phi_2' = e^{i\Lambda}(\phi_1 - i\phi_2). \tag{1.34}$$

Summing both

$$
\begin{aligned}
2\phi_1' &= e^{-i\Lambda}(\phi_1 + i\phi_2) + e^{i\Lambda}(\phi_1 - i\phi_2) \\
&= \phi_1 \cos\Lambda - i\phi_1 \sin\Lambda + i\phi_2 \cos\Lambda + \phi_2 \sin\Lambda + \phi_1 \cos\Lambda + i\phi_1 \sin\Lambda - i\phi_2 \cos\Lambda \\
&+ \phi_2 \sin\Lambda,
\end{aligned} \tag{1.35}
$$

which leads to

$$
\phi_1' = \phi_1 \cos\Lambda + \phi_2 \sin\Lambda. \tag{1.36}
$$

Doing the same for ϕ_2', we obtain

$$
\phi_2' = -\phi_1 \sin\Lambda + \phi_2 \cos\Lambda. \tag{1.37}
$$

Therefore, this transformation is a rotation in the internal space of the field, as shown in Fig.(1.1). Rotations in the plane are usually described by the SO(2) group, but since there is an isomorphism between SO(2) and U(1), we can use the latter as well.

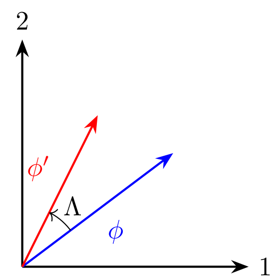

Figura 1.1 – Global transformation as a rotation in the 1-2 plane. Adapted from [8].

So far, we have been working with what we refer to as a global transformation. This is because the parameter Λ is a constant, meaning we apply this transformation uniformly across all points in the field *simultaneously*. But what happens if we make this parameter local? That is, what if $\Lambda = \Lambda(x)$. What we have now what we call a *gauge transformation*. For the field, it remains the same as before

$$
\delta\phi = -i\Lambda(x)\phi. \tag{1.38}
$$

But the derivative changes

$$
\partial_\mu \phi' = \partial_\mu \phi - i\Lambda(x)(\partial_\mu \phi) - i(\partial_\mu \Lambda(x))\phi, \tag{1.39}
$$

and consequently its variation,

$$
\delta(\partial_\mu \phi) = -i\Lambda(\partial_\mu \phi) - i(\partial_\mu \Lambda)\phi. \tag{1.40}
$$

Similarly,

$$
\delta\phi^* = i\Lambda\phi^*, \tag{1.41}
$$

$$
\delta(\partial_\mu \phi^*) = i\Lambda(\partial_\mu \phi^*) + i(\partial_\mu \Lambda)\phi^*. \tag{1.42}
$$

Therefore, the derivative no longer transforms covariantly, i.e., in the same way as ϕ. This implies directly a change in the Lagrangian, as can be seen when computing

$$\begin{aligned}
\delta\mathcal{L} &= \delta(\partial_\mu\partial^\mu\phi^*) - m^2\delta(\phi^*\phi) = \partial_\mu\phi\delta(\partial_\mu\phi\delta(\partial^\mu\phi^*) + \delta(\partial_\mu\phi)\partial^\mu\phi^* \\
&- m^2(\delta\phi^*\phi + \phi^*\delta\phi) \\
&= \partial_\mu\phi(i\Lambda(\partial^\mu\phi^*) + i(\partial^\mu\Lambda)\phi^*) + (-i\Lambda(\partial_\mu\phi) - i(\partial_\mu\Lambda)\phi)\partial^\mu\phi^* \\
&- m^2(i\Lambda\phi^*\phi - i\Lambda\phi^*\phi) \\
&= (\partial_\mu\Lambda)(i\phi^*\partial^\mu\phi - i\phi\partial^\mu\phi^*) = (\partial_\mu\Lambda)j^\mu.
\end{aligned} \tag{1.43}$$

Thus, the Lagrangian is not invariant under gauge transformations! This is problematic, because, taking Maxwell theory as an example, it is known that the photon has only two degrees of freedom, but it is described by a 4-vector A_μ. The way to get out of this problem is to use the gauge freedom of the theory to eliminate two degrees of freedom from A_μ. Therefore we want that our Lagrangians to be gauge invariant.

In order to be gauge invariant, we state that the gauge field A_μ couples directly with the conserved current, that is,

$$\mathcal{L}_1 = -iej^\mu A_\mu, \tag{1.44}$$

where e is a coupling constant that adjusts the units such that eA_μ have dimension of ∂_μ. We demand that A_μ transforms under a gauge transformation as

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\Lambda. \tag{1.45}$$

Hence

$$\delta\mathcal{L}_1 = ie(\delta j^\mu A_\mu + j^\mu\delta A_\mu) = -e\left(\delta j^\mu A_\mu + \frac{1}{e}j^\mu\partial_\mu\Lambda\right), \tag{1.46}$$

with

$$\begin{aligned}
\delta j^\mu &= i(\delta\phi^*)\partial^\mu\phi + i\phi^*\delta(\partial^\mu\phi) - i(\delta\phi)\partial^\mu\phi^* - i\phi\delta(\partial^\mu\phi^*) = -\Lambda\phi^*(\partial^\mu)\phi + \phi^*\Lambda(\partial^\mu\phi) \\
&+ \phi(\partial^\mu\Lambda)\phi^* - \Lambda\phi(\partial^\mu\phi^*) - \phi\Lambda(\partial^\mu\phi^*) + \phi(\partial^\mu\Lambda\phi^*) = 2\phi^*\phi(\partial^\mu\phi),
\end{aligned} \tag{1.47}$$

so that

$$\begin{aligned}
\delta\mathcal{L} + \delta\mathcal{L}_1 &= i\partial_\mu\Lambda(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) - 2e\phi^*\phi(\partial^\mu\Lambda)A_\mu - j^\mu(\partial_\mu\Lambda) \\
&= -2eA_\mu\phi^*\phi(\partial^\mu\Lambda).
\end{aligned} \tag{1.48}$$

We add another term to compensate for this latter,

$$\mathcal{L}_2 = e^2 A_\mu A^\mu\phi^*\phi, \tag{1.49}$$

whose variation is,

$$\begin{aligned}
\delta\mathcal{L}_2 &= e^2\left[2A_\mu(\delta A^\mu)\phi^*\phi + A_\mu A^\mu(\phi^*(\delta\phi) + (\delta\phi^*)\phi)\right] \\
&= e^2\left[2A_\mu\phi^*\phi\left(\frac{1}{e}\partial^\mu\Lambda\right) + A_\mu A^\mu\phi^*(-i\Lambda\phi) + A_\mu A^\mu(i\Lambda\phi^*)\phi\right] \\
&= 2eA_\mu\phi^*\phi(\partial^\mu\Lambda).
\end{aligned} \tag{1.50}$$

Now, we finally have

$$\delta\mathcal{L} + \delta\mathcal{L}_1 + \delta\mathcal{L}_2 = 0. \tag{1.51}$$

Therefore, upon adding the field A_μ, the Lagrangian becomes gauge invariant. But now this field dynamics must be taken into account, so we add a gauge invariant term that contains derivatives of this field. The simplest term is

$$\mathcal{L}_3 = -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.52}$$

The total Lagrangian then becomes

$$\mathcal{L} = (\partial_\mu \phi + ieA_\mu \phi)(\partial^\mu \phi^* - ieA^\mu \phi^*) - m^2\phi^*\phi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.53}$$

If we now compare (1.53) with (1.21) we note that the derivative is replaced by the additional term containing the field A_μ. We define then the *covariant derivative* as

$$\mathcal{D}_\mu \equiv \partial_\mu + ieA_\mu. \tag{1.54}$$

This name is given because now it transforms covariantly under gauge transformations

$$\begin{aligned}\delta(\mathcal{D}_\mu\phi) &= \delta(\partial_\mu\phi) + ie(\delta A_\mu)\phi + ieA_\mu(\delta\phi) = -i\Lambda(\partial_\mu\phi) - i(\partial_\mu)\Lambda\phi + i(\partial_\mu\Lambda)\phi + e\Lambda\phi A_\mu \\ &= -i\Lambda(\partial_\mu\phi + ieA_\mu\phi) = -i\Lambda\mathcal{D}_\mu\phi.\end{aligned} \tag{1.55}$$

Now let us understand the meaning of the coupling constant e. For Maxwell theory, we have that

$$\partial_\mu = (\partial_t, \boldsymbol{\nabla}), \quad A_\mu = (\phi, -\boldsymbol{A}). \tag{1.56}$$

The spatial part of the derivative transforms as

$$\boldsymbol{\nabla} \to \boldsymbol{\nabla} - ie\boldsymbol{A}. \tag{1.57}$$

Quantum mechanically, on the coordinate basis, we have

$$\hat{\boldsymbol{p}} = -i\boldsymbol{\nabla}, \tag{1.58}$$

so that

$$\hat{\boldsymbol{p}} \to \hat{\boldsymbol{p}} - e\boldsymbol{A}. \tag{1.59}$$

This is precisely the momentum that appears in the kinetic term of the Hamiltonian of a particle in an electromagnetic field. Therefore, e can be recognized as the electric charge of the field ϕ. If we now identify the covariant derivative for ϕ^* as

$$\mathcal{D}_\mu\phi^* = (\partial_\mu - ieA_\mu)\phi^*, \tag{1.60}$$

we then conclude that ϕ^* describes a field with charge $-e$.

Indeed, since the dynamics of A_μ have to be taken into account, varying the Lagrangian with respect to it leads to Maxwell equations

$$\begin{aligned}\partial_\nu F^{\mu\nu} &= -ie(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) + 2e^2A^\mu|\phi|^2 \\ &= -ie(\phi^*\mathcal{D}^\mu\phi - \phi\mathcal{D}^\mu\phi*) = -e\mathcal{J}^\mu,\end{aligned} \tag{1.61}$$

where

$$\mathcal{J}^\mu = i(\phi^*\mathcal{D}^\mu\phi - \phi\mathcal{D}^\mu\phi^*), \tag{1.62}$$

is the *covariant current.* Note that due to the anti-symmetric nature of $F_{\mu\nu}$,

$$\partial_\mu\mathcal{J}^\mu = 0. \tag{1.63}$$

There are two interesting remarks to be made. First, some books such as [8] call the global transformation a *gauge transformation of the first kind* and the local a *gauge transformation of the second kind.* We no longer use those terms because these tranformations have different natures. The global transformation is actually a physical symmetry of the theory, in the sense that it has a conserved charge associated to it and takes one state into another. The gauge transformation is not a physical symmetry but rather a redundancy in the theory. It is like a “coordinate change”.

Second, the coupling between the conserved current and the gauge field is no coincidence and is quite general. When we take $\Lambda = \Lambda(x)$, the variation of the Lagrangian must involve the extra terms that will appear and will be given by

$$\delta\mathcal{L}_0 = (\partial_\mu\Lambda)j^\mu + \mathcal{O}(\Lambda^2), \tag{1.64}$$

where $\mathcal{L}_0$ is the variation which do not include any derivatives of Λ and j^μ are just some coefficients. Thus if the equations of motion are satisfied, we must have

$$\int d^4x(\partial_\mu\Lambda^\mu) = 0. \tag{1.65}$$

Integrating by parts we get

$$\partial_\mu j^\mu = 0. \tag{1.66}$$

Thus we obtain a conserved current j^μ. This is just another way to derive Noether's theorem. If we demmand that the Lagrangian is invariant without using the equations of motion, it is necessary to add a field A_μ that transforms as $\delta A_\mu = \partial_\mu\Lambda$, so that

$$\delta\mathcal{L} = \delta\mathcal{L}_0 - \delta A_\mu j^\mu = (\partial_\mu\Lambda)j^\mu - (\partial_\mu\Lambda)j^\mu = 0. \tag{1.67}$$

The extra terms that appear in (1.64) will depend specifically on the Lagrangian that we are working but will always vanish.

1.3 The Yang-Mills Theory

The generalization for the gauge invariance of the electrodynamics was first proposed by Chen Ning Yang and Robert Mills [11], to provide an explanation to the isospin conservation. It turns out that this theory called *Yang-Mills theory* would be generalized to other groups of symmetry $SU(N)$ and would be essential in the development of particle physics as we know today.

1.3.1 The $SU(2)$ symmetry group

The case that we have treated above is built up on the $U(1)$ symmetry group. An obvious generalization would be to consider the $SU(2)$ symmetry group.

Consider the rotations (1.36) e (1.37). If now we extend this to 3 dimensions, obviously, it will be a rotation about an axis that we might call 3 (refer to Fig.(1.2)). The field ϕ should have 3 components as well, i.e., $\boldsymbol{\phi} = \boldsymbol{\phi}(\phi_1, \phi_2, \phi_3)$. An infinitesimal rotation by an angle Λ in this 3-dimensional plane in the internal space is given by

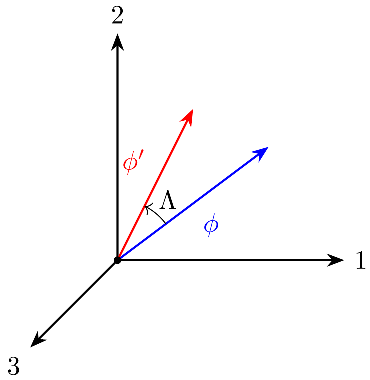

Figura 1.2 – Rotation by an angle Λ_3. Adapted from [8]

$$\delta\boldsymbol{\phi} = -\boldsymbol{\Lambda} \times \boldsymbol{\phi}. \tag{1.68}$$

Now we follow the same recipe as we did for the scalar field. Make $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}(x)$, then

$$\delta\boldsymbol{\phi} = -\boldsymbol{\Lambda}(x) \times \boldsymbol{\phi}, \tag{1.69}$$

and also,

$$\partial_\mu \boldsymbol{\phi'} = \partial_\mu \boldsymbol{\phi} - \partial_\mu \boldsymbol{\Lambda} \times \boldsymbol{\phi} - \boldsymbol{\Lambda} \times \partial_\mu \boldsymbol{\phi}. \tag{1.70}$$

It again does not transform covariantly. To correct this, we introduce a *covariant derivative* given by

$$\boldsymbol{\mathcal{D}}_\mu \boldsymbol{\phi'} = \partial_\mu \boldsymbol{\phi'} + g \boldsymbol{W}_\mu \times \boldsymbol{\phi}, \tag{1.71}$$

where this form is justified because we want something similar to Eq.(1.54). Note that here the field $\boldsymbol{W}_\mu$ plays the role of the field A_μ in the $U(1)$ case, but with the difference that it is a vector in the internal space, in contrast with A_μ that was a scalar. Also, g plays the role of e. From now on, we drop this vector notation since it is explicit that those fields are all vectors in the internal space unless is said the contrary.

Since $\mathcal{D}_\mu$ is a vector, it should transform as

$$\delta(\mathcal{D}_\mu\phi) = -\Lambda(x) \times (\mathcal{D}_\mu\phi). \tag{1.72}$$

To fulfill this requirement, the change in the gauge field W_μ should be similar to that of A_μ, then it is reasonable to write

$$\delta W_\mu = -\Lambda \times W_\mu + \frac{1}{g}\partial_\mu\Lambda. \tag{1.73}$$

Then

$$\begin{aligned} \delta(\mathcal{D}_\mu\phi) &= \delta(\partial_\mu\phi) + g(\delta W_\mu \times \phi + W_\mu \times \delta\phi) \\ &= -\Lambda \times \partial_\mu\phi - \partial_\mu\Lambda \times \phi + g\left((-\Lambda \times W_\mu) + \frac{1}{g}\partial_\mu\Lambda\right) \times \phi - gW_\mu \times (\Lambda \times \phi). \end{aligned} \tag{1.74}$$

Using the vector identity

$$(A\times B)\times C+(B\times C)\times A+(C\times A)\times B = 0 \Rightarrow (A\times B)\times C+B\times(A\times C) = A\times(B\times C), \tag{1.75}$$

we identify

$$\begin{aligned} \delta(\mathcal{D}_\mu\phi) &= -\Lambda \times \partial_\mu\phi - g((\Lambda \times W_\mu) \times \phi + W_\mu \times (\Lambda \times \phi)) \\ &= -\Lambda \times \partial_\mu\phi - g\Lambda \times (W_\mu \times \phi) = -\Lambda \times (\partial_\mu\phi - gW_\mu \times \phi) \\ &= -\Lambda \times \mathcal{D}_\mu\phi, \end{aligned} \tag{1.76}$$

as required.

Now we define a field analog to $F_{\mu\nu}$, we will call it $W_{\mu\nu}$. It is a vector in the internal space, the same way as W_μ, so it should transforms as

$$\delta W_{\mu\nu} = -\Lambda \times W_{\mu\nu}. \tag{1.77}$$

But if we define it analogously to $F_{\mu\nu}$, i.e.,

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu, \tag{1.78}$$

it will not work because

$$\begin{aligned} \delta(W_{\mu\nu}) &= \partial_\mu\left(-\Lambda \times W_\nu + \frac{1}{g}\partial_\nu\Lambda\right) - \partial_\nu\left(-\Lambda \times W_\mu + \frac{1}{g}\partial_\mu\Lambda\right) \\ &= \Lambda \times (\partial_\nu W_\mu - \partial_\mu W_\nu) - (\partial_\mu\Lambda \times W_\nu - \partial_\nu\Lambda \times W_\mu). \end{aligned} \tag{1.79}$$

The last term is unwanted. But noting that

$$\begin{aligned} \delta(gW_\mu \times W_\nu) &= g[(\delta W_\mu) \times W_\nu + W_\mu \times (\delta W_\nu)] \\ &= g\left[(-\Lambda \times W_\mu) \times W_\nu + \frac{1}{g}\partial_\mu\Lambda \times W_\nu - W_\mu \times (\Lambda \times W_\nu) + \frac{1}{g}W_\mu \times \partial_\nu\Lambda\right]. \end{aligned} \tag{1.80}$$

Using again Eq.(1.75),

$$
\begin{aligned}
\delta(gW_\mu \times W_\nu) &= -g\Lambda \times (W_\mu \times W_\nu) + \partial_\mu\Lambda \times W_\nu + W_\mu \times \partial_\nu\Lambda \\
&= -g\Lambda \times (W_\mu \times W_\nu) + \partial_\mu\Lambda \times W_\nu - \partial_\nu\Lambda \times W_\mu .
\end{aligned}
\tag{1.81}
$$

The last term is equal to the negative of the last term in Eq.(1.79). Hence, if we define

$$
W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + gW_\mu \times W_\nu, \tag{1.82}
$$

it cancels out perfectly the unwanted term and then

$$
\delta(W_{\mu\nu}) = -\Lambda \times W_{\mu\nu}. \tag{1.83}
$$

We note that the difference in the definition for $W_{\mu\nu}$ is due to the non-abelian nature of the group. The Lagrangian then will be given by

$$
\mathcal{L} = \frac{1}{2}(\mathcal{D}_\mu\phi)(\mathcal{D}^\mu\phi) - \frac{m^2}{2}\phi\cdot\phi - \frac{1}{4}W_{\mu\nu}\cdot W^{\mu\nu}. \tag{1.84}
$$

Using the Euler-Lagrange equations

$$
\frac{\partial\mathcal{L}}{\partial W_\mu} = \partial_\nu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\nu W_\mu)}\right], \tag{1.85}
$$

we can obtain the equations of motion

$$
g[(\partial_\mu\phi)\times\phi - g(W_\mu\times\phi)\times\phi] = \partial^\nu W_{\mu\nu} + gW^\nu \times W_{\mu\nu}. \tag{1.86}
$$

Or, in terms of the covariant derivative

$$
\mathcal{D}^\nu W_{\mu\nu} = g(\mathcal{D}_\mu\phi)\times\phi \equiv gj_\mu, \tag{1.87}
$$

where j_μ is also a vector on the internal space. This is analogous to the inhomogeneous Maxwell Equations (1.61). But note that there exists an interesting difference about this equation.

In the Maxwell equations, in the absence of matter j_μ, the equations of motion becomes

$$
\partial^\mu F_{\mu\nu} = 0, \tag{1.88}
$$

wich means that there is no source for the electromagnetic field. Whereas in this case, in the absence of matter

$$
\mathcal{D}^\nu W_{\mu\nu} = 0 \Rightarrow \partial^\nu W_{\mu\nu} = -gW_\mu \times W_\nu, \tag{1.89}
$$

which shows that a non-abelian gauge field is a source for itself.

Moreover, there is another interesting aspect. The Bianchi Identity

$$
\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0, \tag{1.90}
$$

leads us to the homogeneous Maxwell equations,

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \quad \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times \boldsymbol{B} = 0. \tag{1.91}$$

The analogous for the field $W_{\mu\nu}$ would be

$$\mathcal{D}_\lambda W_{\mu\nu} + \mathcal{D}_\mu W_{\nu\lambda} + \mathcal{D}_\nu W_{\lambda\mu} = 0. \tag{1.92}$$

But since the above is not written in terms of ordinary derivatives but in terms of covariant derivatives, then we have that

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} \neq 0, \tag{1.93}$$

in other words, non-abelian groups offer the possibility of magnetic monopoles.

1.3.2 Geometric idea of gauge transformations

We now present the geometric interpretation of what is happening in the internal space of the fields.

Recall the $SU(2)$ case. An infinitesimal rotation in the internal space is given by

$$\phi' = \phi - \Lambda \times \phi. \tag{1.94}$$

Since it is infinitesimal, Λ is a small parameter and can be thought of as $\Lambda \to \frac{\Lambda}{N}$, where $N \gg \Lambda$. To obtain a finite rotation, we must perform N successive rotations. Hence, writing in index notation,

$$\phi'_i = \left(\delta_{li} - \epsilon_{ijk}\frac{\Lambda_j}{N}\delta_{lk}\right)^N \phi_l. \tag{1.95}$$

In the limit $N \to \infty$, we obtain

$$\phi'_i = \exp(-\epsilon_{ijk}\Lambda_j\delta_{lk})\phi_l. \tag{1.96}$$

Defining the components of the *generators of the group* $(G_i)_{jk}$ as

$$(G_i)_{jk} = i\epsilon_{ijk}, \tag{1.97}$$

where this notation denotes the jk element of the matrix G_i, we obtain

$$\phi'_i = \exp(i(G_i)_{jl}\Lambda_j)\phi_l. \tag{1.98}$$

Returning to vector notation,

$$\phi' = \exp(-iG \cdot \Lambda)\phi. \tag{1.99}$$

The generators G are explicitly given by

$$G_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.100}$$

The objects G are, in fact, rather abstract. What we are doing when writing the matrices G is constructing what is called a *representation of the group* $SU(2)$. The generators obey the algebra

$$[G_i, G_j] = i\epsilon_{ijk}G_k = f_{ijk}G_k, \tag{1.101}$$

where f_{ijk} are called the structure constants of the group. The generators also satisfy the Jacobi identity,

$$[[M_i, M_j], M_k] + [[M_j, M_k], M_i] + [[M_k, M_i], M_j] = 0, \tag{1.102}$$

which implies, for the structure constants,

$$f_{ijm}f_{mkl} + f_{jkm}f_{mil} + f_{kim}f_{mjl} = 0. \tag{1.103}$$

The components of the matrices (1.100) can be written as

$$(G_i)_{jk} = f_{ijk}, \tag{1.104}$$

which is known as the *adjoint representation*.

Following this idea, the generalization to an N-component vector is straightforward and is given by

$$\psi' = \exp(iG^a\Lambda^a(x))\psi(x) = S(x)\psi(x), \tag{1.105}$$

where we now assume that Λ may depend on position and $a = 1, 2, 3$.

We are now able to understand the geometrical meaning of the objects that arise in Yang–Mills theory. Taking the derivative of Eq. (1.105),

$$\partial_\mu\psi' = S(\partial_\mu\psi) + (\partial_\mu S)\psi. \tag{1.106}$$

As we have seen before, this expression fails to transform covariantly. This occurs because we are performing different rotations of the vectors in the internal space at different spacetime points. Therefore, when we take the derivative, we are comparing vectors that belong to different coordinate systems (see Fig. (1.3)).

To make $\partial_\mu\psi'$ covariant, we must compare $\psi(x)$ with what it would be if transported to a point $x + dx$ while keeping the axes fixed. This procedure is known as *parallel transport*, as illustrated in Fig. (1.4). The resulting vector will be of the form $\psi(x) + \delta\psi$.

It is reasonable to assume that the change in the vector, $\delta\psi$, is proportional to the displacement dx^μ and to the generators G^a, which specify the representation in which we are working. Moreover, there must be an object responsible for transporting the vector ψ; let us denote it by A_μ. Then the change in ψ can be written as

$$\delta\psi = igG^aA^a_\mu dx^\mu\psi. \tag{1.107}$$

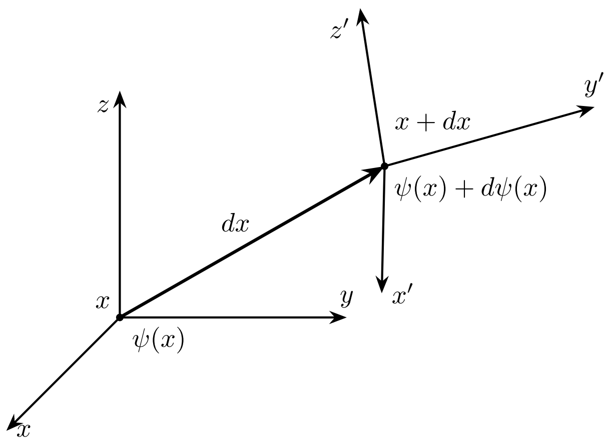

Figura 1.3 – Comparation of two vectors in the internal space. Adapted from [8]

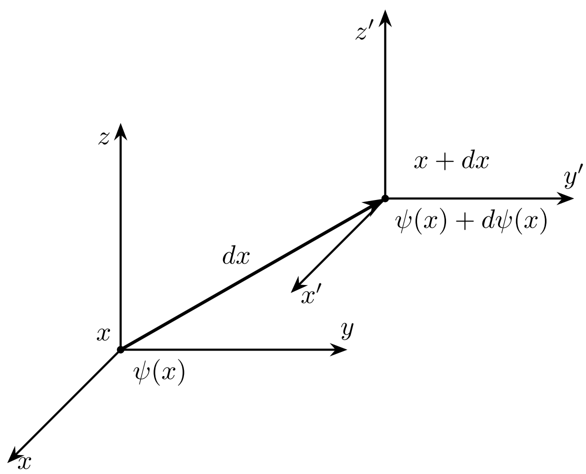

Figura 1.4 – The correct form to compare two vectors in the internal space. Adapted from [8].

We now have two fields: $\psi(x)+d\psi$ and $\psi(x)+\delta\psi$. The covariant derivative evaluates the difference between these two quantities. Hence,

$$D\psi = \psi + d\psi - \psi - \delta\psi = d\psi - igG^a A^a_\mu dx^\mu \psi, \tag{1.108}$$

and dividing by dx^μ,

$$\frac{D\psi}{dx^\mu} \equiv \mathcal{D}_\mu \psi = \partial_\mu \psi - igG^a A^a_\mu \psi, \tag{1.109}$$

which is the covariant derivative for an arbitrary group and field.

Let us verify that this general expression reproduces the previous cases. For the $U(1)$ case, taking $G = -1$ and $g = e$,

$$\mathcal{D}_\mu = \partial_\mu - ieA_\mu. \tag{1.110}$$

The choice $G = -1$ is consistent because the group elements are simply ordinary numbers. For $SU(2)$,

$$(G_a)_{mn} = -i\epsilon_{amn}, \tag{1.111}$$

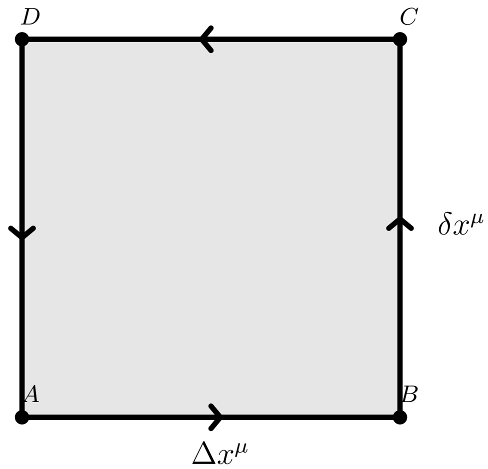

Figura 1.5 – Closed loop to parallel transport the vector ψ. Adapted from [8].

which yields

$$\mathcal{D}_\mu \phi_m = \partial_\mu \phi_m + g\epsilon_{amn} A^a_\mu \phi_n = \partial_\mu \phi + g A_\mu \times \phi, \tag{1.112}$$

in agreement with Eq. (1.71).

It is convenient to write

$$A_\mu = G^a A^a_\mu, \tag{1.113}$$

which makes clear that A_μ is a linear combination of the generators of the group. We say that it is an *algebra-valued vector*. This vector field is called a *gauge field* or a *Yang–Mills gauge potential*.

Since $\mathcal{D}_\mu$ transforms covariantly,

$$\mathcal{D}'_\mu \psi' = S(\mathcal{D}_\mu \psi). \tag{1.114}$$

Proceeding as before, we obtain

$$A'_\mu = S A_\mu S^{-1} - \frac{i}{g}(\partial_\mu S) S^{-1}. \tag{1.115}$$

This is the gauge transformation of the gauge field. Note that the second term makes A_μ transform inhomogeneously. This raises the question of whether the gauge potential can be transformed to zero at every point. If that were possible, the gauge field would have no physical meaning. We can check this by parallel transporting the vector A_μ around a closed loop, like Fig.(1.5). If the final vector is equal to the initial vector, then we successfully found a coordinate system in which the gauge potential goes to zero.

Let's call the initial vector at the point A, $\psi_{A,0}$. So, for the path $A \to B$

$$\psi_B = \psi_{A,0}(x^\mu + \Delta x^\mu) = \psi_{A,0}(x^\mu + \mathcal{D}_\mu \psi_{A,0}(x^\mu)\Delta x^\mu + \frac{1}{2}\mathcal{D}_\mu \mathcal{D}_\nu \psi_{A,0}(x^\mu)\Delta x^\mu \Delta x^\nu. \tag{1.116}$$

Now, for the path $B \to C$

$$\begin{aligned} \psi_C = \psi_B(\Delta x^\mu + \delta x^\mu) &= \psi_B + \mathcal{D}_\mu \psi_B \delta x \mu + \frac{1}{2}\mathcal{D}_\mu \mathcal{D}_\nu \psi_B \delta x^\mu \delta x^\nu \\ &= \left(1 + \delta x^\mu \mathcal{D}_\mu + \frac{1}{2}\delta x^\mu \delta x^\nu \mathcal{D}_\mu \mathcal{D}_\nu\right)\left[\psi_{A,0} + \mathcal{D}_\mu \psi_{A,0}\Delta x^\mu + \frac{1}{2}\mathcal{D}_\mu \mathcal{D}_\nu \psi_{A,0}\Delta x^\mu \Delta x^\nu\right] \\ &= \left(1 + (\delta x^\mu + \Delta x^\mu)\mathcal{D}_\mu + \frac{1}{2}(\delta x\ \mu \delta x^\nu + 2\delta x^\mu \Delta x^\nu + \delta x^\mu \delta x^\nu)\mathcal{D}_\mu \mathcal{D}_\nu\right)\psi_{A,0}. \end{aligned} \tag{1.117}$$

Then the path $C \to D$,

$$\psi_D = \left(1 + \delta x^\mu \mathcal{D}_\mu + \frac{1}{2}\delta x^\mu \delta x^\nu \mathcal{D}_\mu \mathcal{D}_\nu + (\delta x^\mu \Delta x^\nu - \Delta x^\mu \delta x^\nu)\mathcal{D}_\mu \mathcal{D}_\mu \mathcal{D}_\nu\right)\psi_{A,0}. \tag{1.118}$$

Now back to A, we call this vector $\psi_{A,1}$ and

$$\psi_{A,1} = (1 + \delta x^\mu \Delta x^\nu [\mathcal{D}_\mu, \mathcal{D}_\nu])\psi_{A,0}, \tag{1.119}$$

where $[\mathcal{D}_\mu, \mathcal{D}_\nu]$ is the commutator of the covariant derivatives. Explicitly

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = [\partial_\mu - igA_\mu, \partial_\nu - igA_\nu] = -ig(\partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]). \tag{1.120}$$

Defining the *Yang-Mills Field Strenght* $F_{\mu\nu}$ as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \tag{1.121}$$

we have

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = -igF_{\mu\nu}. \tag{1.122}$$

We understand the product $\delta x^\mu \Delta x^\nu$ as the area $S^{\mu\nu}$ of the rectangle. Then

$$\begin{aligned} \psi_{A,1} &= (1 - igS^{\mu\nu}F_{\mu\nu})\psi_{A,0}, \\ \Delta\psi &= -igS^{\mu\nu}F_{\mu\nu}\psi_{A,0} \end{aligned} \tag{1.123}$$

That shows that the vector $\psi_{A,0}$ changes when transported around a closed loop. Hence, the effect of the gauge field is to rotate the vector around a closed loop in the internal space. Using the definition of the field strength and the transformation for the gauge field, it can be shown that $F_{\mu\nu}$ transforms covariantly

$$F'_{\mu\nu} = SF_{\mu\nu}S^{-1}, \tag{1.124}$$

then, if the field strength is non-zero in one coordinate system, it is nonzero in any other coordinate system.

Similar to what we did for the gauge field, let us show that this definition for the field strength reproduces what we had before.

For the $U(1)$ case, since the generators are just ordinary numbers (recall that $G = -1$), what rests is the commutator of two real-valued fields, which commute[3]. Then

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.125}$$

[3] Another way to think is recall that the $U(1)$ group is abelian, so its elements commute.

For the $SU(2)$ case, we have

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - ig[G^b A^b_\mu, G^c A^c_\nu], \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu - igA^b_\mu A^c_\nu [G^b, G^c], \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + gA^b_\mu A^c_\nu \epsilon^{abc} G^a, \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + g(A_\mu \times A_\nu)^a G^a, \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + gA_\mu \times A_\nu, \end{aligned} \tag{1.126}$$

which is what we indeed derived before.

To close this section, let us compare these results with other theories that have a lot of similarities: General Relativity.

While in gauge theory we've been discussing gauge transformations, General Relativity concerns itself with something different: coordinate transformations. In General Relativity, we work with curved spacetime, which changes how we compare two vectors. We need an object called a *connection* to compare vectors in different vector spaces.

This concept parallels what we discussed earlier, where our vectors lie not in spacetime but in the internal space of the field. In this context, the object acting as our "connection" is the gauge potential A_μ, whereas in General Relativity, it is the *Levi-Civita connection* $\Gamma^\mu{}_{\nu\rho}$.

Both transform inhomogeneously. The connection arises due to the curvature of spacetime. This raises the question of whether we can switch to a coordinate system where there is no curvature, meaning the connection vanishes.

In gauge theory, we found that the vector changes when parallel transported around a closed loop. The same applies to General Relativity, where the change in the vector will be proportional to a tensor $R^\alpha{}_{\beta\gamma\delta}$, known as the *Riemann Curvature Tensor*, which measures the curvature of spacetime. Our field strength $F_{\mu\nu}$ is analogous to the curvature tensor. Therefore, in a sense, we can think of working in a curved internal space while spacetime remains flat.

These concepts will become clear when we introduce the idea of Fiber Bundles in Chapter 2, where we will revisit this topic from a completely different perspective.

1.3.3 Generalization for SU(N)

The discussion above is quite general, but let us organize these ideas. The cases treated were for the symmetry groups $U(1)$ and $SU(2)$, respectively. The ideas we present now are valid for $SU(N)$.

The generators of the groups satisfy the commutator

$$[G^a, G^b] = if^{abc} G^c, \tag{1.127}$$

where the i is just to ensure that these generators are hermitian and the coefficients f^{abc} are the structure constants, with $a, b, c = 1, 2, ..., \dim\, G$. We may normalize our Lie algebra by imposing

$$\mathrm{Tr}\; G^a G^b = \frac{1}{2}\delta^{ab}. \tag{1.128}$$

Just as before, we write the gauge fields A_μ as

$$A_\mu = A_\mu^a G^a, \tag{1.129}$$

that emphasizes that A_μ as a Lie algebra valued vector. To clarify what that means, one can think about A_μ as a $N \times N$ hermitian matrix. In this notation, the form for the covariant derivative remains the same for any group

$$\mathcal{D}_\mu = \partial_\mu - \frac{i}{e} A_\mu. \tag{1.130}$$

so as the field strength. which will be given by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ie[A_\mu, A_\nu], \tag{1.131}$$

where e is the coupling constant. Finally, the action that leads to Yang-Mills equations is

$$S_{YM} = -\frac{1}{2}\int d^4x \;\mathrm{tr}\; F^{\mu\nu}F_{\mu\nu}. \tag{1.132}$$

Extremizing it with respect to the field A_μ leads to

$$\mathcal{D}_\mu F^{\mu\nu} = 0, \tag{1.133}$$

which is the non-abelian analog to Maxwell equations.

2 FIBER BUNDLES

The formalism of fiber bundles is deeply connected with the Yang-Mills theory. As we have seen, Yang-Mills Theory started with a generalization of the gauge invariance of electrodynamics, which has the $U(1)$ symmetry group, to higher-symmetry groups. It was noted afterward that the mathematical structure behind Yang-Mills was the formalism of fiber bundles. It is crucial to understand this formalism so we might be able to understand the topological structure of several physical objects that will appear later. Here we follow the discussion of [12, 13].

2.1 Introduction

We denote a fiber bundle by (E, π, X, F, G), where

- E is a topological space called the **total space**;
- X is a topological space called the **base space**;
- π is a map called the **projection**, defined by $\pi : E \to X$;
- F is a topological space called the **standard fiber**. It is homeomorphic to every fiber of the total space, $\pi^{-1}(x) = F_x$, with $x \in X$. The F_x are called the fibers;
- G is a group called the **structure group**.

We also require an open covering of the base space $\{U_\alpha\}$ and a homeomorphism ϕ_α defined by

$$\phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F, \tag{2.1}$$

such that

$$\pi \circ \phi_\alpha^{-1}(x, f) = x, \quad x \in U_\alpha \text{ and } f \in F. \tag{2.2}$$

This homeomorphism ϕ_α is called a **local trivialization**. The bundle structure is represented below.

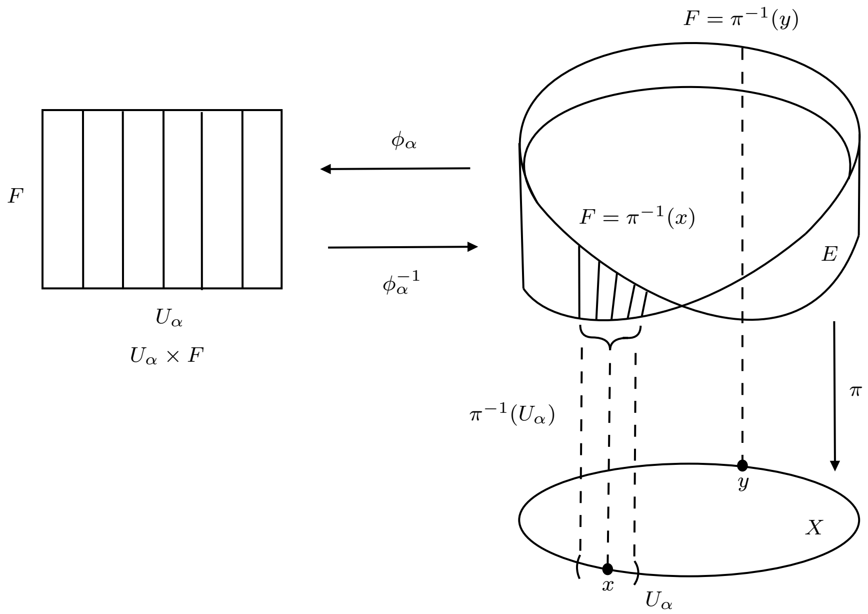

Figura 2.1 – The fiber bundle setup. Adapted from [12].

Let us now understand why the structure group is necessary in this description. Consider two sets of local coordinates (U_α, ϕ_α) and (U_β, ϕ_β) with $U_\alpha \cap U_\beta \neq \emptyset$. Define the map $g_{\alpha\beta}$ by

$$g_{\alpha\beta} \equiv \phi_\alpha \circ \phi_\beta^{-1}. \tag{2.3}$$

This map acts as

$$g_{\alpha\beta} : (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F. \tag{2.4}$$

Keeping the point $x \in U_\alpha \cap U_\beta$ fixed, we obtain

$$g_{\alpha\beta} : F \to F. \tag{2.5}$$

Therefore, the action of this map is to move from one point in the fiber to another; this is why these maps are called **transition functions**. Hence, two points f and f' corresponding to the local trivializations (U_α, ϕ_α) and (U_β, ϕ_β), respectively, are related by

$$f' = g_{\alpha\beta} f. \tag{2.6}$$

Thus, the transition functions tell us how to "glue" one fiber to another over the overlap. These functions satisfy the following properties:

$$g_{\alpha\alpha}(x) = \mathbb{1}, \quad x \in U_\alpha; \tag{2.7}$$

$$g_{\alpha\beta}^{-1}(x) = g_{\beta\alpha}(x), \quad x \in U_\alpha \cap U_\beta; \tag{2.8}$$

$$g_{\alpha\gamma}(x) g_{\gamma\beta}(x) = g_{\alpha\beta}(x), \quad x \in U_\alpha \cap U_\beta \cap U_\gamma. \tag{2.9}$$

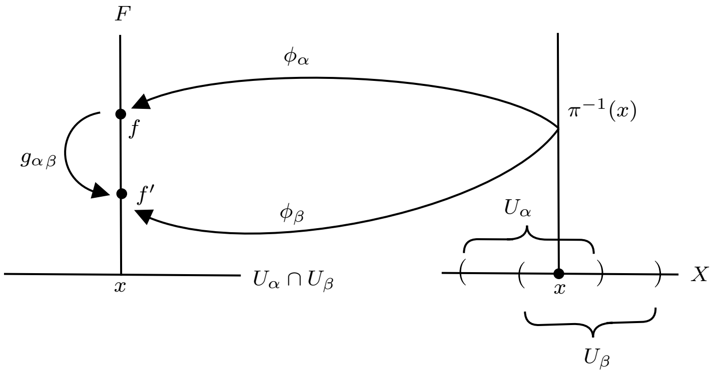

Figura 2.2 – The transition functions. Adapted from [12]

Proposition 2.1 *The transition functions are not unique!*

This can be seen by considering two different fiber bundles, E and E'. Suppose they have the same base space, fiber, and structure group. Let (U_α, ϕ_α) and $(U_\alpha, \varphi_\alpha)$ be their local trivializations, respectively. We define the map

$$h_\alpha \equiv \phi_\alpha \circ \varphi_\alpha^{-1}, \tag{2.10}$$

which is a homeomorphism belonging to the structure group. The transition functions of the two bundles are related by

$$g'_{\alpha\beta}(x) = h_\alpha^{-1}(x) g_{\alpha\beta}(x) h_\beta(x), \quad x \in U_\alpha \cap U_\beta. \tag{2.11}$$

Indeed,

$$h_\alpha^{-1} g_{\alpha\beta} h_\beta = \varphi_\alpha \circ \phi_\alpha^{-1} \circ \phi_\alpha \circ \phi_\beta^{-1} \circ \phi_\beta \circ \varphi_\beta^{-1} = \varphi_\alpha \circ \varphi_\beta^{-1} = g'_{\alpha\beta}. \tag{2.12}$$

We then say that if two bundles differ only by a choice of local coordinates, they are **equivalent**.

Definition 2.1 *Let σ be a map*

$$\sigma : X \to E, \tag{2.13}$$

such that

$$\pi \circ \sigma(x) = x, \quad \forall x \in X.$$

We call this map a ***section****.*

Note that the projection π does not have an inverse, since it maps an entire fiber to a single point. Although we have used the notation π^{-1}, this simply denotes the preimage under π. The section plays the role of a right inverse: given a point in the base space, it selects a point in the corresponding fiber.

We now define some special types of bundles. The section is represented diagrammatically below.

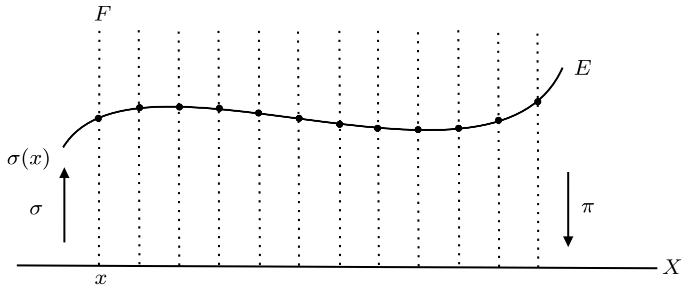

Figura 2.3 – Representation of a section. Adapted from [12]

- **Trivial bundle:** A bundle is called trivial if its transition functions are the identity, that is,

$$g_{\alpha\beta} = \mathbb{1}. \tag{2.14}$$

  In this case, the bundle is simply the direct product of the base space with the fiber,

$$E = X \times F. \tag{2.15}$$

- **Vector bundle:** A bundle whose fiber F is a vector space. It carries a representation $\mathcal{R}$ of the structure group G on the vector space F. A particular case is a **line bundle**, whose fiber F is a one-dimensional vector space.

- **Tangent bundle:** A particular type of vector bundle whose fiber at each point is the tangent space of the manifold.

- **Cotangent bundle:** The dual vector space of the tangent space. This is where one-forms are defined.

- **Principal bundle:** A bundle whose fiber is the structure group itself. We denote it by $P(X, G)$. There is an important theorem concerning principal bundles, which we state below without proof but illustrate with an example.

- **Associated bundle:** A vector bundle is said to be *associated* with a principal bundle if its transition functions are obtained via a representation $\mathcal{R}$ of the structure group acting on the fiber. In other words, we take an element of the principal bundle and represent it in the vector bundle through $\mathcal{R}$.

Theorem 2.1 *The principal bundle $P(X, G)$ and its corresponding bundle (E, π, X, F, G) are trivial if $P(X, G)$ admits a global section.*

An example of this theorem is the Möbius strip. Choosing the structure group to be equal to the fiber, i.e., $F = G = \{\mathbb{1}, g\}$, where g represents a twist, we obtain a double cover of the base space X. Choosing a local coordinate system $\theta \in [0, 2\pi]$, we find that the section satisfies $\sigma(\theta) \neq \sigma(\theta + 2\pi)$. Therefore, we cannot define a global section.

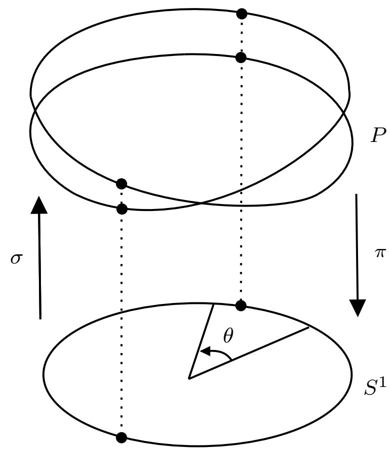

Figura 2.4 – Double cover of the base space. Adapted from [12].

According to the theorem, the Möbius strip and its corresponding principal bundle are nontrivial, which is indeed the case.

Note that although the principal bundle may have no global section, an associated bundle (E, π, X, F, G) might still admit one. Furthermore, several bundles with different fibers but a common base space, structure group, and transition functions share the same principal bundle.

Finally, one may ask what minimal structure is required to reconstruct a bundle. We now show that it is possible to construct a bundle starting from the base space, the fiber, and the transition functions. Start with the disjoint union

$$\tilde{E} = \bigcup_{\alpha} U_\alpha \times F. \tag{2.16}$$

We say that two points (x, f) and (x', f'), with $x \in U_\alpha$ and $f \in F$, are **equivalent** if

$$x = x', \quad f' = g_{\alpha\beta}(x) f. \tag{2.17}$$

We denote this equivalence relation by $\sim$. The fiber bundle is then defined as

$$E = \tilde{E}/\sim, \tag{2.18}$$

with projection

$$\begin{aligned} \pi : &E \to X \\ &[(x, f)] \mapsto x, \end{aligned} \tag{2.19}$$

where $[(x, f)]$ denotes the equivalence class of (x, f). The local trivialization is defined by

$$\begin{aligned} \phi_\alpha^{-1} : &U_\alpha \times F \to \pi^{-1}(U_\alpha) \\ &(x, f) \mapsto [(x, f)]. \end{aligned} \tag{2.20}$$

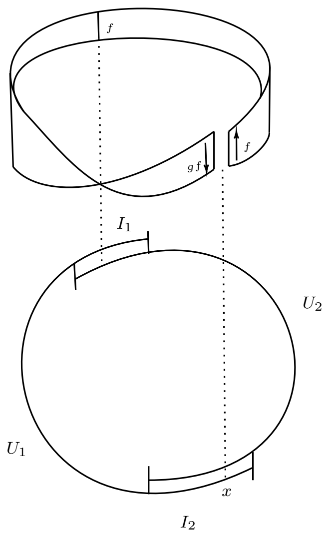

Figura 2.5 – Reconstruction of the Möbius strip. Adapted from [12]

Thus,

$$\pi \circ \phi_\alpha^{-1}(x, f) = x. \tag{2.21}$$

As an example, let us reconstruct the Möbius strip.

As an example, let us reconstruct the Möbius strip. In this case, the base space is S^1, the fiber is a line segment, and the structure group contains two elements $G = \{\mathbb{1}, g\}$, where g represents a twist. We define two open subsets U_1 and U_2 of S^1 such that $U_1 \cap U_2 = I_1 \cup I_2$. Explicitly, the transition functions are

$$\begin{aligned} g_{11} &= g_{22} = \mathbb{1}, \\ g_{12} &= \begin{cases} \mathbb{1}, & \text{if } x \in I_1, \\ g, & \text{if } x \in I_2, \end{cases} \\ g_{21} &= g_{12}^{-1}. \end{aligned} \tag{2.22}$$

In the region I_1, there is only one point (x, f), whereas in the region I_2 there are two points, $\{(x, f), (x, gf)\}$. However, the equivalence relation identifies these two points. When we glue them together according to this identification, we obtain the twist characteristic of the Möbius strip.

2.2 Connection

2.2.1 Abstract definition

We are interested in comparing different points in the fiber. We have tangent vectors at our disposal. However, as the name suggests, these vectors are tangent to points in the bundle. What we seek instead is a vector that points away from a fiber, more specifically in the direction of another fiber. This can be achieved using what we call a **connection**.

Definition 2.2 *A connection is a decomposition of the tangent space T_uP, where u is a point in the principal bundle $P(X,G)$. It satisfies*

$$\begin{aligned} &i) \quad T_uP = V_uP \oplus H_uP, \\ &ii) \quad R_{*g}H_uP = H_{ug}P. \end{aligned} \tag{2.23}$$

The sets V_uP and H_uP are called the **vertical subspace** and **horizontal subspace**, respectively. We will define formally these spaces below. R_g denotes the map implementing the right action of the symmetry group G on the principal bundle P. Therefore, we use the induced map (push-forward) R_*[1].

Note that "vertical" and "horizontal" are merely names, since we do not assume a metric structure that would allow us to define orthogonality between these subspaces.

Condition $i)$ states that any vector $X \in T_uP$ can be decomposed as $X = X^V + X^H$, where $X^V \in V_uP$ and $X^H \in H_uP$. Condition $ii)$ states that once a horizontal vector is parallel transported, all other horizontal vectors in the same fiber are transported accordingly.

To define tangent vectors in the principal bundle, take a point $u \in P(X,G)$ and use the right-translation exponential map

$$R_{\exp(tA)}u = u\exp(tA) = \sigma(t,u), \tag{2.24}$$

which defines a flow, where $A \in \mathfrak{g}$ and $\mathfrak{g}$ is the Lie algebra of G. The vectors belonging to V_uP are generated by such elements A. They are denoted by X_A and defined as

$$X_A f(u) \equiv \frac{d}{dt}\Big|_{t=0} f(u\exp(tA)), \tag{2.25}$$

where $f : P \to \mathbb{R}$ is a smooth function. These vectors preserve the Lie algebra structure:

$$[X_A, X_B] = X_{[A,B]}. \tag{2.26}$$

The vertical space can also be characterized by the fact that motion along a fiber does not change the corresponding point in the base space:

$$V_uP \equiv \{X \in T_uP \mid \pi_* X = 0\}., \tag{2.27}$$

where π_* is the push-forward associated with the projection π.

[1] For a more detailed discussion on Lie algebras and other aspects of differential geometry, see [13].

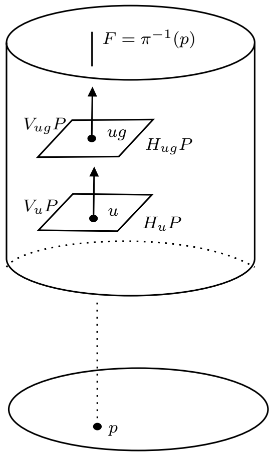

Figura 2.6 – Decomposition of the tangent space into the vertical and horizontal subspace. Adapted from [12].

Definition 2.3 (Horizontal Lift) *We say that $\tilde{X}$ is a horizontal lift of X if*

$$\begin{aligned} &i) \quad \pi_* \tilde{X} = X, \\ &ii) \quad \tilde{X}^V = 0. \end{aligned} \tag{2.28}$$

Definition 2.4 *Let $\gamma : [0,1] \to M$ be a curve. A curve $\tilde{\gamma} : [0,1] \to P$ is a horizontal lift of γ if*

$$\begin{aligned} &i) \quad \pi \circ \tilde{\gamma} = \gamma, \\ &ii) \quad \text{the tangent vectors to } \tilde{\gamma} \text{ belong to } H_{\tilde{\gamma}(t)}P. \end{aligned} \tag{2.29}$$

Theorem 2.2 *Let $\gamma : [0,1] \to M$ be a curve and $u \in \pi^{-1}(\gamma(0))$. Then there exists a unique horizontal lift $\tilde{\gamma} : [0,1] \to P$ such that $\tilde{\gamma}(0) = u$.*

This means that the parallel transport of a fiber point along the curve $\gamma(t)$ is given by the horizontal lift $\tilde{\gamma}(t)$. Note that if a base curve $\gamma : [0,1] \to M$ is closed, its horizontal lift need not be closed. For example, $\tilde{\gamma}(1) = \tilde{\gamma}(0)h$ with $h \in G$. This follows from the concept of **holonomy**, which is related to loss of geometrical data when parallel transporting some vector around a closed loop. The set of such elements $\{h\}$ is what we call the holonomy.

2.2.2 Connection one-form

We now define an object that will help us connect with physics later on and that implies Definition 2.2.

Definition 2.5 *The connection one-form ω is a* ***Lie algebra valued one-form****, that is, $\omega \in T^*P \otimes \mathfrak{g}$, defined as a projection of the tangent space onto the vertical subspace,*

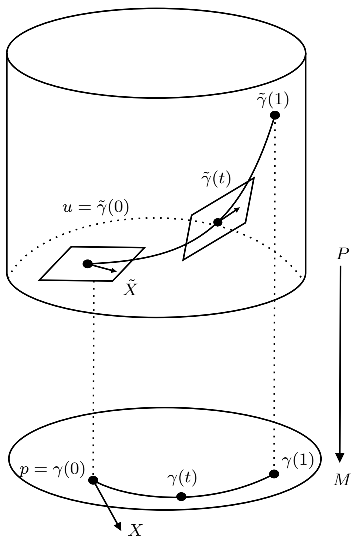

Figura 2.7 – Setup for the horizontal lift. Adapted from [12].

satisfying

$$\begin{aligned} &i) \qquad \omega(X_A) = A, \quad A \in \mathfrak{g}, \\ &ii) \qquad R_g^* \omega = Ad_{g^{-1}} \omega, \end{aligned} \tag{2.30}$$

where R^* *is the induced map* ***pull back*** *and* $Ad_{g^{-1}}$ *is defined as the adjoint map given by*

$$\begin{aligned} &Ad_g : \mathfrak{g} \to \mathfrak{g}, \\ &A \mapsto gAg^{-1}, \quad g \in G, A \in \mathfrak{g}. \end{aligned} \tag{2.31}$$

For some vector $X \in T_uP$ we have

$$R_g^* \omega_{ug}(X) = \omega_{ug}(R_{g*}X) = g^{-1}\omega_u(X)g. \tag{2.32}$$

From $i)$, we can define the horizontal subspace as the kernel of ω[2] as

$$H_uP \equiv \{X \in T_uP \mid \omega(X) = 0\}. \tag{2.33}$$

This again leads to the splitting

$$T_uP = V_uP \oplus H_uP.$$

Condition $ii)$ implies condition $ii)$ of Definition 2.2, as we show below.

We start by defining H_uP as in Eq.(2.33). Let $X_u \in T_uP$, then $(R_g)_*X_u \in T_{ug}P$. We have

$$(R_g)^* \omega_u(X) = \omega_u(R_{g*}X) = g^{-1}\omega_u(X)g. \tag{2.34}$$

[2] Again, the horizontal and vertical vectors are just names and do not have a notion of orthogonality. The way we are defining them is to characterize which vector belongs to which set.

If $X_u \in H_uP$, then $\omega_u(X) = 0$, and therefore $\omega_u(R_{g*}X) = 0$. Hence $R_{g*}X_u \in H_{ug}P$. Since $(R_g)_*$ is an invertible map, any vector $Y \in H_{ug}P$ can be written as $Y = R_{g*}X_u$ for some $X_u \in H_uP$. This proves that $H_{ug}P = R_{g*}H_uP$ for every element of G.

Until now, everything seems too abstract. But now we can finally understand the bridge between this formalism and physics.

Proposition 2.2 *Locally, the connection 1-form ω represents the Yang–Mills gauge potential A.*

Let U_α be an open covering of M. We have the section defined above, given by

$$\sigma_\alpha : U_\alpha \to \pi^{-1}(U_\alpha).$$

This induces the pullback map

$$\sigma_\alpha^* : T^*P \to T^*U_\alpha. \tag{2.35}$$

We then define the gauge potential A as

$$A \equiv \sigma_\alpha^* \omega. \tag{2.36}$$

On the other hand, we have the following theorem.

Theorem 2.3 *Let A_α be a Lie algebra valued 1-form defined on a patch U_α and $\sigma_\alpha : U_\alpha \to \pi^{-1}(U_\alpha)$ a section. There exists exactly one connection 1-form ω on $\pi^{-1}(U_\alpha)$ such that $A_\alpha = \sigma_\alpha^*\omega$, given by*

$$\omega\big|_{U_\alpha} = g_\alpha^{-1}\pi^* A_\alpha g_\alpha + g_\alpha^{-1} d_p g_\alpha, \tag{2.37}$$

where d_p is the exterior derivative on the bundle and g_α is a coordinate in the fiber given by the local trivialization,

$$\begin{aligned} &\phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F, \\ &u \mapsto (p, g_\alpha). \end{aligned} \tag{2.38}$$

Thus,

$$\phi_\alpha(u) = (p, g_\alpha), \quad \textit{with } u = \sigma_\alpha(p) g_\alpha. \tag{2.39}$$

Proof. *a*) First, let us show that (2.37) reproduces (2.36). Take a tangent vector $X \in T_pM$ and lift it to the bundle as $\sigma_{*\alpha}X \in T_{\sigma_\alpha(p)}P$. Note that $\pi_* \circ \sigma_{*\alpha} = \mathbb{1}_{T_pM}$ because $\pi(\sigma_\alpha(p)) = p$. Also, it follows from Eq.(2.39) that $g_\alpha = e$ over σ_α. Then

$$\sigma_\alpha^*\omega(X) = \omega(\sigma_{*\alpha}X) = e^{-1}\pi^* A_\alpha(\sigma_{*\alpha}X)e + e^{-1}d_p e(\sigma_{*\alpha}X) = A_\alpha(\pi_*\sigma_{*\alpha}X) = A_\alpha(X). \tag{2.40}$$

b) We now demonstrate that (2.37) satisfies the properties of a connection.

For axiom i): taking a vector field $X_A \in V_uP$, then $\pi_* X_A = 0$. Hence

$$\begin{aligned}\omega(X_A) &= g_\alpha^{-1}\pi^* A_\alpha(\pi_* X_A) g_\alpha + g_\alpha^{-1} d_p g_\alpha(X_A) \\ &= g_\alpha^{-1}(u)\frac{d}{dt}\bigg|_{t=0} g_\alpha(u\exp(tA)) \\ &= g_\alpha^{-1}(u) g_\alpha(u)\frac{d}{dt}\bigg|_{t=0} \exp(tA) = A.\end{aligned} \tag{2.41}$$

For axiom ii): take a tangent vector $X \in T_uP$ and a group element $h \in G$. Recalling that $g_\alpha(uh) = g_\alpha(u)h$, we have

$$\begin{aligned}R_h^*\omega(X) &= h^{-1}g_\alpha^{-1}(u)\pi^* A_\alpha(X) g_\alpha(u) h + h^{-1} g_\alpha^{-1}(u) d_p g_\alpha(u)(X) h \\ &= h^{-1}\omega(X) h = Ad_{h^{-1}}\omega(X).\end{aligned} \tag{2.42}$$

Since the form ω is defined globally, we have

$$\omega\big|_{U_\alpha} = \omega\big|_{U_\beta}, \quad U_\alpha \cap U_\beta \neq \emptyset. \tag{2.43}$$

Using Eq.(2.37) we obtain

$$g_\alpha^{-1}\pi^* A_\alpha g_\alpha + g_\alpha^{-1} d_p g_\alpha = g_\beta^{-1}\pi^* A_\beta g_\beta + g_\beta^{-1} d_p g_\beta. \tag{2.44}$$

In the overlap, the fiber coordinates are related by $g_\beta = h_{\beta\alpha} g_\alpha$. Then

$$\begin{aligned}g_\alpha^{-1}\pi^* A_\alpha g_\alpha + g_\alpha^{-1} d_p g_\alpha &= h_{\beta\alpha}^{-1} g_\alpha^{-1}\pi^* A_\beta h_{\beta\alpha} g_\alpha + h_{\beta\alpha}^{-1} g_\alpha^{-1} d_p(h_{\beta\alpha} g_\alpha) \\ &= h_{\beta\alpha}^{-1} g_\alpha^{-1}\pi^* A_\beta h_{\beta\alpha} g_\alpha + h_{\beta\alpha}^{-1} d_p h_{\beta\alpha} + g_\alpha^{-1} d_p g_\alpha.\end{aligned} \tag{2.45}$$

Hence

$$g_\alpha^{-1}\pi^* A_\alpha g_\alpha = h_{\beta\alpha}^{-1} g_\alpha^{-1}\pi^* A_\beta h_{\beta\alpha} g_\alpha + h_{\beta\alpha}^{-1} d_p h_{\beta\alpha}. \tag{2.46}$$

Acting with g_α on the left and then with g_α^{-1} on the right, we obtain

$$\pi^* A_\alpha = h_{\beta\alpha}\pi^* A_\beta h_{\beta\alpha}^{-1} + h_{\beta\alpha}^{-1} d_p h_{\beta\alpha}. \tag{2.47}$$

But this is what happens to the potential when pulled back. Its transformation on the base space is given by

$$A_\alpha = h_{\beta\alpha} A_\beta h_{\beta\alpha}^{-1} + h_{\beta\alpha}^{-1} d h_{\beta\alpha}. \tag{2.48}$$

This is what we previously called the **gauge transformation** of the potential. We will see later that this reproduces exactly the transformations we have seen before.

2.2.3 Curvature

To finish our discussion, we now introduce the notion of curvature in the fibers. To do so, we first introduce the concept of a **covariant derivative.**

Consider a p-form ω on a principal bundle $P(M,G)$ with values in a vector space

$$\omega : \underbrace{T^*P \otimes ... \otimes T^*P}_{\text{p-times}} \to V. \tag{2.49}$$

This form can be written as

$$\omega = \omega^a \otimes e_a \in \Lambda^p(P) \otimes V, \tag{2.50}$$

with a sum implicit in the index a, where $a = 1, 2, 3..., k$, the $\{e_\alpha\}$ form a set of basis vectors for the space V, and k is its dimension.

Let $X_1, ..., X_{p+1}$ be a set of tangent vectors on the bundle P, then the **exterior covariant derivative** of a p-form ω is defined by

$$D\omega(X_1, ..., X_{p+1}) = d_p\omega(X_1^H, ..., X_{p+1}^H). \tag{2.51}$$

From this, we define the curvature to be the covariant derivative of the connection 1-form, i.e.,

$$\Omega \equiv D\omega \in \Lambda^2(P) \otimes \mathfrak{g}. \tag{2.52}$$

It is easy to show that

$$R_g^*\Omega = Ad_{g^{-1}}\Omega = g^{-1}\Omega g, \quad g \in G. \tag{2.53}$$

Theorem 2.4 *Ω and ω satisfy the Cartan structure equation.*

$$\Omega(X,Y) = d_p\omega(X,Y) + [\omega(X), \omega(Y)], \tag{2.54}$$

for $X, Y \in T_uP$, where $[\omega(X), \omega(Y)]$ is the Lie bracket between the elements $\omega(X)$ and $\omega(Y)$. In terms of 2-forms we have

$$\Omega = d_p\omega + \omega^2, \tag{2.55}$$

with

$$\omega^2 = \omega\omega = \frac{1}{2}[\omega, \omega]. \tag{2.56}$$

Proof. Let us consider three cases.

First case: $X, Y \in H_uP$. Then, by the definition of the horizontal space, $\omega(X) = \omega(Y) = 0$. Hence

$$\Omega(X,Y) = D\omega(X,Y) = d_p\omega(X^H, Y^H) = d_p\omega(X,Y). \tag{2.57}$$

Second case: $X \in H_uP,\ Y \in V_uP$. This implies that

$$\Omega(X,Y) = d_p\omega(X^H, Y^H) = (d_p\omega)_{\mu\nu}(X^H)^\mu(Y^H)^\nu = 0, \tag{2.58}$$

because $Y^H = 0$. Recalling the relation

$$d_p\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]) = X\omega(Y) - \omega([X,Y]). \tag{2.59}$$

Since $Y \in V_uP$ we can write $Y = X_V$ for some $V \in \mathfrak{g}$. Then $\omega(X_V) = V$, with V constant. Since X acts as a derivative, $X\omega(Y) = 0$. The second term vanishes because $[X, Y] \in H_uP$. This can be seen by noting that the Lie derivative is defined as

$$[Y, X] = \mathcal{L}_X Y = \lim_{t=0} \frac{1}{t}[R_{g^{-1}(t)*}X - X]. \tag{2.60}$$

Since $X \in H_uP$, then $R_{g^{-1}(t)*}X \in H_uP$ as well. Therefore $d_p\omega(X, Y) = 0$.

Third case: $X, Y \in V_uP$. We have (for the same reason as in Eq.(2.58)) that $\Omega(X, Y) = 0$. Hence

$$d_p\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]) = -\omega([X, Y]). \tag{2.61}$$

Since both X and Y belong to the vertical space, their commutator will also be a vector field that belongs to it. Then we can write $[X, Y] = X_A$, for some $A \in \mathfrak{g}$. On the other hand, we can write the X and Y fields as $X = X_B$ and $Y = X_C$ for $B, C \in \mathfrak{g}$. Then

$$X_A = [X_B, X_C]. \tag{2.62}$$

Consider the commutator

$$[\omega(X), \omega(Y)] = [\omega(X_B), \omega(X_C)] = [B, C]. \tag{2.63}$$

Since these fields belong to $\mathfrak{g}$, the Lie algebra structure is preserved, that is,

$$[X_B, X_C] = X_{[B,C]}. \tag{2.64}$$

This implies that

$$[\omega(X), \omega(Y)] = \omega([X, Y]). \tag{2.65}$$

Therefore

$$d_p\omega(X, Y) + [\omega(X), \omega(Y)] = 0. \tag{2.66}$$

Which is in accordance with what was expected. These three cases prove that, indeed, Ω can be written as in Eq.(2.54).

Proposition 2.3 *The curvature 2-form Ω is locally the Yang-Mills field strength, given by*

$$F = \sigma^*\Omega \in \Lambda^2(P) \otimes \mathfrak{g}, \tag{2.67}$$

where σ^* is the pull-back induced by the section, just as for the gauge potential.

Note that, using Eq.(2.55)

$$\begin{aligned} F = \sigma^*\Omega &= \sigma^*(d_p\omega) + \sigma^*(\omega\omega) \\ &= d\sigma^*\omega + \sigma^*\omega\sigma^*\omega = dA + A^2. \end{aligned} \tag{2.68}$$

In components, $F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu$ and $A = A_\mu dx^\mu$. Then

$$\begin{aligned}\frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu &= \partial_\mu A_\nu dx^\mu \wedge dx^\nu + \frac{1}{2}[A_\mu, A_\nu]dx^\mu \wedge dx^\nu \\ \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu &= \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu \wedge dx^\nu + \frac{1}{2}[A_\mu, A_\nu]dx^\mu \wedge dx^\nu,\end{aligned} \tag{2.69}$$

which leads to

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{2.70}$$

This is the Yang-Mills field strength as we have seen before.

Last, we show its transformation. Consider the overlap $U_\alpha \cap U_\beta \neq \emptyset$. We note that

$$F_\beta = dA_\beta + A_\beta^2. \tag{2.71}$$

Using the transformation

$$A_\beta = h_{\alpha\beta}^{-1}A_\alpha h_{\alpha\beta}, \tag{2.72}$$

we obtain

$$\begin{aligned}F_\beta &= dA_\beta + A_\beta^2 = d(h_{\alpha\beta}^{-1}A_\alpha h_{\alpha\beta}) + (h_{\alpha\beta}^{-1}A_\alpha h_{\alpha\beta})^2 \\ &= dh_{\alpha\beta}^{-1}A_\alpha h_{\alpha\beta} + h_{\alpha\beta}^{-1}dA_\alpha h_{\alpha\beta} + dh_{\alpha\beta}^{-1}dh_{\alpha\beta} + h_{\alpha\beta}^{-1}A_\alpha^2 h_{\alpha\beta} + h_{\alpha\beta}^{-1}dh_{\alpha\beta}h_{\alpha\beta}^{-1}dh_{\alpha\beta} \\ &+ h_{\alpha\beta}^{-1}dh_{\alpha\beta}h_{\alpha\beta}^{-1}A_\alpha h_{\alpha\beta}.\end{aligned} \tag{2.73}$$

Since $h_{\alpha\beta}h_{\alpha\beta}^{-1} = \mathbb{1} \Rightarrow dh_{\alpha\beta}^{-1}h_{\alpha\beta} + h_{\alpha\beta}^{-1}dh_{\alpha\beta} = 0 \Rightarrow dh_{\alpha\beta}^{-1}h_{\alpha\beta} = -h_{\alpha\beta}^{-1}dh_{\alpha\beta}$. Then

$$F_\beta = h_{\alpha\beta}^{-1}dA_\alpha h_{\alpha\beta} + h_{\alpha\beta}^{-1}A_\alpha^2 h_{\alpha\beta} = h_{\alpha\beta}^{-1}(dA_\alpha + A_\alpha^2)h_{\alpha\beta} = h_{\alpha\beta}^{-1}F_\alpha h_{\alpha\beta}. \tag{2.74}$$

This is the gauge transformation for the Yang-Mills field strength.

2.2.4 Bridging Yang-Mills Theory and Fiber Bundles

Now that we understand Yang-Mills theory and the formalism of fiber bundles, we can describe the bridge between these two subjects.

We have already seen that the Yang-Mills potential is locally the pull-back of the connection 1-form defined globally over the principal bundle, and that the Yang-Mills field strength is locally the pull-back of the curvature 2-form, also defined globally over the principal bundle.

The map used to perform the pull-back is the section σ. This is what we previously understood as the field in Yang-Mills theory. Therefore, the internal space of the field is the principal bundle. When we were performing the parallel transport of vectors in the field's internal space, we were in fact parallel transporting vectors on the principal bundle.

The definition of the exterior covariant derivative is consistent with our previous notion of the covariant derivative because, at that stage, it was a derivative acting on objects defined in the internal space of the field. Now we know that the internal space of the field is the principal bundle. Thus, it makes sense to define it as above. It is a derivative that acts on the principal bundle over horizontal vectors, which are precisely the ones being parallel transported.

Also, we now see from a different perspective why we call the gauge potential Lie-algebra valued. We have seen before that it is an object that can be written as a linear combination of the Lie group generators. Now we see that it is a direct product of the space of p-forms with the set of left-invariant vector fields.

2.3 Physics in terms of fiber bundles

To make things clearer and to understand where the physics we are used to appears in this formalism, we will discuss two examples: Maxwell theory and the magnetic monopole.

2.3.1 Maxwell Theory

The fiber bundle structure of Maxwell theory is quite simple and does not have many interesting aspects, but it provides valuable insight.

We start by writing the Maxwell equations in terms of differential forms. We know that in covariant form they are given by

$$\partial_\mu \star F^{\mu\nu} = 0 \quad \text{(homogeneous)}, \tag{2.75}$$

$$\partial_\mu F^{\mu\nu} = j^\nu \quad \text{(inhomogeneous)}, \tag{2.76}$$

where $\star$ denotes the Hodge star operator. We can relate these components to a 2-form by

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu, \tag{2.77}$$

$$\star F = \frac{1}{2} \star F_{\mu\nu} dx^\mu \wedge dx^\nu. \tag{2.78}$$

In matrix form,

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}, \tag{2.79}$$

and

$$\star F^{\mu\nu} = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & -E_3 & E_2 \\ B_2 & E_3 & 0 & -E_1 \\ B_3 & -E_2 & E_1 & 0 \end{pmatrix}. \tag{2.80}$$

Hence we can write the 2-form F as

$$F = E_i dx^0 \wedge dx^i - \frac{1}{2} B_i \epsilon_{ijk} dx^j \wedge dx^k. \tag{2.81}$$

Acting with the exterior derivative, we obtain

$$dF = \frac{1}{2}(\partial_k E_i - \partial_i E_k - \partial_t B_j \epsilon_{jik}) dx^0 \wedge dx^i \wedge dx^k - \partial_i B_i dx^1 \wedge dx^2 \wedge dx^3, \tag{2.82}$$

where we have used $dx^i \wedge dx^j \wedge dx^k = \epsilon^{ijk} dx^1 \wedge dx^2 \wedge dx^3$ and $\epsilon_{ijk}\epsilon^{ljk} = 2\delta_i^l$. The homogeneous Maxwell equations are given by

$$\epsilon_{jlm}\partial_l E_m + \partial_t B_j = 0 \tag{2.83}$$

$$\partial_i B_i = 0. \tag{2.84}$$

Multiplying by ϵ_{jik} and using again the contraction of the Levi-Civita symbol, we obtain

$$\partial_k E_i - \partial_i E_k - \partial_t B_j \epsilon_{jik} = 0. \tag{2.85}$$

Thus the homogeneous Maxwell equations in terms of differential forms are precisely

$$dF = 0. \tag{2.86}$$

For the inhomogeneous equations, we start by writing

$$\star F = -B_i dx^0 \wedge dx^i - \frac{1}{2} E_i \epsilon_{ijk} dx^j \wedge dx^k. \tag{2.87}$$

Taking the exterior derivative,

$$d \star F = \frac{1}{2}(\partial_i B_k - \partial_k B_i - \partial_t E_j \epsilon_{ijk}) dx^0 \wedge dx^i \wedge dx^k - \partial_i E_i dx^1 \wedge dx^2 \wedge dx^3. \tag{2.88}$$

The inhomogeneous equations can be written as

$$\partial_i B_k - \partial_k B_i - \partial_t E_j \epsilon_{ijk} = j_j \epsilon_{jik} \tag{2.89}$$

$$\partial_i E_i = \rho. \tag{2.90}$$

Defining the current 1-form,

$$j = j_\mu dx^\mu = j_0 dx^0 - j_i dx^i, \tag{2.91}$$

and taking its dual,

$$\star j = j_0 \star dx^0 - j_i \star dx^i = \rho dx^1 \wedge dx^2 \wedge dx^3 - \frac{1}{2} j_j \epsilon_{jik} dx^0 \wedge dx^i \wedge dx^k, \tag{2.92}$$

where $j_0 \equiv \rho$, we find

$$d \star F = - \star j. \tag{2.93}$$

Thus the four Maxwell equations in terms of differential forms become

$$dF = 0 \quad \text{(homogeneous)}, \tag{2.94}$$

$$d\star F = -\star j \quad \text{(inhomogeneous)}. \tag{2.95}$$

Note that since F is closed, we have

$$F = dA, \tag{2.96}$$

because $d^2 = 0$. For the same reason, A is defined up to the derivative of a 0-form,

$$A' = A + d\Lambda, \tag{2.97}$$

which is the gauge transformation. From this we can write down the action for Maxwell theory in terms of differential forms. Since we are integrating over spacetime, the action must be given in terms of 4-forms. There are two 4-forms compatible with the symmetries of the theory that reproduce the action with which we are already familiar, namely

$$A \wedge \star j = A_\mu j^\mu dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3, \tag{2.98}$$

and

$$F \wedge \star F = \frac{1}{2} F_{\mu\nu} F^{\mu\nu} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3. \tag{2.99}$$

The oriented volume element is $dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 = d^4x$, then

$$S[A] = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu j^\mu \right) = \int \left(-\frac{1}{2} F \wedge \star F - A \wedge \star j \right). \tag{2.100}$$

Note that under a gauge transformation the action becomes

$$S[A + d\Lambda] - S[A] = -\int d\Lambda \wedge \star j = \int \Lambda \wedge d \star j = 0, \tag{2.101}$$

where we have used Stokes' theorem and discarded the boundary term. This corresponds to current conservation,

$$d \star j = 0. \tag{2.102}$$

Therefore, to construct the fiber bundle, we take the structure group to be the $U(1)$ gauge group, the base space to be the 4-dimensional Minkowski spacetime, and the fiber to be the set of all possible $U(1)$ group elements. The principal bundle is a trivial bundle given by $P = \mathbb{R}^4 \times U(1)$ because the base space is contractible. The connection is the gauge field A, the curvature is the field strength F with the factor i from the Lie algebra multiplying it, and the transition functions are given by

$$\begin{aligned} h_{\alpha\beta} : & U_\alpha \cap U_\beta \to U(1) \\ & h_{\alpha\beta}(x) = e^{i\Lambda(x)}, \quad x \in U_\alpha \cap U_\beta. \end{aligned} \tag{2.103}$$

The compatibility condition (2.48) gives the gauge transformation

$$A'(x) = A(x) + i d\Lambda(x), \tag{2.104}$$

which is Eq.(2.97) but with the factor of i from the Lie algebra.

2.3.2 The Dirac Monopole

The Dirac monopole is far more interesting than usual Maxwell theory because it has a non-trivial fiber bundle structure. This was proposed by Paul Dirac in 1931 [14].

We start with a magnetic monopole with magnetic charge g localized at the origin, given by

$$\rho_m(\boldsymbol{x}) = 4\pi g \delta^3(\boldsymbol{x}), \tag{2.105}$$

with the corresponding magnetic field

$$\boldsymbol{B}(\boldsymbol{x}) = g\frac{\hat{r}}{r^2}, \quad \hat{r} = \frac{\boldsymbol{r}}{|\boldsymbol{r}|}. \tag{2.106}$$

Note that a possible potential that reproduces this field is

$$\boldsymbol{A}_+ = g\left(\frac{1-\cos\theta}{r\sin\theta}\right)\hat{\boldsymbol{\varphi}}. \tag{2.107}$$

However, Eq.(2.107) is not defined for $\theta = \pi$. We can circumvent this problem by using the potential

$$\boldsymbol{A}_- = g\left(\frac{-1-\cos\theta}{r\sin\theta}\right)\hat{\boldsymbol{\varphi}}. \tag{2.108}$$

But now this potential has a singularity at $\theta = 0$.

In general, it is not possible to write a potential $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ without singularities in $\boldsymbol{A}$; let us check this. Suppose that we can write $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ without singularities. Then, using Stokes' theorem and considering that S^2 is the boundary of some manifold Σ,

$$\int_{S^2} d\boldsymbol{a} \cdot \boldsymbol{B} = \int_\Sigma d^3x\ \boldsymbol{\nabla} \cdot \boldsymbol{B} = \int_\Sigma d^3x\ \rho(\boldsymbol{x}) = 4\pi g. \tag{2.109}$$

However, using Stokes' theorem in the "opposite" way,

$$\int_{S^2} d\boldsymbol{a} \cdot \boldsymbol{B} = \int_{S^2} d\boldsymbol{a} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A}) = \int_{\partial S^2} d\boldsymbol{x} \cdot \boldsymbol{A} = 0. \tag{2.110}$$

The integral vanishes because S^2 has no boundary. This gives a contradiction. Therefore, we could not have used Stokes' theorem in the first place, which means that our potential must indeed have a singularity. Then which potential should we use? The answer is: both. Because of the two singularities, we need one potential that avoids $\theta = 0$ and another that avoids $\theta = \pi$. These two regions can be seen in Fig.(2.8). Therefore, we define

$$\boldsymbol{A}_+ = g\left(\frac{1-\cos\theta}{r\sin\theta}\right)\hat{\boldsymbol{\varphi}}, \quad \text{for } \theta < \pi - \epsilon : U_+, \tag{2.111}$$

$$\boldsymbol{A}_- = g\left(\frac{-1-\cos\theta}{r\sin\theta}\right)\hat{\boldsymbol{\varphi}}, \quad \text{for } \theta > \epsilon : U_-. \tag{2.112}$$

In the intersection region, it is reasonable to expect that these potentials are related by a gauge transformation,

$$\boldsymbol{A}_- = \boldsymbol{A}_+ + \boldsymbol{\nabla}\Lambda. \tag{2.113}$$

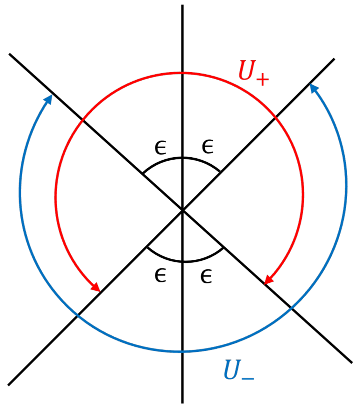

Figura 2.8 – Region for the definition of the potentials. Adapted from [12]

Recalling the expression for $\boldsymbol{\nabla}\Lambda$ in polar coordinates, for $\theta \neq 0, \pi$, we have

$$A_{\varphi_-} = A_{\varphi_+} + \frac{1}{r\sin\theta}\frac{\partial\Lambda}{\partial\varphi}. \tag{2.114}$$

It is easy to solve this equation, and we find that this gauge transformation exists and is given by

$$\Lambda = -2g\varphi. \tag{2.115}$$

We now move on to a quantum mechanical approach to this problem, which leads to a very interesting result.

Start with the Schrödinger equation for a particle with charge e in the presence of a magnetic field $\boldsymbol{B}$,

$$\left[\frac{1}{2m}\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 + e\phi\right]\psi = i\hbar\frac{\partial\psi}{\partial t}, \tag{2.116}$$

where we have inserted the appropriate units again. This equation is invariant under the following transformations,

$$\boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\Lambda, \tag{2.117}$$

$$\phi' = \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t}, \tag{2.118}$$

$$\psi' = \exp\left[i\frac{e}{\hbar c}\Lambda\right]\psi. \tag{2.119}$$

Just as the potentials are connected by a gauge transformation, the wavefunctions are as well,

$$\psi_- = \exp\left[-i\frac{2eg}{\hbar c}\varphi\right]\psi_+. \tag{2.120}$$

The wavefunctions must be single-valued in order for the probability not to be ambiguous. This means

$$\varphi = 0 \Rightarrow \psi_+ = \psi_-, \tag{2.121}$$

$$\varphi = 2\pi \Rightarrow \psi_- = \exp\left[-i\frac{2eg}{\hbar c}2\pi\right]\psi_+ \equiv \psi_+, \tag{2.122}$$

which implies

$$eg = \pm \frac{\hbar c}{2} n, \quad n = 0, 1, ... \tag{2.123}$$

Thus the product of the electric charge and the magnetic charge must be quantized. We are now able to understand what is happening mathematically in the structure of spacetime.

First, recall the expression for the gradient in polar coordinates,

$$\boldsymbol{\nabla} = \hat{\boldsymbol{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\varphi}} \frac{1}{r \sin\theta} \frac{\partial}{\partial \varphi} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta}. \tag{2.124}$$

We want to write it in the cotangent basis, that is,

$$d = dr \frac{\partial}{\partial r} + d\varphi \frac{\partial}{\partial \varphi} + d\theta \frac{\partial}{\partial \theta}. \tag{2.125}$$

If we compare the basis terms, we obtain

$$\hat{\boldsymbol{r}} \Leftrightarrow dr, \tag{2.126}$$

$$\hat{\boldsymbol{\varphi}} \Leftrightarrow r \sin\theta d\varphi, \tag{2.127}$$

$$\hat{\boldsymbol{\theta}} \Leftrightarrow r d\theta. \tag{2.128}$$

Note that this is not an equality since it compares a one-form with a vector. There are some subtleties involved, but we will not delve into them here. With this, we can write (2.111) as

$$A_{\pm} = g \left(\frac{\pm 1 - \cos\theta}{r \sin\theta} \right) r \sin\theta d\varphi = g(\pm 1 - \cos\theta) d\varphi. \tag{2.129}$$

Again, the 1-forms A_+ and A_- are linked by a gauge transformation (in natural units),

$$A_+ = A_- + d\Lambda = A_- + 2g d\varphi = A_- + n d\varphi. \tag{2.130}$$

Taking the exterior derivative,

$$\begin{aligned} dA_+ = dx^{\mu} \frac{\partial A_+}{\partial x^{\mu}} &= \left(dr \frac{\partial}{\partial r} + d\varphi \frac{\partial}{\partial \varphi} + d\theta \frac{\partial}{\partial \theta} \right) A_+ = d\theta \frac{\partial A_+}{\partial \theta} = g \sin\theta d\theta \wedge d\varphi \\ &= \frac{n}{2} \sin\theta d\theta \wedge d\varphi. \end{aligned} \tag{2.131}$$

The same holds for A_-,

$$dA_- = \frac{n}{2} \sin\theta d\theta \wedge d\varphi. \tag{2.132}$$

Despite the two potentials giving the same field strength, it is important to note that

$$\begin{cases} F = dA_+ & \text{in the region } U_+, \\ F = dA_- & \text{in the region } U_-. \end{cases} \tag{2.133}$$

Thus, in the case of the Dirac monopole, it is impossible to define an exact 2-form globally, since we need at least two potentials to avoid the singularity.

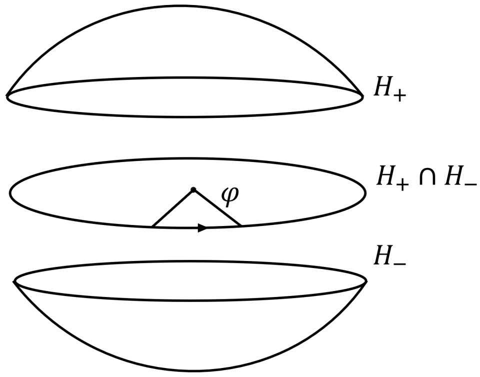

Figura 2.9 – Bundle structure for the Dirac monopole. Adapted from [12]

We now proceed to the fiber bundle structure of the Dirac monopole. The base space we are working on is $\mathbb{R}^3 \backslash \{0\}$. This space is homotopic to the sphere, i.e., $\mathbb{R}^3 \backslash \{0\} \sim S^2$. We use the $U(1)$ group as the structure group and the fiber as all possible elements of $U(1)$, so it is a principal bundle $P(S^2, U(1))$. Since the fibers are rather abstract, we use the wavefunction $\psi(x)$ as a representation of the associated vector bundle, which is a line bundle. Our bundle will be covered by the sets H_+ and H_-, defined as

$$H_+ \equiv \{(r, \theta, \varphi) | \, 0 < r < +\infty, 0 < \varphi \leq 2\pi, 0 < \theta \leq \pi\}, \tag{2.134}$$

$$H_- \equiv \{(r, \theta, \varphi) | \, 0 < r < -\infty, 0 < \varphi \leq 2\pi, 0 < \theta \leq -\pi\}, \tag{2.135}$$

these sets are represented in Fig.(2.9). Locally, the bundle has the coordinates $(\theta, \varphi, e^{i\alpha_+}) \in H_+ \times U(1)$ and $(\theta, \varphi, e^{i\alpha_-}) \in H_- \times U(1)$, with $0 \leq \theta < \pi, 0 \leq \varphi < 2\pi$.

The transition functions are defined in the overlap $H_+ \cap H_-$ and link both the charts and the wavefunctions,

$$e^{i\alpha_+} = h_{\mp} e^{i\alpha_-}, \tag{2.136}$$

$$\psi_+(x) = h_{\mp} \psi_-(x). \tag{2.137}$$

They are constructed in such a way that they glue all the fibers together when completing a full turn around the equator. Hence, one choice that satisfies this condition is

$$h_{\mp} = e^{in\varphi}. \tag{2.138}$$

In natural units, this corresponds again to the Dirac quantization condition. The connection is constructed using the gauge potentials,

$$\mathcal{A}_{\pm} = iA_{\pm}, \tag{2.139}$$

which include the correct factor of i from the Lie algebra. From (2.48),

$$\mathcal{A}_+ = h_{\mp}^{-1} \mathcal{A}_- h_{\mp} + h_{\mp}^{-1} dh_{\mp} = \mathcal{A}_- + ind\varphi, \tag{2.140}$$

which corresponds to (2.130). The curvature tensor is

$$\mathcal{F} = iF, \tag{2.141}$$

or

$$\mathcal{F} = d\mathcal{A}_{\pm}. \tag{2.142}$$

From this construction, we note again that each $\mathcal{A}$ is defined only in a given region (the northern hemisphere of S^2 and the southern hemisphere). We therefore conclude from the bundle structure that this curvature 2-form is not globally exact, but only locally exact.

3 SPONTANEOUS SYMMETRY BREAKING, GOLDSTONE THEOREM AND THE HIGGS MECHANISM

3.1 Spontaneous Symmetry Breaking

From Noether's theorem, we conclude that symmetries play an extremely important role in physical theories. But one question that may arise is: what if there are some circumstances in which some symmetries of our theory no longer prevail? This is when the phenomenon of *spontaneous symmetry breaking (SSB)* arises. It consists of some symmetries of the theory not being realized in a different state. It might seem at first glance a very simple feature, but as we will see, it has enormous consequences. Here we follow the discussion of [8–10].

3.1.1 Spontaneous breaking of a global U(1) symmetry

We start with the simplest Lagrangian where this phenomenon manifests, the interacting ϕ^4 theory,

$$\mathcal{L} = (\partial_\mu \phi)(\partial^\mu \phi)^* - m^2|\phi|^2 - \lambda|\phi|^4. \tag{3.1}$$

The potential is

$$V(\phi) = m^2|\phi|^2 + \lambda|\phi|^4. \tag{3.2}$$

Its minima occur at

$$\begin{aligned} &\frac{\partial V}{\partial|\phi|} = 2m^2|\phi| + 4\lambda|\phi|^3 = 0, \\ &2|\phi|\left(m^2 + 2\lambda|\phi|^2\right) = 0. \end{aligned} \tag{3.3}$$

With solutions

$$|\phi| = 0, \tag{3.4}$$

$$|\phi|^2 = -\frac{m^2}{2\lambda}. \tag{3.5}$$

For $m^2 > 0$, $|\phi| = 0$ is a minimum point. Since we are solving for the modulus of the field, the complex solutions are not physical. For $m^2 < 0$ we must analyze further. Taking the second derivative

$$\frac{\partial^2 V}{\partial \phi^2} = 2m^2 + 12\lambda|\phi|^2. \tag{3.6}$$

Substituting the solution for $|\phi|^2$ into (3.6) leads to $-4m^2$, which is positive. This shows that both are minimum points. For a constant field $\phi = v$ that satisfies the equations of motion, the classical expectation value of this field must be equal to v everywhere in space.

The quantum expectation value is connected to the classical one by taking the limit $\hbar \to 0$ in the path integral[1],

$$\lim_{\hbar \to 0} \langle 0|\phi|0\rangle = \lim_{\hbar \to 0} \int D\phi \, e^{\frac{iS}{\hbar}} \phi = v. \tag{3.7}$$

In terms of the vacuum expectation value (VEV), it is easy to understand why we say that some symmetry is broken. This Lagrangian has a discrete $\mathbb{Z}_2$ symmetry $\phi \to -\phi$. Since we have

$$v = \pm \left(\frac{-m^2}{2\lambda} \right)^{1/2}, \tag{3.8}$$

there must exist two different states that lead to these solutions. Let us call them $|0^+\rangle$ and $|0^-\rangle$ with

$$\langle ^+0|\phi|0^+\rangle = \left(\frac{-m^2}{2\lambda} \right)^{1/2}, \tag{3.9}$$

$$\langle ^-0|\phi|0^-\rangle = -\left(\frac{-m^2}{2\lambda} \right)^{1/2}. \tag{3.10}$$

Therefore, the symmetry must take one vacuum into the other. But first, one of them must be chosen. Once we choose a vacuum, we may expand the field around it, for instance, choosing (3.9),

$$\phi = \left(\frac{-m^2}{2\lambda} \right)^{1/2} + \tilde{\phi}. \tag{3.11}$$

If we substitute this field into the Lagrangian, we get

$$\mathcal{L} = \partial_\mu \tilde{\phi} \partial^\mu \tilde{\phi}^* + \frac{m^4}{2\lambda} - m^2 \left(\frac{-m^2}{2\lambda} \right)^{1/2} \tilde{\phi} - m^2 \left(\frac{-m^2}{2\lambda} \right) \tilde{\phi}^* - m^2 |\tilde{\phi}|^2 + \text{quartic terms}. \tag{3.12}$$

This Lagrangian is no longer invariant under $\tilde{\phi} \to -\tilde{\phi}$. Thus we note that the symmetry has been broken. In other words, when we have a non-trivial VEV for the field, we say that the field condenses and the vacuum does not share the same symmetries as the Lagrangian. Therefore the symmetry is said to be *spontaneously broken*.

There is a very interesting aspect involving the presence of spontaneous symmetry breaking. Another internal symmetry that this Lagrangian has is a continuous global $U(1)$ symmetry $\phi \to e^{i\Lambda}\phi$. Note that this symmetry is also spontaneously broken because the Lagrangian (3.12) is not invariant under it.

It is convenient to reparametrize the complex field $\phi(x)$ in terms of two real fields $\rho(x)$ and $\theta(x)$

$$\phi(x) = (\rho(x) + v) e^{i\theta(x)}, \tag{3.13}$$

where we choose the vacuum[2]

$$\langle \rho \rangle = 0, \quad \langle \theta \rangle = 0. \tag{3.14}$$

[1] There are higher order corrections, but for our analysis this is sufficient.

[2] In this chapter, every expectation value will be assumed to be calculated in the vacuum unless stated otherwise.

We now substitute this field into the Lagrangian,

$$\begin{aligned}\mathcal{L} &= \partial_\mu\rho(x)\partial^\mu\rho(x) + (\rho+v)^2(\partial_\mu\theta)(\partial^\mu\theta) - m^2(\rho(x)+v)(\rho(x)+v) \\ &- \lambda[(\rho(x)+v)(\rho(x)+v)]^2 \\ &= \partial_\mu\rho\partial^\mu\rho + (\rho+v)^2(\partial_\mu\theta)(\partial^\mu\theta) - m^2\rho^2 - 2m^2v\rho - m^2v^2 \\ &- \lambda(\rho^4 + 4v\rho^3 + 6v^2\rho^2 + 4v^3\rho + v^4).\end{aligned} \tag{3.15}$$

Using (3.8) we obtain

$$\mathcal{L} = \partial_\mu\rho(x)\partial^\mu\rho(x) + (\rho+v)^2(\partial_\mu\theta)(\partial^\mu\theta) + \lambda\rho^4 + 4v\lambda\rho^3 + 4\lambda v^2\rho^2 - \lambda v^4. \tag{3.16}$$

Thus we see that the mass term for the $\theta(x)$ field has vanished from the Lagrangian and only the field $\rho(x)$ is left with a mass

$$m_\rho^2 = 4\lambda v^2. \tag{3.17}$$

The massless field $\theta(x)$ is called a *Goldstone boson* and it is the result of the spontaneous breaking of the continuous $U(1)$ symmetry.

To visualize this, we recall that the field can be written as $\phi = \phi_1 + i\phi_2$. We can see from Fig.(3.1) that the minimum of the potential lies along $|\phi| = v$ and the other possible vacua are connected by a rotation.

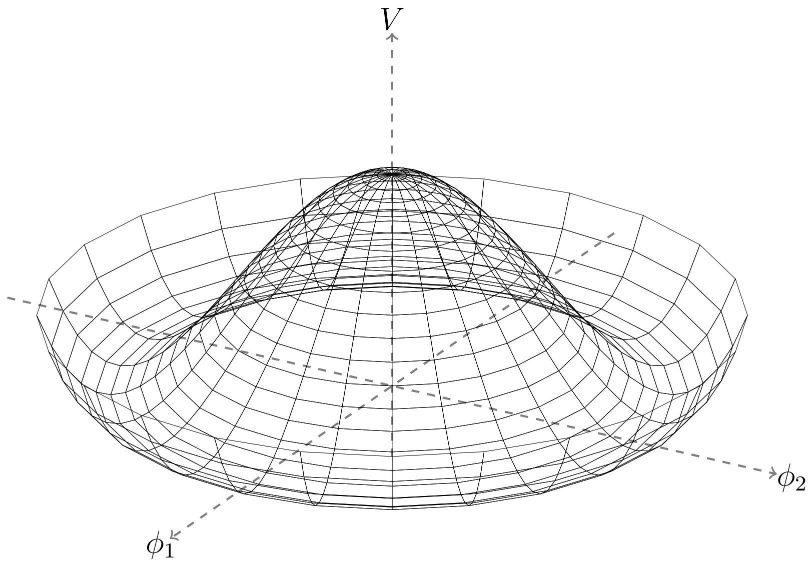

Figura 3.1 – Mexican hat potential. The minima are at $|\phi| = v$. The other minima are connected by a rotation. Adapted from [8].

This is precisely the rotation in the internal space that we have seen in Chapter 1. This is further evidence of why the massless particle results from the breaking of the $U(1)$ symmetry. With this we can enunciate the *Goldstone theorem*.

Theorem 3.1 (Goldstone Theorem) *To each broken continuous global symmetry, there exists a massless field associated with it.*

The Goldstone theorem will be proved in the next section for the general case, but before we do that, it is interesting to analyze the case of a higher symmetry group such as $SO(3)$, where we can understand how many Goldstone bosons will appear in our theory.

Consider the $SO(3)$ group. As we have seen before, the field has an internal structure so we can interpret ϕ_i as a vector in the internal space. The Lagrangian is given by

$$\mathcal{L} = \partial_\mu \phi_i \partial^\mu \phi_i - \frac{m^2}{2}(\phi_i \phi_i) - \lambda(\phi_i \phi_i)^2, \; i = 1, 2, 3. \tag{3.18}$$

This has a similar structure to the Lagrangian for the complex scalar field; thus the minima are situated at $\langle \phi \rangle = v$ for $m^2 < 0$, where

$$\langle \phi \rangle = \langle (\phi_1^2 + \phi_2^2 + \phi_3^2)^{1/2} \rangle = \left(-\frac{m^2}{4\lambda} \right)^{1/2} \equiv v. \tag{3.19}$$

We again have a degeneracy and must choose one vacuum. The one that we choose is

$$\langle \boldsymbol{\phi} \rangle = v \hat{\boldsymbol{e}}_{\mathbf{3}}. \tag{3.20}$$

Note that this choice is not invariant under the whole group $SO(3)$ but is invariant under a subgroup of $SO(3)$ given by the rotations around the $\hat{\boldsymbol{e}}_{\mathbf{3}}$ direction only. Nevertheless, $V(\phi)$ remains invariant under $SO(3)$.

Just as before, we may expand the field around the VEV. In this case, we will expand around ϕ_3, since this is the vacuum that we have chosen. Hence

$$\phi_3 = \chi + v. \tag{3.21}$$

We now plug this into the Lagrangian. For simplicity we can just evaluate the potential terms, since the kinetic term does not give any mass terms, and we can neglect the other terms that are not quadratic in the fields since those also do not give mass terms. Then

$$\begin{aligned} V &= \frac{m^2}{2}(\phi_1^2 + \phi_2^2 + (\chi + v)^2) + \lambda(\phi_1^2 + \phi_2^2 + (\chi + v)^2)^2 \\ &= \frac{m^2}{2}(\phi_1^2 + \phi_2^2 + \chi^2) + \lambda(6v^2\chi^2 + 2\phi_1^2 + 2\phi_2^2) + \text{linear} + \text{cubic} + \text{quartic terms}. \end{aligned} \tag{3.22}$$

Then, using (3.19) we find

$$V = 8v^2 \lambda \chi^2 + \text{linear} + \text{cubic} + \text{quartic terms}. \tag{3.23}$$

The field χ has a mass $m_\chi^2 = 8v^2\lambda$ and the mass terms for the two fields ϕ_1 and ϕ_2 have disappeared. Thus there are 2 Goldstone bosons associated with the $O(3)$ broken symmetry.

This simple example lets us build the insight that the number of Goldstone bosons is related to the dimension of the symmetry group. Note that in this case, there is a subgroup of $O(3)$ that leaves our vacuum choice invariant. It is the group of rotations around the 3-direction. This is the $SO(2)$ group, which has one generator. Thus, the number of Goldstone bosons is given by the dimension of the group $SO(3)/SO(2)$. This can be seen in a more general way by considering the potential and expanding it about its minimum

$$V(\phi) = V(\phi_0) + \frac{1}{2}\left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\right)\Bigg|_{\phi=\phi_0} \chi_i\chi_j + \mathcal{O}(\chi^3), \tag{3.24}$$

with $\chi(x) = \phi(x) - \langle\phi(x)\rangle$ and ϕ_0 being the point where V has a minimum. Then we know that

$$M_{ij} \equiv \left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\right)\Bigg|_{\phi=\phi_0} \geq 0, \tag{3.25}$$

since M_{ij} is a mass matrix. Since $V(\phi)$ is invariant under transformations of the group G, this induces the expansion of the potential,

$$V(\phi_0) = V(U(g)\phi_0) + \frac{1}{2}\left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\right)\Bigg|_{\phi=\phi_0} \delta\phi_i\delta\phi_j + \mathcal{O}(\delta\phi^3), \tag{3.26}$$

where $\delta\phi$ is the transformation of the field under the symmetry. Then, for the invariance of the potential under G to hold, we must have

$$M_{ij}\delta\phi_i\delta\phi_j = 0. \tag{3.27}$$

We have two possible cases.

If $U(g)$ is a symmetry of the vacuum, then

$$U(g)\langle\phi\rangle = \langle\phi\rangle \Rightarrow \delta\phi = 0, \tag{3.28}$$

and $g \in H$. If $U(g)$ is not a symmetry of the vacuum, then

$$U(g)\langle\phi\rangle \neq \langle\phi\rangle \Rightarrow \delta\phi \neq 0, \tag{3.29}$$

and $g \in G$. Therefore, in this latter case, for (3.27) to hold we must have

$$\left(\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\right)\Bigg|_{\phi=\phi_0} = 0. \tag{3.30}$$

This shows that indeed there are massless particles arising from the fact that there is some symmetry group that maintains the vacuum invariant, and the number of massive particles is the dimension of the group H. The number of massless particles will then be given by the total number of particles minus those that do have mass. This is precisely the dimension of the coset G/H. Note that we must write it in terms of a coset since the elements of G that are not elements of H do not form a group, as the identity is in H.

3.2 Goldstone Theorem

To prove the Goldstone theorem in a more general way, we must understand how SSB works in terms of the states.

Recall that the symmetry operator is defined as

$$U = e^{i\alpha Q}. \tag{3.31}$$

If the vacuum is invariant under the symmetry,

$$U|0\rangle = |0\rangle \Rightarrow Q|0\rangle = 0. \tag{3.32}$$

But if the symmetry is spontaneously broken,

$$U|0\rangle = |0\rangle' \Rightarrow Q|0\rangle \neq 0. \tag{3.33}$$

Note that since U is a symmetry of the Hamiltonian, $[U, H] = 0$. This shows that there exist different states with the same energy. This is just what we have seen before, where the different degenerate states are linked by the symmetry transformation.

It is convenient to determine whether a symmetry is broken or not. To do that, we define a parameter called the *order parameter*, which computes precisely this. By construction, the order parameter is an operator that has zero expectation value in the symmetric state and a non-zero expectation value in the symmetry-broken phase. That is,

$$\langle\psi|\mathcal{O}(x)|\psi\rangle \begin{cases} = 0, & \text{symmetric state,} \\ \neq 0, & \text{symmetry-broken phase.} \end{cases} \tag{3.34}$$

The operator $\mathcal{O}(x)$ is defined as

$$\mathcal{O}(x) \equiv [Q, \phi'(x)], \tag{3.35}$$

with $\phi'(x)$ an arbitrary field called the *interpolating field.* It is possible to find $\phi'(x)$ such that

$$[Q, \phi'(x)] = \phi(x). \tag{3.36}$$

There is also a definition equivalent to Eq.(3.34) that states that if the symmetry is broken, then

$$\lim_{|x-y|\to\infty} \langle\mathcal{O}(x)\mathcal{O}(y)\rangle = \langle\mathcal{O}(x)\rangle\langle\mathcal{O}(y)\rangle \neq 0. \tag{3.37}$$

Then, by our definition, SSB happens when the field $\phi(x)$ has a non-vanishing VEV. Using (3.36) we have

$$\begin{aligned} &\langle 0|[Q, \phi'(x)]|0\rangle \neq 0, \\ &\langle 0|Q\phi'(x) - \phi'(x)Q|0\rangle \neq 0. \end{aligned} \tag{3.38}$$

This shows that, indeed, $Q|0\rangle = 0$ cannot hold in the case of SSB. Introducing a complete set of eigenstates of the Hamiltonian $|n, \boldsymbol{k}\rangle$ labeled by their energy E_n and momentum $\boldsymbol{k}$, respectively,

$$\langle 0|[Q, \phi']|0\rangle = \sum_n \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[\langle 0|Q|n, \boldsymbol{k}\rangle\langle n, \boldsymbol{k}|\phi'|0\rangle - \text{c.c.}\right], \tag{3.39}$$

where c.c. stands for the complex conjugate. Using

$$Q = \int_\Omega d^3x\, j^0(x, t) \tag{3.40}$$

with Ω a finite volume, we have

$$\begin{aligned}\langle 0|[Q, \phi']|0\rangle = \sum_n \int_\Omega d^3\boldsymbol{x} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} &[\langle 0|e^{-i(Ht-\boldsymbol{p}\cdot\boldsymbol{x})} j^0(0,0) e^{i(Ht-\boldsymbol{p}\cdot\boldsymbol{x})}|n, \boldsymbol{k}\rangle \\ &\times \langle n, \boldsymbol{k}|\phi'|0\rangle - \text{c.c.}],\end{aligned} \tag{3.41}$$

where we have used

$$j^0(x, t) = U(x, t) j^0(0, 0) U^\dagger(x, t). \tag{3.42}$$

Assuming that the vacuum is translationally invariant and applying the exponential operators to its eigenstates,

$$\langle 0|[Q, \phi']|0\rangle = \sum_n \int_\Omega d^3\boldsymbol{x} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} [\langle 0|j^0(0,0)|n, \boldsymbol{k}\rangle\langle n, \boldsymbol{k}|\phi'|0\rangle e^{i(E_n t - \boldsymbol{k}\cdot\boldsymbol{x})} - \text{c.c.}]. \tag{3.43}$$

Using

$$\int_\Omega d^3\boldsymbol{x}\; e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = (2\pi)^3 \delta^3(\boldsymbol{k}), \tag{3.44}$$

we obtain

$$\langle 0|[Q, \phi']|0\rangle = \sum_n \int d^3\boldsymbol{k}\, \delta^3(\boldsymbol{k}) [e^{iE_n t}\langle 0|j^0(0,0)|n, \boldsymbol{k}\rangle\langle n, \boldsymbol{k}|\phi'|0\rangle - \text{c.c.}]. \tag{3.45}$$

Therefore, these matrix elements must be non-zero. Both contain information about the Goldstone boson, but since j^0 is related to the symmetry that was spontaneously broken, it is reasonable to associate this with the Goldstone boson. Thus we understand that the conserved current creates a particle with momentum $\boldsymbol{k}$ from the vacuum. This result was obtained using specifically j^0, but in general we have

$$\text{Goldstone boson} \sim \langle 0|j_\mu|n, \boldsymbol{k}\rangle. \tag{3.46}$$

To see that these excitations are massless, we recall that Q is time independent. If ϕ is also time independent, taking the time derivative of (3.45) leads to

$$\partial_t(\langle 0|[Q, \phi']|0\rangle) = \sum_n \int d^3\boldsymbol{k}\; i\delta^3(\boldsymbol{k}) E_n (e^{iE_n t}\langle 0|j^0(0,0)|n, \boldsymbol{k}\rangle\langle n, \boldsymbol{k}|\phi'|0\rangle - \text{c.c.}) = 0. \tag{3.47}$$

This shows that as $k \to 0$ we must have $E_n \to 0$, which implies that these excitations are massless.

As a final comment, note that we have called these excitations Goldstone bosons, but how do we know that they are indeed bosons? The answer is that they arise from the SSB of a bosonic symmetry. There exists the possibility of having Goldstone fermions as well. This is related to the SSB of fermionic symmetries, which are generated by *Grassmann variables*, but we will not discuss this here.

3.3 The Higgs Mechanism

We can now turn to the SSB of gauge symmetries. As we have said before, this terminology is not completely well defined since gauge transformations are not physical symmetries. Nonetheless, we will see that the result of this SSB gives a very interesting outcome. We will start with the $U(1)$ symmetry group to gain insight into how things are done, and then we will move to the $SO(3)$ symmetry group, which will help us generalize the idea.

3.3.1 Symmetry breaking of U(1) gauge symmetry

Consider the $U(1)$ gauge transformation

$$\phi(x) \to e^{i\Lambda(x)}\phi. \tag{3.48}$$

For the Lagrangian to remain invariant we must introduce the covariant derivative,

$$\mathcal{L} = (\partial_\mu + ieA_\mu)\phi(\partial^\mu - ieA^\mu)\phi^* - m^2\phi\phi^* - \lambda(\phi\phi^*)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{3.49}$$

For $m^2 < 0$ the VEV is

$$|\phi| = \left(-\frac{m^2}{2\lambda}\right)^{1/2} = v. \tag{3.50}$$

Thus, we can again expand the field and reparametrize it around its vacuum,

$$\phi(x) = v + \frac{\phi_1(x) + i\phi_2(x)}{\sqrt{2}} = v + \frac{\phi_1}{\sqrt{2}} + i\frac{\phi_2}{\sqrt{2}}, \tag{3.51}$$

and substitute back into $\mathcal{L}$. Again, we only want to evaluate the mass terms, that is, the terms quadratic in the fields. We need the terms

$$\phi\phi^* = \left(v + \frac{\phi_1}{\sqrt{2}}\right)^2 + \frac{\phi_2^2}{2} = \frac{1}{2}\phi_1^2 + \frac{1}{2}\phi_2^2 + \sqrt{2}v\phi_1 + v^2 \tag{3.52}$$

$$(\phi\phi^*)^2 = \left(v^2 + \sqrt{2}v\phi_1\right)^2 + (v^2 + \sqrt{2}v\phi_1)(\phi_1^2 + \phi_2^2) \tag{3.53}$$

$$= 3v^2\phi_1^2 + v^2\phi_2^2. \tag{3.54}$$

We may now substitute these terms into the Lagrangian. As we did before, we will evaluate only the terms that are at most quadratic in the fields, since we are interested in the masses,

$$\begin{aligned}
\mathcal{L} &= \frac{1}{4}\partial_\mu(\phi_1 + i\phi_2)\partial^\mu(\phi_1 - i\phi_2) - \frac{ie}{2\sqrt{2}}(\partial_\mu\phi_1 + i\partial_\mu\phi_2)A^\mu\left(v + \frac{\phi_1 - i\phi_2}{\sqrt{2}}\right) \\
&+ \frac{ie}{2\sqrt{2}}A_\mu\left(v + \frac{\phi_1 + i\phi_2}{\sqrt{2}}\right)(\partial_\mu\phi_1 - i\partial_\mu\phi_2) + e^2A_\mu A^\mu|\phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 \\
&- \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
&= \frac{1}{4}(\partial_\mu\phi_1)^2 + \frac{1}{4}(\partial_\mu\phi_2)^2 - \sqrt{2}evA_\mu\partial^\mu\phi_2 + e^2v^2A_\mu A^\mu - 2\lambda v^2\phi_1^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
&+ \text{non-quadratic terms.}
\end{aligned} \tag{3.55}$$

We have collected all the quadratic terms, but there is one that is somewhat confusing. It is the mixing between the gauge field and the ϕ_2 component. To gain a better understanding of this term, recall that for infinitesimal transformations $\phi \to e^{i\Lambda(x)}\phi$,

$$\phi'(x) = \phi(x) + i\Lambda\phi(x).$$

In terms of the components,

$$v + \frac{1}{\sqrt{2}}(\phi_1' + i\phi_2') = v + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) + i\Lambda v + \frac{i\Lambda}{\sqrt{2}}(\phi_1 + i\phi_2).$$

Note that the expectation value of the field does not change. Thus, the components transform as

$$\phi_1' = \phi_1 - \Lambda\phi_2, \tag{3.56}$$

$$\phi_2' = \phi_2 + \Lambda\phi_1 + \sqrt{2}\Lambda v. \tag{3.57}$$

The components transform inhomogeneously. This shows that we can choose a gauge such that $\phi_2 = 0$. Doing that we obtain

$$\mathcal{L} = \frac{1}{4}(\partial_\mu\phi_1)^2 + e^2v^2A_\mu A^\mu - 2\lambda v^2\phi_1^2 - \frac{1}{4}(F_{\mu\nu})^2 + \text{non-quadratic terms.} \tag{3.58}$$

We now observe two interesting things. The first is that the scalar field ϕ_2 has completely vanished. This is different from the previous case, in which only the mass term for ϕ_2 vanished; here the field itself disappears. The second is that the gauge field has now acquired a mass; it is as if the photon has become massive.

The interpretation is that one component of the scalar field combines with the longitudinal part of the gauge field to produce a new massive degree of freedom. This is known as the *Higgs mechanism.* The counting of the degrees of freedom proceeds as follows

- **Symmetry breaking of global U(1) symmetry**
  2 massive scalar fields $\to$ 1 massive scalar field + 1 massless scalar field.

- **Symmetry breaking of U(1) gauge symmetry**
  2 massive scalar fields + 1 massless gauge field $\to$ 1 massive scalar field + 1 massive gauge field.

Note that the number of degrees of freedom before and after the symmetry breaking matches. In the case of the global symmetry we have already checked this. In the case of the gauge symmetry, before the symmetry breaking we had 2 degrees of freedom from the scalar field and 2 degrees of freedom from the massless gauge field. After the symmetry breaking we obtain 1 degree of freedom from the scalar field and 3 degrees of freedom from the massive gauge field.

Let us now study this within another symmetry group that will help us generalize the idea.

3.3.2 Symmetry breaking of $SO(3)$ gauge symmetry

Consider the Lagrangian (3.18) but in the presence of a gauge field,

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi_i)(D^\mu\phi_i) - \frac{m^2}{2}\phi_i\phi_i - \lambda(\phi_i\phi_i)^2 - \frac{1}{4}(F^i_{\mu\nu})^2, \tag{3.59}$$

with

$$D_\mu\phi_i = \partial_\mu\phi_i + g\epsilon_{ijk}A^j_\mu\phi_k, \tag{3.60}$$

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^j_\mu + g\epsilon^{ijk}A^j_\mu A^k_\nu. \tag{3.61}$$

The minimum of the potential is at

$$v = \left(-\frac{m^2}{4\lambda}\right)^{1/2},$$

and we choose

$$\langle\boldsymbol{\phi}\rangle = v\hat{\boldsymbol{e}}_{\boldsymbol{3}}. \tag{3.62}$$

The physical fields are then ϕ_1, ϕ_2 and $\chi = \phi_3 - v$. In terms of these fields, the Lagrangian becomes

$$\begin{aligned}\mathcal{L} = &\frac{1}{2}[(\partial_\mu\phi_1)^2 + (\partial_\mu\phi_2)^2 + (\partial_\mu\chi)^2] + gvA^\mu_2\partial_\mu\phi_1 - gvA^\mu_1\partial_\mu\phi_2 + \frac{g^2v^2}{2}[(A^\mu_1)^2 + (A^\mu_2)^2]\\ &- 4v^2\lambda\chi^2 - \frac{1}{4}(F^i_{\mu\nu})^2 + \text{non-quadratic terms.}\end{aligned} \tag{3.63}$$

We encounter the same problem here as before. We can perform a gauge transformation to set $\phi_1 = \phi_2 = 0$. We will then see that the remaining terms correspond to the mass terms that can already be identified in this Lagrangian. Therefore, we again obtain the Higgs mechanism, but in this case the field χ has acquired a mass $m^2_\chi = 4v^2\lambda$ and the components 1 and 2 of the gauge field have acquired masses $m^2_{A_1} = m^2_{A_2} = \frac{gv^2}{2}$. Summarizing

- **Symmetry breaking of global $SO(3)$ symmetry**
  3 massive scalar fields → 1 massive scalar field + 2 massless scalar fields.

- **Symmetry breaking of $SO(3)$ gauge symmetry**
  3 massive scalar fields + 3 massless gauge fields → 1 massive scalar field + 2 massive gauge fields + 1 massless gauge field.

The degrees of freedom are preserved as well. Before the symmetry breaking of the $SO(3)$ gauge symmetry we had 3 degrees of freedom from the scalar field and 6 degrees of freedom from the gauge fields (2 for each component). After the symmetry breaking, we have 1 degree of freedom from the scalar, 6 degrees of freedom from the massive gauge fields (3 for each component), and 2 degrees of freedom from the massless component of the gauge field.

This allows us to conclude, based on our previous discussion in terms of groups, that the number of degrees of freedom for an arbitrary group is given by

- Number of massless gauge fields: $\dim(H)$,
- Number of massive gauge fields: $\dim(G/H)$,
- Total number of gauge fields: $\dim(G)$.

4 TOPOLOGICAL FIELD THEORIES

There is a class of field theories whose degrees of freedom do not depend on the metric of the spacetime but rather on the topology. These are called *topological field theories (TFT's).* TFT's are essential to describe topological phase transitions and are a cutting-edge field of study.

In this chapter, we will start by studying one of the simplest, but richest, theory with topological degrees of freedom. This is called *Chern-Simons theory.* Chern Simons theory is of great importance because it is the Effective Field Theory (EFT) for the Quantum Hall Effect (QHE). We will then proceed to the *BF theory*, which similar to Chern-Simons, but with an increased number of degrees of freedom.

4.1 Quantum Hall Effect

The Quantum Hall Effect (QHE) is a remarkable physical phenomenon that arises under extreme conditions, namely low temperatures and strong magnetic fields, where the Hall conductance becomes quantized. The QHE's deep connection to topology is revealed through its effective field theory description, which is a gauge theory in (2+1) dimensions called the *Chern-Simons Theory,* which, as we will see, depends only on the underlying topology of the theory. The Chern-Simons theory is the first theory that we will study that has topological degrees of freedom. In this section, we give a brief overview of the QHE to motivate some physical arguments that we will use when discussing Chern-Simons theory. For a more detailed description, see [15].

4.1.1 Classical Hall Effect

The Hall Effect consists of a plane with a current I (let us suppose that it lies in the x-direction), and a magnetic field $\boldsymbol{B}$ applied in the z-direction. The magnetic field bends the current towards the y-direction, which produces a potential difference V_H, where H stands for Hall. The scheme can be seen in the figure below

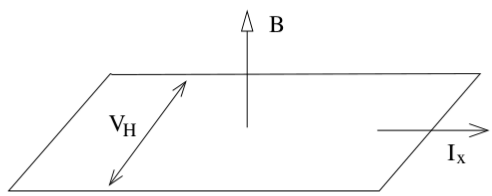

Figura 4.1 – Scheme for the Quantum Hall Effect. Taken from [15].

The electrons move in this plane in circular motion due to the magnetic field $\boldsymbol{B}$ with a frequency,

$$\omega_c = \frac{eB}{m}. \tag{4.1}$$

This is called the *cyclotron frequency.*

4.1.1.1 Drude Model

We can describe this effect classically using what we call the *Drude Model.* Suppose we add an electric field $\boldsymbol{E}$ which will be responsible for accelerating the electrons in the absence of the magnetic field. We may consider the collisions that the electrons have within the material with whatever may impede their motion. This was a first attempt to describe a conductor as if it were composed of billiard balls. The equation of motion of this system is given by Newton's second law

$$m\frac{d\boldsymbol{v}}{dt} = -e\boldsymbol{E} - e\boldsymbol{v} \times \boldsymbol{B} - \frac{m\boldsymbol{v}}{\tau}, \tag{4.2}$$

where τ is the average time between collisions of the particles. We want the stationary solutions, that is,

$$\frac{d\boldsymbol{v}}{dt} = 0.$$

So we must solve the equation

$$e\boldsymbol{E} + e\boldsymbol{v} \times \boldsymbol{B} + \frac{m\boldsymbol{v}}{\tau} = 0. \tag{4.3}$$

Recalling that the current density is related to the velocity through the equation,

$$\boldsymbol{J} = -ne\boldsymbol{v}. \tag{4.4}$$

In matrix notation

$$\begin{pmatrix} 1 & \omega_c\tau \\ -\omega_c\tau & 1 \end{pmatrix} \boldsymbol{J} = \frac{e^2 n\tau}{m}\boldsymbol{E}. \tag{4.5}$$

If we invert this relation,

$$\boldsymbol{J} = \sigma\boldsymbol{E}. \tag{4.6}$$

This is *Ohm's Law*, which states how the current behaves in the presence of an electric field. We define

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}, \tag{4.7}$$

as the *conductivity tensor*. This is something new that appears only in the presence of a magnetic field. Just as in the scalar case, the resistivity is the inverse of the conductivity; therefore

$$\rho = \sigma^{-1}. \tag{4.8}$$

More explicitly,

$$\rho = \frac{1}{\sigma_{DC}} \begin{pmatrix} 1 & \omega_c \tau \\ -\omega_c \tau & 1 \end{pmatrix}. \tag{4.9}$$

The off-diagonal terms are responsible for the Hall effect. It can be shown that

$$\rho_{xx} = \frac{m}{ne^2 \tau}, \quad \rho_{xy} = \frac{B}{ne}, \tag{4.10}$$

which, graphically, is represented as

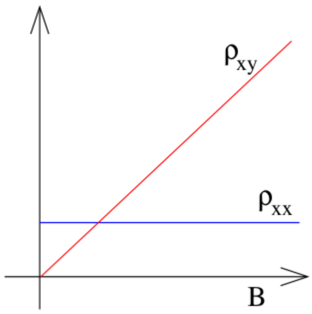

Figura 4.2 – Expected resistivities for the classical Hall effect. Taken from [15].

4.1.2 Integer Quantum Hall Effect

The Quantum Hall Effect (QHE) was first discovered experimentally in 1980 [1]. It consists of the same setup, but now in a regime where the magnetic field $\boldsymbol{B}$ is strong and the temperature is low. The graph found for the resistivities was the one shown below.

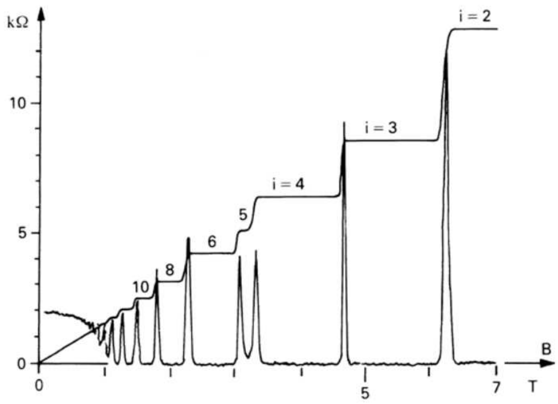

Figura 4.3 – Resistivities for the integer quantum Hall effect. Taken from [15].

This is completely different from what we expected classically. The plateaus (that is, the regions in which the graph is flat) represent the transverse resistivity ρ_{xy} and the spikes correspond to the longitudinal resistivity ρ_{xx}. This is interesting because the resistivity over the plateaus has the values

$$\rho_{xy} = \frac{2\pi}{e^2} \frac{1}{\nu}, \ \nu \in \mathbb{Z}, \tag{4.11}$$

where the integer ν is called the *filling fraction*. The center of these plateaus occurs when the magnetic field assumes the values

$$B = \frac{2\pi n}{\nu e} = \frac{n}{\nu}\Phi_0. \tag{4.12}$$

In contrast, the values of ρ_{xx} are zero most of the time, except when ρ_{xy} changes its value. Then it exhibits spikes.

What is fascinating about this effect is that a macroscopic quantity is quantized, rather than a microscopic one. Another interesting aspect is that, in components, the conductivity is

$$\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}, \quad \sigma_{xy} = \frac{-\rho_{xy}}{\rho_{xx}^2 + \rho_{xy}^2}. \tag{4.13}$$

If $\rho_{xy} = 0$ we have

$$\sigma_{xx} = \frac{1}{\rho_{xx}}. \tag{4.14}$$

This is a result that we are already familiar with, but in the case $\rho_{xx} = 0$ with $\rho_{xy} \neq 0$ we have

$$\sigma_{xx} = 0. \tag{4.15}$$

At first sight, this case suggests that our system is simultaneously a perfect conductor and a perfect insulator. But these are just names. What these expressions are really telling us is that, for $\sigma_{xx} = 0$, no current is flowing in the longitudinal direction (as in an insulator), and $\rho_{xx} = 0$ tells us that no energy is being dissipated (as in a conductor).

It is common to work with the QHE in terms of the conductivity,

$$\sigma_{xy} = \frac{e^2}{2\pi}\nu. \tag{4.16}$$

4.1.3 Fractional Quantum Hall Effect

The fractional QHE is even more exciting. The setup is the same as in the integer QHE, but we now consider interactions between the electrons in the plane. What was found experimentally for the resistivities was

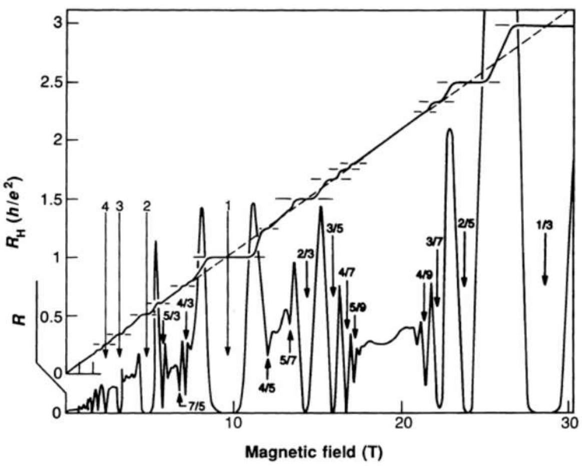

Figura 4.4 – Resistivities for the fractional quantum Hall effect. Taken from [15].

This shows that, in this case, the filling fractions may also assume *fractional* values. This is the reason for the name. We will dive deeper into the fractional QHE later when discussing its effective field theory.

4.2 Chern-Simons Theory

One way we can describe the Quantum Hall Effect and obtain the results we have discussed is by relying on *Chern-Simons Theory*. Chern-Simons theory is a gauge theory that is only well defined in odd dimensions. In our case, we will study it in $D = 3$[1].

It works this way because the new term that we add to the action, the *Chern-Simons* term, is

$$S_{CS} = \frac{k}{4\pi} \int d^3x \; \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho. \tag{4.17}$$

Note that indeed this theory does not work in 3+1 dimensions because of the Levi-Civita symbol: the indices do not match. The constant k is called the *level* of the Chern-Simons term.

At first glance, the Chern-Simons action does not seem gauge invariant because it depends explicitly on the field A_μ. But under a gauge transformation,

$$S_{CS} \to S_{CS} + \frac{k}{4\pi} \int d^3x \; \partial_\mu (\omega \epsilon^{\mu\nu\rho} \partial_\nu A_\rho).$$

Thus, it changes by a total derivative. For some topologies, we can simply discard this term. But in general, the boundary will have physical contributions. The Chern-Simons equations of motion are given by

$$F_{\mu\nu} = 0. \tag{4.18}$$

This shows that this theory has trivial dynamics. This happens because the Chern-Simons term is a purely *topological* term, in the sense that it does not depend on the spacetime metric. This can be seen explicitly if we write the action in terms of differential forms,

$$S_{CS} = \int A \wedge dA. \tag{4.19}$$

and compare it with the Maxwell action,

$$S_M = -\frac{1}{4e^2} \int F \wedge \star F. \tag{4.20}$$

To take the Hodge dual $\star F$ we need to use a metric, which contains information about the geometry of the space. But the Chern-Simons term, in contrast, does not involve any dual. Therefore, it depends only on the topology of the manifold.

In (2+1) dimensions, parity is defined as

$$x^0 \to x^0, \; x^1 \to -x^1, \; x^2 \to x^2,$$

[1] From now on, our convention is that D denotes the spacetime dimension and d denotes only the spatial dimension.

because if we were to define it just as in 4 dimensions, that is, $\boldsymbol{x} \to -\boldsymbol{x}$, since the matrix is 3-dimensional, it would have unit determinant and therefore correspond to a rotation. Therefore, we need to define it this way. Note that the axis chosen to be reflected is completely arbitrary, and it would work equally well if we chose x^2 to be reflected. Under these transformations, the gauge field transforms as

$$A_0 \to A_0,\ A_1 \to -A_1,\ A_2 \to A_2.$$

The integration measure $\int\ d^3x$ is invariant under parity because even if $x_1 \to -x_1$, the integration limits also change. But the integrand is not invariant, that is,

$$\epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho \to -\epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho.$$

Thus, the Chern-Simons term can only arise in theories that break parity.

We know that the conserved current couples to the gauge field as

$$S = \int d^3x\ J^\mu A_\mu.$$

This shows that

$$\frac{\delta S}{\delta A_\mu} = J^\mu.$$

This implies

$$J_i = -\frac{k}{2\pi}\epsilon_{ij} E_j. \tag{4.21}$$

From (4.6),

$$\boldsymbol{J} = \sigma \boldsymbol{E}.$$

Hence,

$$\sigma_{xy} = \frac{k}{2\pi}.$$

This coincides with the Hall conductivity if we identify

$$k = e^2 \nu. \tag{4.22}$$

But there is nothing that tells us that ν should be quantized. We will see how this works shortly.

4.2.1 Topological mass generation

Before we talk about the quantization of the Chern-Simons level, let us study the effect of adding a Chern-Simons term to other gauge theories. Consider the Maxwell-Chern-Simons Lagrangian

$$\mathcal{L}_{\text{MCS}} = -\frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + \frac{k}{4\pi}\epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho. \tag{4.23}$$

The equations of motion are

$$\partial_\mu F^{\mu\nu} + \frac{ke^2}{4\pi}\epsilon^{\nu\alpha\beta}F_{\alpha\beta} = 0. \tag{4.24}$$

Introducing the dual field $\tilde{F}^\mu$,

$$\tilde{F}^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho}F_{\nu\rho} \Rightarrow \tilde{F}_\mu = \frac{1}{2}\epsilon_{\mu\alpha\beta}F^{\alpha\beta}, \tag{4.25}$$

we can invert this relation by multiplying both sides by a Levi-Civita tensor, which gives

$$F^{\sigma\rho} = \epsilon^{\mu\sigma\rho}\tilde{F}_\mu. \tag{4.26}$$

Then we obtain

$$\epsilon^{\alpha\mu\nu}\partial_\mu\tilde{F}_\alpha + \frac{ke^2}{2\pi}\tilde{F}^\nu = 0.$$

This can be re-written as

$$\left(\partial_\mu\partial^\mu + \left(\frac{ke}{2\pi}\right)^2\right)\tilde{F}^\sigma = 0. \tag{4.27}$$

This is the equation of motion for a field with mass $m = \frac{ke}{4\pi}$.

Another way to see this is through the propagator. Recall from the LSZ reduction formula that, in order to compute the S-matrix elements, we need to evaluate the following expression:

$$\begin{aligned}\langle p_1 \dots p_n|S|q_1 \dots q_n\rangle = \prod_{r=1}^{n} \lim_{p_r^2\to m_r^2}(p_r^2 - m_r^2)J_{i_r}(p_r) \prod_{s=1}^{m} \lim_{p_s^2\to m_s^2}(q_s^2 - m_s^2)J_{i_s}(p_s) \\ \times G_{i_1\dots j_m}(p_1, \dots, p_n, -q_1, \dots, -q_m),\end{aligned} \tag{4.28}$$

where $G_{i_1\dots j_m}(p_1, \dots, p_n, q_1, \dots, q_m)$ is the propagator. Since propagators are often of the form $\frac{1}{p^2-m^2}$, the poles are interpreted as the masses of the field.

Rewriting the Lagrangian as

$$\mathcal{L}_{\text{MCS}} = A_\mu\left[\frac{1}{2e^2}g^{\mu\nu}\Box - \frac{k}{4\pi}\epsilon^{\mu\rho\nu}\partial_\rho - \frac{1}{2e^2}\left(\frac{1}{\xi} - 1\right)\partial^\mu\partial^\nu\right]A_v,$$

with a gauge-fixing term $-\frac{1}{2\xi e^2}(\partial_\mu A^\mu)^2$, we identify the term between brackets as the inverse of the propagator,

$$(\Delta^{-1})^{\mu\nu} = \frac{1}{2e^2}g^{\mu\nu}\Box - \frac{k}{4\pi}\epsilon^{\mu\rho\nu}\partial_\rho - \frac{1}{2e^2}\left(\frac{1}{\xi} - 1\right)\partial^\mu\partial^\nu. \tag{4.29}$$

In momentum space this is written as

$$(\Delta^{-1}(p))^{\mu\nu} = \frac{1}{2}g^{\mu\nu}p^2 + \frac{ik}{4\pi}\epsilon^{\mu\nu\rho}p_\rho + \frac{1}{2}\left(1 - \frac{1}{\xi}\right)p^\mu p^\nu. \tag{4.30}$$

To compute the propagator, we first introduce the following operators:

$$\omega^{\mu\nu} \equiv \frac{p^\mu p^\nu}{p^2},\ S^{\mu\nu} \equiv i\epsilon^{\mu\nu\rho}p_\rho,\ \theta^{\mu\nu} = g^{\mu\nu} - \omega^{\mu\nu}. \tag{4.31}$$

These form a closed algebra:

$$\omega^2 = 0, \;\; \omega \cdot \theta = 0, \;\; \theta^2 = 0, \;\; \omega \cdot S = 0, \;\; S^2 = -p^2 \cdot \theta, \;\; \theta \cdot S = S. \tag{4.32}$$

The propagator in this basis should have the form

$$\Delta_{\mu\nu}(p) = A\omega_{\mu\nu} + B\theta_{\mu\nu} + CS_{\mu\nu}, \tag{4.33}$$

with A, B, C complex numbers. Thus, we need to solve the equation

$$\Delta_{\mu\alpha}(\Delta^{\alpha\nu})^{-1} = \delta_{\mu}{}^{\nu}, \tag{4.34}$$

for A, B and C. The propagator is then given by

$$\Delta_{\mu\nu} = \left(\frac{p^2 g_{\mu\nu} - p_\mu p_\nu - \frac{ike^2}{2\pi}\epsilon_{\mu\nu\rho}p^\rho}{p^2(p^2 - (\frac{ke^2}{2\pi})^2)} + \xi \frac{p_\mu p_\nu}{p^4} \right). \tag{4.35}$$

Thus, we again obtain that the mass of the gauge field is $\left(\frac{ke}{2\pi}\right)^2$. Note that it seems that we also have a pole at $p^2 = 0$, but that pole is actually "fake". This can be seen from Eq.(4.28). The conserved currents can be written in a basis that depends on the momentum p and the polarization vectors ϵ. Therefore, products of the form $J\Delta J$ either vanish or cancel the p^2 in the denominator. The only case where this does not follow straightforwardly is the term with the Levi-Civita symbol. In this case, the conservation law $p_\mu j^\mu = 0$ imposes a constraint on the components, which cancel each other.

This shows that the coupling with the Chern-Simons term works as a kind of "Higgs mechanism", giving a mass to the gauge field. We call this a "topological mass" due to the nature of the Chern-Simons theory.

We may also include the Higgs mechanism. This is described by the Maxwell-Chern-Simons-Higgs Lagrangian

$$\mathcal{L}_{\text{MCSH}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho + (\mathcal{D}_\mu\phi)^*\mathcal{D}^\mu\phi - V(|\phi|), \tag{4.36}$$

where $V(|\phi|)$ is a symmetry-breaking potential with a non-trivial minimum $\langle\phi\rangle = v$. As we did in Chapter 3, we may parametrize the field as $\phi = \rho e^{i\varphi}$, leading to

$$\mathcal{L}_{MCSH} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho + \partial_\mu\rho\partial^\mu\rho + \rho^2(eA_\mu + \partial_\mu\varphi)^2 - V(\rho). \tag{4.37}$$

Expanding the field around the minimum as $\rho = \tilde{\rho} + v$, we obtain

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho + \partial_\mu\tilde{\rho}\partial^\mu\tilde{\rho} + v^2(eA_\mu + \partial_\mu\varphi)^2 - V(|\phi|) \\ &+ \text{cubic terms} + \text{quartic terms}.\end{aligned} \tag{4.38}$$

To compute the propagator for A_μ, we only need the quadratic terms involving A_μ. We can choose a gauge-fixing term that eliminates the field φ,

$$A_\mu \to -A_\mu + \frac{1}{e^2}\partial_\mu\varphi. \tag{4.39}$$

If we use the same algorithm that we used for the Maxwell-Chern-Simons Lagrangian, the propagator for the gauge field will be

$$\begin{aligned}\Delta_{\mu\nu} = &\frac{e^2(p^2 - m_H^2)}{(p^2 - m_+^2)(p^2 - m_-^2)}\left[g_{\mu\nu} - \frac{p_\mu p_\nu}{(p^2 - \xi m_H^2)} - \frac{i}{2\pi}\frac{ke^2\epsilon_{\mu\nu\rho}p^\rho}{(p^2 - m_H^2)}\right] \\ &+ e^2\xi\frac{p_\mu p_\nu(p^2 - (\frac{ke^2}{2\pi})^2 - m_H^2)}{(p^2 - m_+^2)(p^2 - m_-^2)(p^2 - \xi m_H^2)},\end{aligned} \tag{4.40}$$

where $m_H^2 = 2e^2v^2$ is the Higgs mass, i.e., the mass generated by the Higgs mechanism, and

$$m_\pm^2 = m_H^2 + \frac{(ke^2)^2}{8\pi^2} \pm \frac{ke^2}{4\pi}\sqrt{\left(\frac{ke^2}{2\pi}\right)^2 + 4m_H^2}, \tag{4.41}$$

are the topological masses of the gauge field.

Thus, we have the following situation. In the unbroken vacuum, the complex scalar field has two massive degrees of freedom and the gauge field has one massive degree of freedom arising from the Chern-Simons term. In the broken vacuum, one component of the scalar field combines with the longitudinal degree of freedom of the gauge field, resulting in one massive degree of freedom for the scalar field and two massive degrees of freedom for the gauge field.

Therefore, the photon acquires a topological mass due to the Chern-Simons term and a Higgs mass due to the Higgs mechanism. Note that we can also work with a Chern-Simons-Higgs Lagrangian if we take the limit

$$e^2 \to \infty, \quad k = \text{fixed}. \tag{4.42}$$

In this case, the Lagrangian becomes

$$\mathcal{L}_{\text{CSH}} = \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_v A_\rho + (\mathcal{D}_\mu\phi)^*\mathcal{D}^\mu\phi - V(|\phi|). \tag{4.43}$$

To compute the degrees of freedom, we can take this limit in the Maxwell-Chern-Simons-Higgs propagator, leading to

$$\Delta_{\mu\nu} = \frac{1}{p^2 - \left(\frac{4\pi v^2}{k}\right)^2}\left[\frac{4\pi v^2}{k}g_{\mu\nu} - \frac{p_\mu p_\nu}{2v^2} + \frac{i}{k}\epsilon_{\mu\nu\rho}p^\rho\right], \tag{4.44}$$

where we identify a mass pole at $p^2 = \left(\frac{4\pi v^2}{k}\right)^2$.

Thus, in the unbroken vacuum the gauge field is non-propagating and the scalar field has two massive degrees of freedom. In the broken vacuum, one of the components of the scalar field combines with the longitudinal component of the gauge field, giving it a mass $\frac{4\pi v^2}{k}$. Note that we can rewrite $m_\pm$ as

$$m_\pm = \frac{ke^2}{4\pi}\left(\sqrt{1 + \frac{32\pi^2 v^2}{(ke^2)^2}} \pm 1\right). \tag{4.45}$$

Taking the limit $e^2 \to \infty$,

$$m_\pm \sim \frac{ke^2}{4\pi}\left(1 + \frac{16\pi^2 v^2}{(ke^2)^2} \pm 1\right) = \frac{ke^2}{4\pi} + \frac{4v^2}{k\pi} \pm \frac{ke^2}{4\pi}. \tag{4.46}$$

Thus, we have $m_+ \to \infty$ and $m_- \to 2v^2/2$, highlighting that the gauge field indeed has only one massive degree of freedom.

4.2.2 Quantization of the Chern-Simons Level

We have seen above that the Chern-Simons theory describes the correct expression for the Hall conductivity, but there was nothing telling us that the filling factor ν should be quantized. We now proceed to show this.

The quantization of the Chern-Simons level is not easy to perform given the fact that the Chern-Simons theory is a constrained theory and therefore should be quantized using Dirac brackets [16]. Nonetheless, we can obtain the results for the QHE without having to enter this heavy mathematical formalism. In order to do this, we must first perform a very common technique in field theory called a *Wick rotation,* which enables us to connect statistical mechanics with the path integral formalism.

For simplicity, we will do this for a system consisting of a single particle in a potential $V(q)$. The partition function for this system is

$$Z[\beta] = \mathrm{Tr}\, e^{-\beta H}, \tag{4.47}$$

where $\beta = 1/T$ and H is the Hamiltonian. Recall from the path integral formalism that a probability amplitude is given by[2]

$$\langle q_f | e^{-iHt} | q_i \rangle = \int_{q(0)=q_i}^{q(t)=q_f} \mathcal{D}q\; e^{iS}. \tag{4.48}$$

We note that these two objects have a similar form. To get rid of the factor of i, we can change the variable t to

$$\tau = it, \tag{4.49}$$

which is what we call a Wick rotation. Hence, the action of the particle changes as

$$S = \int_0^{-i\tau} \frac{dt}{d\tau'} d\tau'\; L(\dot{q}, q) = -i \int_0^{-i\tau} d\tau' \left[-\frac{m}{2}\left(\frac{dq}{d\tau}\right)^2 - V(q) \right] \equiv -S_E, \tag{4.50}$$

where S_E is called the *Euclidean action*. We will use this notation only in this section to make the distinction clear, but in what follows we will no longer use it. It can be seen that the Euclidean action is being used when there are no factors of i accompanying the action.

[2] For the reader who is unfamiliar with the path integral formalism, a nice discussion can be found in [13].

Now, to obtain the β factor we want, we can say that the particle propagates during a Euclidean time $\tau = \beta$. Then the amplitude becomes

$$\langle q_f | e^{-H\beta} | q_i \rangle = \int_{q(0)=q_i}^{q(\beta)=q_f} \mathcal{D}q \, e^{-S_E}.$$

To obtain the trace, we must recall that the trace is a sum over the *same* states. In this amplitude, we have different states $|q_i\rangle$ and $|q_f\rangle$. Therefore, we must demand that the particle return to its initial position. We do this by making the time periodic,

$$\tau \sim \tau + \beta.$$

Therefore, the partition function becomes

$$Z[\beta] = \int \mathcal{D}q \, e^{-S}, \tag{4.51}$$

where we have dropped the S_E notation. Note that this mapping implies that our time, the Euclidean time, has become an S^1. This must be taken into account in our further calculations. We can now proceed to the quantization.

For simplicity, consider our space to be an S^2. Therefore, our spacetime is the manifold $\mathcal{M} = S^1 \times S^2$. There is another simplification in our calculations. Suppose we have a magnetic monopole at the origin with flux

$$\int_{S^2} d\boldsymbol{a} \cdot \boldsymbol{B} = g, \tag{4.52}$$

where g is quantized via the Dirac condition from Chapter 2,

$$eg = 2\pi n, \quad n \in \mathbb{Z}. \tag{4.53}$$

This may seem biased, since magnetic monopoles do not exist (yet). However, what we are assuming is that this condition should hold even in the presence of monopoles. A different and more mathematical approach, using wavefunctionals, can be found in [17–19].

The Dirac condition in terms of the field strength (considering $g = 1$) is

$$\int_{S^2} F_{12} = \frac{2\pi}{e}. \tag{4.54}$$

The gauge transformation is, as usual,

$$A_\mu \to A_\mu + \partial_\mu \Lambda.$$

Because of the periodicity of our time, we can choose a gauge parameter that is periodic in time as well,

$$\Lambda = \frac{2\pi\tau}{e\beta}. \tag{4.55}$$

In this way, the gauge transformation for A_0 is given by

$$A_0' = A_0 + \frac{2\pi}{e\beta}. \tag{4.56}$$

These are called *large gauge transformations* and are related to gauge transformations that are periodic and have a non-trivial winding number. As we have seen, under gauge transformations the action changes as

$$\begin{aligned} S_{CS} \to S_{CS} &+ \frac{k}{4\pi}\int d^3x\ \partial_\mu(\omega\epsilon^{\mu\nu\rho}\partial_\nu A_\rho) \\ &= S_{CS} + \frac{2\pi k}{e^2}, \end{aligned} \tag{4.57}$$

which highlights the fact that the action is not invariant under large gauge transformations. However, in QFT the partition function is a more fundamental object than the action. Therefore, it is not a problem that the action changes under large gauge transformations, as long as the partition function does not change. From (4.51) we have

$$Z[A_\mu] = e^{iS[A_\mu]}. \tag{4.58}$$

Thus, we see that it is invariant if

$$\frac{k}{e^2} \in \mathbb{Z}. \tag{4.59}$$

But from Eq. (4.22),

$$k = e^2\nu.$$

Therefore, ν is indeed an integer. Hence, we have successfully shown that the Chern-Simons level is quantized and that it describes the correct conductivities of the integer QHE. We now proceed to describe the fractional QHE.

4.2.3 Effective field theory for the fractional QHE

The fractional quantum Hall effect can be described using Chern-Simons theory as well. The novelty is that we now take into account the interaction between the electrons.

We start by introducing an emergent $U(1)$ gauge field a_μ, associated with the collective motion of the underlying electrons. This is different from the gauge field A_μ from Maxwell theory. The simplest term that we can add to the action involving this field is the Chern-Simons term,

$$S_{CS} = \frac{k}{4\pi}\int d^3x\ \epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho. \tag{4.60}$$

Despite the fact that the field a_μ is not the Maxwell field, it follows the same properties that we have seen before. It has its own dynamics given by

$$S[a_\mu] = -\frac{1}{4}\int d^3x\ f_{\mu\nu}f^{\mu\nu}, \tag{4.61}$$

and it works as a topological mass generator as well.

The states that describe the fractional QHE are the Laughlin states [20]. We will now write down an action that is able to describe those states and therefore reproduce the expected values for the Hall conductivity. First, we need a coupling between the fields A_μ and a_μ. In general, the coupling involving gauge fields is through a conserved current. Luckily, we have a perfectly conserved current to couple with. It is

$$j^\mu = \frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu a_\rho, \tag{4.62}$$

where its conservation results from the contraction of the two derivatives with the Levi-Civita symbol. We can interpret this as if the magnetic flux of a_μ is the electric charge that couples to A_μ. This magnetic flux also follows the Dirac quantization condition

$$\frac{1}{2\pi}\int_{S^2} f_{12} = \frac{\hbar}{e}. \tag{4.63}$$

The normalization ensures that the minimum allowed charge is $\int J^0 = e$, as it should be. The effective action is then

$$S_{eff}[A;a] = \int d^3x\ \frac{1}{2\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + \ldots\ . \tag{4.64}$$

The first term is the coupling and the second is the Chern-Simons term for a_μ. We have used m instead of k because it is a different term (although it has the same functional form) from the Chern-Simons term for the integer QHE. The other terms that may arise vanish at large distances and therefore we do not need them for our conclusions. There could also be a Chern-Simons term for the A_μ field, but this is just the term we had in the integer QHE and we would simply obtain again the conclusion that the level k should be quantized. For the same reasons that k is quantized, m is also quantized. Computing the equation of motion for a_μ, we have

$$f_{\mu\nu} = \frac{1}{m}F_{\mu\nu}. \tag{4.65}$$

This has the solution

$$a_\mu = \frac{1}{m}A_\mu. \tag{4.66}$$

Substituting back into the action,

$$S_{eff}[A] = \int d^3x\ \frac{1}{4\pi m}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho. \tag{4.67}$$

This term is again a Chern-Simons term but with level $1/m$. Thus, following the same procedure, we obtain the Hall conductivity to be

$$\sigma_{xy} = \frac{e^2}{2\pi}\frac{1}{m}. \tag{4.68}$$

This shows that indeed the Hall conductivity in the fractional QHE may assume fractional values as well, with the filling factor being $\nu = 1/m$. There are some subtleties in this derivation because the effective action has some issues when integrating over an S^2. Moreover, the fields are constrained by the Dirac quantization condition, and the equation of motion is not satisfied when both fields have a single unit of flux. However, these subtleties lie in the manifold structure of our space and are related to the choice of charts, something that we will not explore further.

4.2.4 Charged excitations

In Chern-Simons theory we may have two types of charged excitations: the *quasi-holes* and the *quasi-particles*. In what follows, we will derive a remarkable property of these excitations, which is verified in the QHE. We will not make any distinction in this section, and every excitation will be called a quasi-particle.

Take the Chern-Simons action and add a term for the current of the quasi-particles $\tilde{j}^\mu$,

$$S[a;A] = \int d^3x \frac{1}{2\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + a_\mu\tilde{j}^\mu + \dots. \tag{4.69}$$

Since A_μ is a background field, we can set it to zero. The equations of motion are then given by

$$\frac{1}{2\pi}f_{\mu\nu} = \frac{1}{k}\epsilon_{\mu\nu\rho}\tilde{j}^\rho. \tag{4.70}$$

The simplest charge we can consider is a static charge at the origin. It has the components

$$\tilde{j}^1 = \tilde{j}^2 = 0,\ \tilde{j}^0 = e\delta^2(x). \tag{4.71}$$

The choice of the charge of these particles is due to Dirac quantization. Therefore, the equations of motion become

$$\frac{1}{2\pi}f_{12} = \frac{e}{k}\delta^2(x). \tag{4.72}$$

This shows that the role of the Chern-Simons term is to attach a magnetic flux e/k to each particle of charge e.

From Eq. (4.62), we note that the charge of these particles is given by the zeroth component. Thus,

$$j^0 = \frac{1}{2\pi}\epsilon^{ij}\partial_i a_j = \frac{1}{2\pi}f_{12} = \frac{e}{k}\delta^2(x). \tag{4.73}$$

This shows that the electron charge in Chern-Simons theory is fractional. Therefore, the charged excitations in Chern-Simons theory carry a fraction of the electron charge, with a magnetic flux attached to them. Below, we show schematically the quasi-particles in Chern-Simons theory.

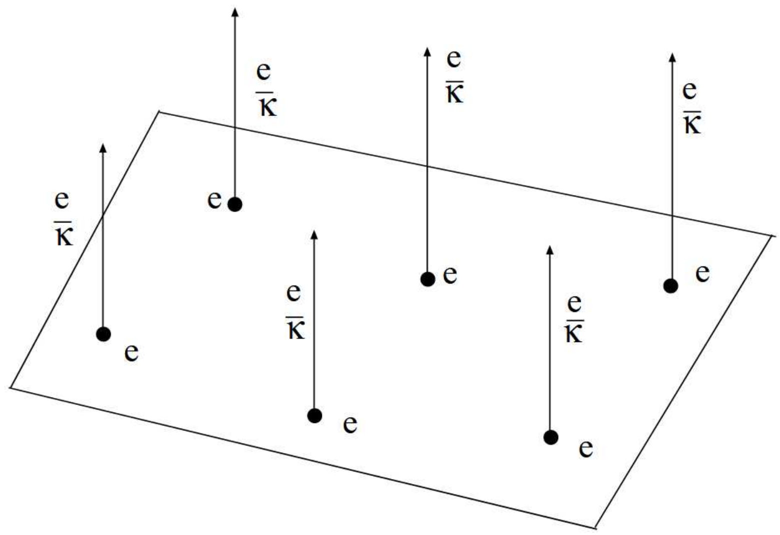

Figura 4.5 – Quasi-particles in Chern-Simons theory. Taken from [17].

Consider now a more general current,

$$j^0(\boldsymbol{x},t) = e\sum_{a=1}^{N}\delta^2(\boldsymbol{x}-\boldsymbol{x}_a(t)),\ \boldsymbol{j}(\boldsymbol{x},t) = e\sum_{a=1}^{N}\dot{\boldsymbol{x}}_a(t)\delta^2(\boldsymbol{x}-\boldsymbol{x}_a(t)). \tag{4.74}$$

Imposing $\partial_i a_i = 0$, we conclude that

$$a_i = \epsilon_{ijk}\partial_j\chi_k. \tag{4.75}$$

Solving the two-dimensional Green's function in the gauge $a_0 = 0$, we find

$$a_i(\boldsymbol{x},t) = \frac{e}{k}\sum_{a=1}^{N}\epsilon^{ij}\frac{x^j - x_a^j}{|\boldsymbol{x}-\boldsymbol{x}_a(t)|^2}. \tag{4.76}$$

Using the identity

$$\partial_i\mathrm{arg}(\boldsymbol{x}) = -\epsilon_{ij}\frac{x^j}{|\boldsymbol{x}|^2},\quad \mathrm{arg}(\boldsymbol{x}) = -\arctan\left(\frac{y}{x}\right), \tag{4.77}$$

we obtain

$$a_i(\boldsymbol{x}) = \frac{e}{k}\sum_a \partial_i\,\mathrm{arg}(\boldsymbol{x}-\boldsymbol{x}_a). \tag{4.78}$$

One might think that this field is a pure gauge and that we could transform it to zero by choosing

$$a_i' = a_i + \partial_i\Lambda,\quad \Lambda = -\frac{e}{k}\sum_a \mathrm{arg}(\boldsymbol{x}-\boldsymbol{x}_a). \tag{4.79}$$

However, under gauge transformations the wavefunction of the electrons changes as

$$\psi \to e^{ie\Lambda}\psi = \exp\left(-\frac{e^2}{k}\sum_a \mathrm{arg}(\boldsymbol{x}-\boldsymbol{x}_a)\right)\psi. \tag{4.80}$$

Thus, we see that this wavefunction is not single-valued and therefore the phase cannot be ignored.

In fact, it is this non-trivial residual phase that gives a very interesting aspect of this theory. Consider moving a particle around another. Since these particles have magnetic flux attached to them, upon returning to the same point they acquire an Aharonov-Bohm phase $\exp\left(ie\oint_C \boldsymbol{A}\cdot d\boldsymbol{x}\right)$. Hence

$$\exp\left(\frac{i\hbar}{k}\sum_{a=1}^{N}\oint \partial_i \arg(\boldsymbol{x}-\boldsymbol{x}_a)dx^i\right) = e^{\frac{2\pi i}{k}}. \tag{4.81}$$

Exchanging two particles is equivalent to performing half of a full rotation followed by a translation. This gives the phase $e^{\frac{\pi i}{k}}$. Therefore, we conclude that the exchange phase is fractional. Braiding is a way of measuring the statistics of two particles. Since this phase is neither an integer multiple of 2π (which would correspond to bosons) nor a half-integer multiple (which would correspond to fermions), we call these particles with unusual statistics *anyons.*

4.3 BF Theory

4.3.1 BF theory from Higgs model

We now proceed to show how to obtain BF theory starting from the usual Abelian Higgs model discussed previously. The Lagrangian is

$$\mathcal{L} = |(\partial_\mu + iNA_\mu)\phi|^2 - V(|\phi|) - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu, \tag{4.82}$$

where the field ϕ has charge N. In the Higgs phase we obtain

$$\mathcal{L} = v^2(\partial_\mu\varphi + NA_\mu)^2 - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu. \tag{4.83}$$

One might think of integrating out the field φ to obtain an effective Lagrangian, but we cannot perform a straightforward integration since ϕ admits vortex solutions. This implies that φ has non-trivial winding. First, we need to split the phase φ as

$$\varphi = \tilde{\varphi} + \eta, \tag{4.84}$$

where $\tilde{\varphi}$ carries information about the winding and the vortex positions, and η is a smooth function. We then introduce the dual field ξ_μ with the Lagrangian

$$\mathcal{L} = -\frac{1}{2v^2}\xi_\mu^2 + \xi_\mu(\partial^\mu\varphi - NA^\mu) - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu. \tag{4.85}$$

Note that computing the equations of motion for ξ_μ and substituting back into this Lagrangian leads us back to (4.83). Now we use (4.84),

$$\mathcal{L} = -\frac{1}{2v^2}\xi_\mu^2 + \xi_\mu(\partial_\mu\tilde{\varphi} + \partial_\mu\eta - NA_\mu) - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu. \tag{4.86}$$

When we integrate over η using a Gaussian integral, we obtain the constraint $\partial_\mu \xi^\mu = 0$. We may then write

$$\xi^\mu = \frac{iN}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu b_\rho, \tag{4.87}$$

for some vector field b_ρ. The factor $iN/2\pi$ will be explained below. Then we get

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2v^2}\epsilon^{\mu\nu\rho}\epsilon_{\mu\alpha\beta}\partial_\nu b_\rho \partial_\alpha b_\beta + \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}\partial^\nu b^\rho(\partial^\mu\tilde{\varphi} - NA^\mu) - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu \\ &= \frac{1}{2v^2}(f_{\mu\nu}^b)^2 + \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}\partial^\nu b^\rho(\partial^\mu\tilde{\varphi} - NA^\mu) - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu,\end{aligned} \tag{4.88}$$

where $f_{\mu\nu}^b = \partial_\mu b_\nu - \partial_\nu b_\mu$. Integrating the term $\epsilon_{\mu\nu\rho}\partial^\nu b^\rho \partial^\mu \tilde{\varphi}$ by parts, we obtain

$$\mathcal{L} = \frac{1}{2v^2}(f_{\mu\nu}^b)^2 - \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}b^\rho\partial^\nu\partial^\mu\tilde{\varphi} - \frac{iN^2}{2\pi}\epsilon_{\mu\nu\rho}\partial^\nu b^\rho A^\mu - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu. \tag{4.89}$$

It may appear that the second term is identically zero, but this is not true. The derivatives do not commute due to the non-trivial winding of $\tilde{\varphi}$. Note that the quantity that couples to b_0 is $\frac{N}{2\pi}\epsilon_{ij}\partial_i\partial_j\tilde{\varphi} = \frac{N}{2\pi}\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\tilde{\varphi})$. Integrating this over a region that contains the vortex leads to

$$\frac{N}{2\pi}\int d^2x\; \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\tilde{\varphi}) = \frac{N}{2\pi}\underbrace{\oint d\boldsymbol{x}\cdot\nabla\tilde{\varphi}}_{2\pi} = N. \tag{4.90}$$

This implies that this term is the density of vortices, that is, the zeroth component of some vortex current $\tilde{j}^\mu$,

$$\tilde{j}_\mu = \frac{N}{2\pi}\epsilon_{\mu\nu\rho}\partial_\nu\partial_\rho\tilde{\varphi}, \tag{4.91}$$

where the factor $\frac{N}{2\pi}$ ensures that we obtain the total charge when integrating over all space. We thus end up with the Lagrangian

$$\mathcal{L} = \frac{1}{2v^2}(f_{\mu\nu}^b)^2 - \tilde{j}_\mu b^\mu - \frac{iN^2}{2\pi}\epsilon_{\mu\nu\rho}\partial^\nu b^\rho A^\mu - \frac{1}{4}F_{\mu\nu}^2 - NA_\mu j^\mu. \tag{4.92}$$

From (4.83), note that if we perform the gauge transformation

$$A_\mu \to A_\mu - \frac{1}{N}\partial_\mu\eta, \tag{4.93}$$

then at low energies, when $v^2 \to \infty$, we have

$$A_\mu \to -\frac{1}{N}a_\mu, \tag{4.94}$$

where $a_\mu = \partial_\mu\tilde{\varphi}$. Hence, we obtain the BF Lagrangian

$$\mathcal{L} = \frac{1}{2v^2}(f_{\mu\nu}^b)^2 - \frac{1}{4}(f_{\mu\nu}^a)^2 - b^\mu\tilde{j}_\mu + a_\mu j^\mu + \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}\partial^\nu b^\rho a^\mu. \tag{4.95}$$

5 HIGHER-FORM SYMMETRIES

Higher-form symmetries generalize the conventional global symmetries by acting on extended operators, such as line or surface operators, rather than local, point-like ones. These ideas were considered first in [3], but the foundational work was [2]. These symmetries are characterized by topological symmetry operators, and their action is defined via the linking between the manifold supporting the symmetry operator and the manifold supporting the charged operator. A key consequence is that the associated conserved currents become higher-degree differential forms, in contrast to the standard 1-form currents of ordinary symmetries.

To build intuition, we will first revisit the topological interpretation of ordinary symmetry operators before extending these ideas to the higher-form case. Here we follow the discussion of [5, 21–23].

5.1 0-form symmetries

We have seen in Chapter 1 that another way to derive Noether's theorem is to make the parameter of the transformation, say α, local. That is,

$$\alpha \to \alpha(x).$$

This leads to

$$\delta S = -\int d^4x \; \alpha(x) \partial_\mu j^\mu. \tag{5.1}$$

Since α is arbitrary, if the equations of motion are satisfied, we have

$$\partial_\mu j^\mu = 0. \tag{5.2}$$

where j^μ is precisely the conserved current. Note that the conserved current can be seen as a 1-form, and it originates from the symmetry parameter α, which can be seen as a 0-form. Therefore, we now call the "ordinary" symmetries *0-form symmetries*, because the parameter of the transformation is a closed 0-form.

In terms of differential forms, the conservation law is written, as we have seen in Chapter 2,

$$d \star j = 0, \tag{5.3}$$

where j is the current 1-form and $\star : \Lambda^p \to \Lambda^{D-p}$ is the Hodge star map that takes a p-form and associates it with a $(D - p)$-form. The conserved charge is analogous to the one we have seen in Chapter 1,

$$Q \equiv \int d^{D-1}x \; j^0. \tag{5.4}$$

In terms of differential forms, we write it as

$$Q = \int_{\Sigma_{D-1}} \star j, \tag{5.5}$$

where Σ_{D-1} is a $(D-1)$-dimensional manifold. Note that now the manifold on which the charge is defined is not necessarily a spatial manifold. We then say that these charges are integrated over a codimension-1 manifold, which is a submanifold of the D-dimensional manifold. To understand how these operators act on the fields, we will make use of the *Ward identity*[1],

$$\partial_\mu \langle j^\mu(x)\phi(y)\rangle = -i\delta^{(D)}(x-y)\langle\delta\phi(y)\rangle. \tag{5.6}$$

Integrating over a D-dimensional manifold Ω with boundary Σ_{D-1}, the left-hand side becomes

$$\int_\Omega d\langle\star j\phi(y)\rangle = \int_{\Sigma_{D-1}} \langle\star j\phi(y)\rangle = \langle Q(\Sigma_{D-1})\phi(y)\rangle, \tag{5.7}$$

where we have used Stokes' theorem. For the right-hand side, we have

$$-i\int_\Omega d^Dx\ \delta^{(D)}(x-y)\langle\delta\phi(y)\rangle, \tag{5.8}$$

where we identify

$$\int_\Omega d^Dx\ \delta^{(D)}(x-y) = \begin{cases} 1, & y\in\Omega \\ 0, & y\notin\Omega \end{cases} \tag{5.9}$$

as the *linking* between the point y and the manifold Ω. Thus, we define

$$\int_\Omega d^Dx\ \delta^{(D)}(x-y) \equiv \mathrm{Link}(\Sigma_{D-1}, y). \tag{5.10}$$

Hence,

$$\langle Q(\Sigma_{D-1})\phi(y)\rangle = -i\mathrm{Link}(\Sigma_{D-1}, y)\langle\delta\phi(y)\rangle. \tag{5.11}$$

This expression tells us that the charge acts on the local field, and it only has a non-vanishing action if the point where the local operator is defined lies within the manifold on which the charge is defined. We can see that the link is invariant under a deformation of the manifold of the type $\Omega' = \Omega + \Omega_0$, with $\partial\Omega' = \Sigma'$, $\partial\Omega = \Sigma$, and $y \notin \Omega_0$. Then

$$\langle Q(\Sigma + \partial\Omega_0)\phi(y)\rangle = \int_{\Omega\cup\Omega_0} \langle d\star j\phi(y)\rangle = \int_\Omega \langle d\star j\phi(y)\rangle + \int_{\Omega_0}\langle d\star j\phi(y)\rangle. \tag{5.12}$$

Since $y\notin\Omega_0$, we have $\mathrm{Link}(\Sigma_0, y) = 0$, and therefore

$$\langle Q(\Sigma + \partial\Omega_0)\phi(y)\rangle = \int_\Omega \langle d\star j\phi(y)\rangle = \langle Q(\Sigma)\phi(y)\rangle. \tag{5.13}$$

Thus, the computed charge is invariant. That is the reason why we call these operators *topological*. The unitary operator is defined by taking the exponential of the conserved charge,

$$U_g(\Sigma_{D-1}) = \exp(i\alpha_a Q_a) = \exp\left(i\alpha\int\star j\right), \tag{5.14}$$

[1] A derivation of the Ward identity can be found in [5, 9].

where $g \in G$ is an element of some group G and α_a are the parameters of the transformation. In the case $G = U(1)$, we have $g = e^{i\alpha}$. We then say that this is a topological codimension-1 operator.

In terms of the unitary operator, the action on local operators is given by

$$\langle U_g(\Sigma_{D-1})\mathcal{O}(x)\rangle = \langle \mathcal{O}'(x)\rangle. \tag{5.15}$$

From now on, we will use $\mathcal{O}(x)$ to generalize the expressions to any local operator (not necessarily fields). These expressions are evaluated inside correlators as a consequence of the Ward identity, but we can understand this in another way as well. Since the manifold on which the topological operators are defined need not be a spatial manifold, it does not make sense to compare two different operators at different times, since these belong to different Hilbert spaces. However, inside correlation functions these operations are well defined. To understand pictorially the action of topological operators, consider the figure below.

Let $U(S^{D-1})$ be an operator defined on a $(D-1)$-sphere. As we have seen, these operators act on local operators $\mathcal{O}(x)$ and transform them into another operator $\mathcal{O}'(x)$ as shown in the figure

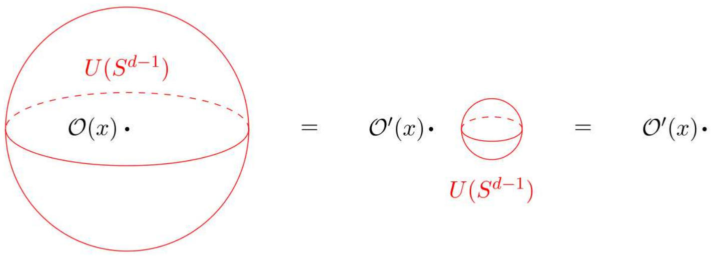

Figura 5.1 – The link of the $(D-1)$-sphere and the local operator $\mathcal{O}(x)$. Taken from [21].

The operator $U(S^{D-1})$ has a non-trivial link with $\mathcal{O}(x)$. Note that it is impossible to remove the operator $\mathcal{O}(x)$ from the manifold S^{D-1} without tearing it apart. Once we do this, we compute the action of the symmetry operator on $\mathcal{O}$. Now $U(S^{D-1})$ is topologically equivalent to a point; thus, we can deform it again and we are left only with the operator $\mathcal{O}'(x)$.

5.1.1 0-form Groups

We define the 0-form group $G^{(0)}$ to be the group composed of the elements g that parametrize the operator U_g.

The composition of two topological operators is given by the composition law of the group,

$$\langle U_g(\Sigma_{D-1})U_{g'}(\Sigma_{D-1})\rangle = \langle U_{gg'}(\Sigma_{D-1})\rangle. \tag{5.16}$$

Thus, inserting the composition of topological operators into correlation functions is equivalent to inserting a single topological operator with the composed group element. This is called the fusion rule,

$$U_g \otimes U_{g'} = U_{gg'}. \tag{5.17}$$

Therefore, the local operators transform under representations of the 0-form group $\mathcal{R}(g)$,

$$\langle U_g(\Sigma_{D-1})\mathcal{O}(x)\rangle = \mathcal{R}(g)\langle \mathcal{O}(x)\rangle. \tag{5.18}$$

From now on, we will drop the correlation function notation, but it is implicitly understood unless stated otherwise.

5.2 p-form Symmetries

p-form symmetries are straightforward generalizations of 0-form symmetries. The parameter of the transformation is a p-form, and the associated currents are $(p+1)$-forms. Therefore, the symmetry operators associated with these symmetries are codimension-$(p+1)$ operators.

To see why it must be a codimension-$(p+1)$ operator, we go back to Eq. (5.1). For a D-dimensional spacetime, and in the language of differential forms, this equation becomes

$$\delta S = \int \star j \wedge d\alpha. \tag{5.19}$$

Now α is a p-form; consequently, $d\alpha$ is a $(p+1)$-form. In order to match the dimension of spacetime, $\star j$ must be a $(D-p-1)$-form. This implies that the conserved charge is

$$Q = \int_{\Sigma_{D-p-1}} \star j, \tag{5.20}$$

and the unitary operator is defined as before,

$$U_g(\Sigma_{D-p-1}) = \exp\left(i\alpha \int_{\Sigma_{D-p-1}} \star j\right), \tag{5.21}$$

which is now understood as a codimension-$(p+1)$ operator. There is one subtlety regarding the parameter α. One may question why the parameter α that parametrizes the unitary operator is a number—should it not be a p-form? The answer is that these are the components of the p-form. Therefore, for each component we have a different unitary operator. This expression is simply written in a concise and general manner.[2]

The group of p-form symmetries $G^{(p)}$ consists of the elements that parametrize the operator $U_g(\Sigma_{D-p-1})$. Just as before, the composition of these operators is given by

$$U_g(\Sigma_{D-p-1})U_{g'}(\Sigma_{D-p-1}) = U_{gg'}(\Sigma_{D-p-1}). \tag{5.22}$$

[2] I thank Gustavo Yoshitome for clarifying this point to me.

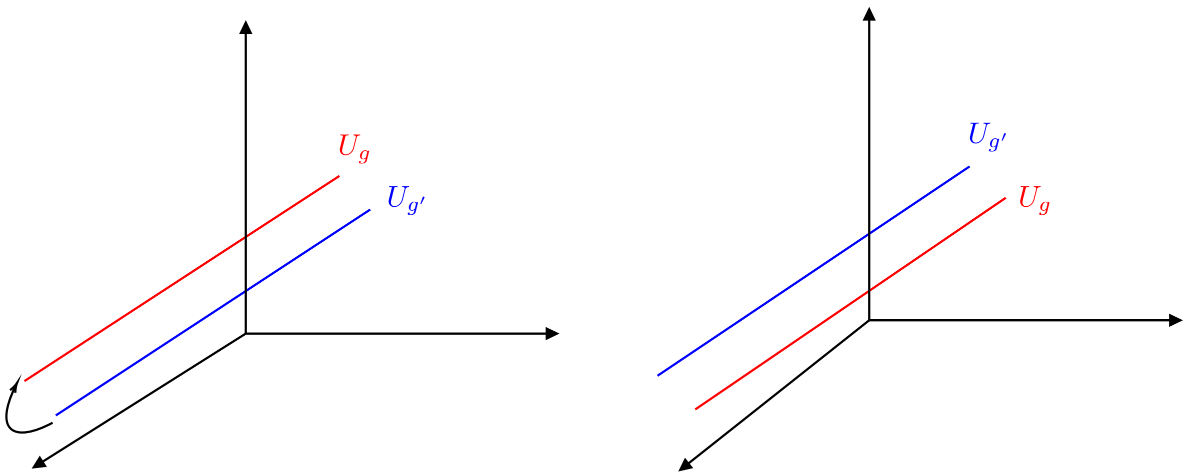

Figura 5.2 – The operators U_g and $U_{g'}$ defined on spacetime may have their ordering interchanged without one crossing the other.

Moreover, the p-form symmetry group is abelian for $p \geq 1$. This follows because we can deform one operator around another without ever crossing them, as shown in Fig. (5.2). We must now determine which objects are charged under these symmetries, that is, which objects the p-form operators act on.

5.2.1 Action of p-form symmetries

The 0-form symmetries act on local operators, which are supported on 0-dimensional regions. We then say that local operators are *charged* under the 0-form symmetry. We now wish to understand which objects are charged under p-form symmetries. It is straightforward to note that for $p \geq 1$, a p-form symmetry never acts on a local operator because we can always deform the symmetry operator in such a way that it does not link with the point. Therefore, a p-form symmetry acts on objects that have dimension $q \geq p$.

Case $q = p$

We define an operator $U_g(\Sigma_{D-p-1})$ on a $(D-p-1)$-dimensional manifold that links with an operator $\mathcal{O}(M_p)$ defined on a p-dimensional manifold. We then deform $U_g(\Sigma_{D-p-1})$ across $\mathcal{O}(M_p)$, leaving behind a phase $\phi(g)$ such that

$$\phi(g) \in \mathbb{C}^{\times}, \quad \mathbb{C}^{\times} \equiv \mathbb{C} \backslash \{0\}.$$

The action of the operator $U_g(\Sigma_{D-p-1})$ on $\mathcal{O}(M_p)$ is represented below.

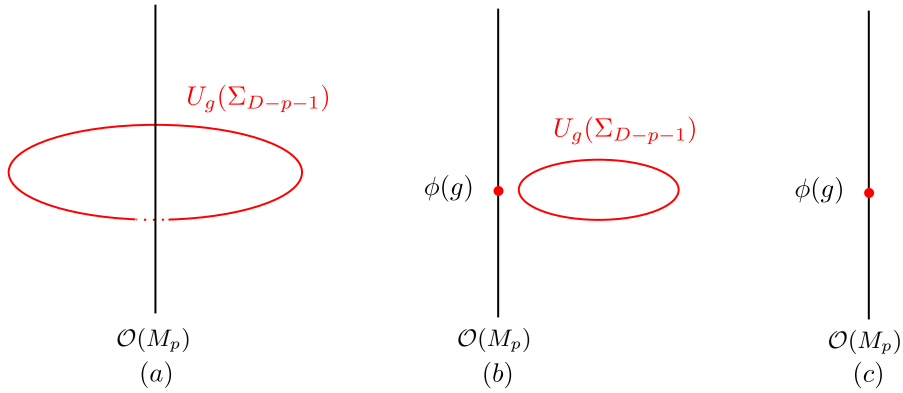

Figura 5.3 – Link between extended operators. In (a) the topological operator $U_g(\Sigma_{D-p-1})$ and the extended operator $\mathcal{O}(M_p)$ have a non-trivial link. In (b), $U_g(\Sigma_{D-p-1})$ is deformed and passes through $\mathcal{O}(M_p)$, leaving behind a phase $\phi(g)$. In (c), $U_g(\Sigma_{D-p-1})$ and $\mathcal{O}(M_p)$ no longer have a link; then $U_g(\Sigma_{D-p-1})$ is deformed to a point.

The action of the topological extended operators is given in terms of the Ward identity,

$$U_g(\Sigma_{D-p-1})\mathcal{O}(M_p) = e^{iq\lambda \text{Link}(\Sigma,M)}\mathcal{O}(M_p)U_g(\Sigma'_{D-p-1}), \tag{5.23}$$

where Σ'_{D-p-1} is a manifold homotopic to Σ_{D-p-1} that does not link with M_p. The composition law of the group induces a composition of the functions $\phi(g)$,

$$\phi(g)\phi(g') = \phi(gg'). \tag{5.24}$$

Therefore, the functions $\phi(g)$ form a one-dimensional representation of the p-form group $G^{(p)}$. Since $G^{(p)}$ is an abelian group, Schur's lemma tells us that every irreducible representation of this group is one-dimensional. Hence, the functions $\phi(g)$ are precisely the irreducible representations of the group $G^{(p)}$. Therefore, the operators $\mathcal{O}(M_p)$ transform under the fundamental representations of the p-form group, which shows that the objects charged under p-form symmetries are extended operators supported on p-dimensional manifolds. The case $q > p$ is beyond the scope of this work. A discussion of this subject can be found in references such as [4, 24].

5.2.1.1 Higher-form symmetries in D-dimensional $U(1)$ gauge theory

Let A be a 1-form $U(1)$ gauge field with

$$F = dA, \tag{5.25}$$

where F is a 2-form. In $D = 4$ this is precisely Maxwell theory. From (5.25) we have

$$dF = d\star(\star F) = 0. \tag{5.26}$$

Therefore, the Noether current is the $(D-2)$-form (up to normalizations)

$$J^m = \frac{1}{2\pi}\star F. \tag{5.27}$$

The conserved charge is

$$Q^m = \frac{1}{2\pi}\int_{\Sigma_2} \star(\star F) = \frac{1}{2\pi}\int_{\Sigma_2} F, \tag{5.28}$$

and the unitary operator is

$$U_g^m(\Sigma_2) = \exp\left(\frac{i\alpha_m}{2\pi}\int_{\Sigma_2} F\right). \tag{5.29}$$

The index m stands for *magnetic symmetry*. Below we explain the reason for this name. For $p \geq 1$-forms, we integrate the charge over a $(D-p-1)$-dimensional manifold. Since this manifold is 2-dimensional, $D-p-1=2 \Rightarrow p = D-3$. Thus, the magnetic symmetry is a $(D-3)$-form symmetry.

The other conservation law follows from the vacuum Maxwell equations,

$$d \star F = 0, \tag{5.30}$$

which implies that in this case the conserved current is the $(D-2)$-form,

$$J^e = \frac{1}{e^2} \star F. \tag{5.31}$$

The conserved charge is

$$Q^e = \frac{1}{e^2}\int_{\Sigma_{D-2}} \star F, \tag{5.32}$$

and the unitary operator is

$$U_g^e(\Sigma_{D-2}) = \exp\left(\frac{i\alpha_e}{e^2}\int_{\Sigma_{D-2}} \star F\right). \tag{5.33}$$

This is the *electric symmetry*. In this case, $D-p-1 = D-2 \Rightarrow p = 1$. Therefore, the electric symmetry is a 1-form symmetry, independent of the dimension.

Hence, this $U(1)$ gauge theory has two higher-form symmetries: an *electric 1-form symmetry* and a *magnetic* $(D-3)$*-form symmetry*. We represent the symmetry group as

$$U(1)_e^1 \times U(1)_m^{D-3}. \tag{5.34}$$

For the case $D=4$, the charged objects under the electric 1-form symmetry are the *Wilson lines* $W_{q_e}(L)$,

$$W_{q_e}(L) \equiv \exp\left(iq_e \int_L A\right), \tag{5.35}$$

where q_e is the charge of the Wilson line. The Wilson line is interpreted as the worldline of a probe particle with electric charge q_e.

Written explicitly in components,

$$W_{q_e}(L) = \exp\left(iq_e \oint dx^\mu A_\mu\right) = \exp\left(iq_e \oint dx^0 A_0 + iq_e \oint dx^i A_i\right). \tag{5.36}$$

Note that if we perform a large gauge transformation in A_0,

$$A_0 \to A_0 + \partial_0 \Lambda, \quad \Lambda = \frac{2\pi}{L_0} x^0, \tag{5.37}$$

we see that, for the zeroth component of the Wilson line,

$$\exp\left(iq_e \oint dx^0 A_0\right) \exp\left(iq_e \oint dx^0 \frac{2\pi}{L_0}\right) = \exp\left(iq_e \oint dx^0 A_0\right) \exp(2\pi i q_e). \tag{5.38}$$

Hence, in order for the Wilson line to be gauge invariant, we must have $q_e \in \mathbb{Z}$.

From the Maxwell action, one can compute the canonical momentum associated with the dynamical gauge field. It is precisely J^e. This implies that the action of the electric 1-form symmetry on the gauge field is a shift by a flat connection,

$$A \to A + \alpha_e. \tag{5.39}$$

Consequently, its action on the Wilson lines is

$$U_g^e(\Sigma_{D-2}) W_{q_e}(L) = e^{iq_e\alpha_e} W_{q_e}(L) U_g^e(\Sigma'_{D-2}), \tag{5.40}$$

where Σ'_{D-2} is homotopic to Σ_{D-2} and does not link with L.

The magnetic $(D-3)$-form symmetry acts on 't Hooft lines, which are interpreted as the worldlines of probe particles with magnetic charge q_m. To write explicitly the 't Hooft line, we introduce the dual field $\hat{A}$ following from Eq. (5.30),

$$\star F = d\hat{A}. \tag{5.41}$$

Thus, we define the 't Hooft line as

$$T(L) \equiv \exp\left(iq_m \int_L \hat{A}\right). \tag{5.42}$$

Analogously to the electric 1-form symmetry, the magnetic $(D-3)$-form symmetry acts by shifting the dual field $\hat{A}$ by flat connections,

$$\hat{A} \to \hat{A} + \alpha_m. \tag{5.43}$$

Consequently, the action on 't Hooft lines gives

$$U_g^m(\Sigma_2) T(L) = e^{iq_m\alpha_m} T(L) U_g^m(\Sigma'_2), \tag{5.44}$$

where Σ'_2 is homotopic to Σ_2 and does not link with L.

5.2.1.2 Higher-form symmetries in Chern-Simons theory

In Chapter 3 we introduced Chern-Simons theory, but we did not discuss one of its most important features, namely its degeneracy. We now demonstrate this by analyzing the higher-form symmetries of the theory.

Consider the Chern-Simons action,

$$S_{CS} = \int d^3x \frac{k}{4\pi} \epsilon_{\mu\nu\rho} a_\mu \partial_\nu a_\rho.$$

We can extract the commutation relations from it. Choosing the gauge $a_0 = 0$, we obtain

$$S_{CS} = \int d^3x \frac{k}{4\pi} \epsilon^{ij} a_i \dot{a}_j. \tag{5.45}$$

The canonical momenta are

$$\Pi_j = \frac{\partial \mathcal{L}}{\partial \dot{a}_j} = \frac{k}{4\pi} \epsilon_{ij} a_i. \tag{5.46}$$

Imposing the commutation relation

$$[a_i(x), \Pi_j(y)] = i\delta_{ij}\delta(x-y), \tag{5.47}$$

we obtain

$$\left[a_i(x), \frac{k}{4\pi}\epsilon_{kj} a_k(y)\right] = i\delta_{ij}\delta(x-y). \tag{5.48}$$

Multiplying by ϵ^{jm} on both sides,

$$[a_i(x), a_m(y)] = \frac{4\pi i}{k}\epsilon_{im}\delta(x-y). \tag{5.49}$$

In particular,

$$[a_1(x), a_2(y)] = \frac{4\pi i}{k}\delta(x-y). \tag{5.50}$$

Thus, the components of the gauge field do not commute with each other. We construct the Wilson lines of the theory,

$$W^1(C_1) \equiv \exp\left(iq \oint_{C_1} dx^1 a_1\right), \tag{5.51}$$

$$W^2(C_2) \equiv \exp\left(iq \oint_{C_2} dx^2 a_2\right), \tag{5.52}$$

and compute W^1W^2. Using the Baker–Campbell–Hausdorff formula,

$$W^1W^2 = \exp\left(i\oint_{C_1} dx\, a_1 + i\oint_{C_2} dy\, a_2 + \frac{1}{2}\oint dx\, dy\, [a_1, a_2]\right) = W^2W^1 \exp\left(\frac{2\pi i}{k}\oint dx\, dy\, \delta^2(x-y)\right), \tag{5.53}$$

where we have taken $q = 1$ for simplicity. Therefore,

$$W^1W^2 = e^{-\frac{2\pi i}{k}} W^2W^1. \tag{5.54}$$

We may also compute the conserved quantities. Writing the Chern-Simons action in terms of differential forms,

$$S_{CS} = \int \frac{k}{4\pi} a \wedge da, \tag{5.55}$$

we see from the equations of motion for a that

$$da = 0. \tag{5.56}$$

This is precisely the conservation law for a 1-form symmetry. The conserved charge is then

$$Q(C) = q\oint_C a, \tag{5.57}$$

for some curve C and parameter q. The unitary operator is

$$W(C) = \exp\left(iq \oint_C a\right). \tag{5.58}$$

It is now clear that this is a Wilson line in Chern-Simons theory and that q is its charge. Using the same method as for Wilson lines in Maxwell theory, one can show that $q \in \mathbb{Z}$. We may label the Wilson lines by an integer representing their charge, e.g., $W_n(C)$.

From (5.54) we see that Wilson lines act on other Wilson lines. However, from (5.58) we also see that Wilson lines are symmetry operators. Thus, in Chern-Simons theory, Wilson lines are both the charged objects and the symmetry operators. Since these are symmetries,

$$[H, W] = 0. \tag{5.59}$$

Therefore, the energy eigenstates are also eigenstates of W^1 and W^2, but they cannot be simultaneous eigenstates of both, since W^1 and W^2 do not commute. This reveals that the theory has a degeneracy. To determine the order of the degeneracy, apply an energy eigenstate $|n\rangle$ to (5.54),

$$W^1W^2|n\rangle = e^{-\frac{2\pi i}{k}}W^2W^1|n\rangle = \omega_1 e^{-\frac{2\pi i}{k}}W^2|n\rangle, \tag{5.60}$$

where ω_1 is the eigenvalue of W^1. Now apply W^2 k times in this equation. Using the commutation relation to reorder the operators, each reordering produces a phase $e^{-\frac{2\pi i}{k}}$, yielding

$$W^1(W^2)^k|n\rangle = \omega_1(W^2)^k|n\rangle. \tag{5.61}$$

This shows that after k applications, $(W^2)^k|n\rangle$ is again an eigenstate of W^1 with the same eigenvalue, and therefore the degeneracy is of order k.

More generally, applying W^1 n times and W^2 m times, we obtain

$$(W^1)^n(W^2)^m = e^{-\frac{2\pi i}{k}nm}(W^2)^m(W^1)^n, \quad m, n \in \mathbb{Z}. \tag{5.62}$$

Thus, $(W^1)^k$ or $(W^2)^k$ behave as the identity. It might then appear that the lines with charges n and $n+k$ can be identified. However, this identification is not straightforward. To see this, consider the link between two Wilson lines, where one line extends to infinity and the other is closed around the first line, as represented in Fig. (5.4),

We can interpret these lines as the worldlines of particles with statistics ν and spin $s = \nu/2$. The statistics is determined from the phase $e^{\pm i\nu\pi}$ that the particles acquire upon exchange. This is equivalent to performing a half-rotation and then translating the particle. From (5.54),

$$W_n(C)W_n(C') = e^{-\frac{2\pi i n^2}{k}}W_n(C')W_n(C). \tag{5.63}$$

This shows that

$$\nu = \frac{n^2}{k} \bmod 2 \quad \Rightarrow \quad s = \frac{n^2}{2k} \bmod 1. \tag{5.64}$$

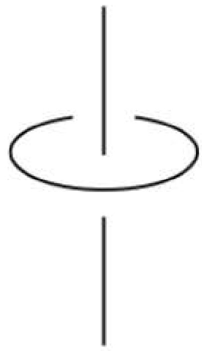

Figura 5.4 – Link between two Wilson lines. Taken from [5].

For a line with charge $n+k$, the statistics becomes

$$\nu = \frac{(n+k)^2}{k} = \frac{n^2}{k} + 2n + k. \tag{5.65}$$

This leads to two conclusions:

- **For even k:** writing $k = 2m$, $m \in \mathbb{Z}$, and since ν is defined mod 2, $\nu(n)$ and $\nu(n+k)$ are identified. Therefore, there are k independent lines.
- **For odd k:** writing $k = 2m+1$, $m \in \mathbb{Z}$, we find that $\nu(n)$ and $\nu(n+k)$ differ mod 1. To identify the lines consistently, we must take $k \to 2k$. Therefore, there are $2k$ independent lines.

5.2.2 Spontaneous symmetry breaking of higher-form symmetries

At the end of the day, we are working with symmetries. Therefore, we can use the framework we are familiar with and apply it to higher-form symmetries as well. One interesting feature that symmetries exhibit, and which we have shown in Chapter 3, is that they can be spontaneously broken.

Recall our definition of SSB. To determine whether or not a symmetry is spontaneously broken, we evaluate the expectation value of an operator $\mathcal{O}(x)$ that is charged under the symmetry. We then concluded that

$$\langle \mathcal{O}(x) \rangle \begin{cases} = 0, & \text{symmetric phase,} \\ \neq 0, & \text{symmetry broken phase.} \end{cases} \tag{5.66}$$

For p-form symmetries, this behavior can be translated as follows. The charged operators are p-dimensional operators. In general, the expectation value of such operators can be written as

$$\langle \mathcal{O}(\Sigma_p) \rangle \sim e^{-f(\Sigma_p)}, \tag{5.67}$$

where $f(\Sigma_p)$ is an arbitrary function. Define the function $\mathrm{Vol}(\Sigma_p)$. Note that this function does not compute the volume enclosed by Σ_p, but rather the appropriate p-dimensional measure. For example, for $p = 1$, the "volume" is the perimeter of the curve Σ_p. Therefore, we may reproduce the behavior of Eq. (5.66) as

$$\lim_{\mathrm{Vol}(\Sigma_p)\to\infty} \left(\frac{f(\Sigma_p)}{\mathrm{Vol}(\Sigma_p)} \right) = \begin{cases} \infty, & \text{symmetric phase,} \\ \text{finite}, & \text{symmetry broken phase.} \end{cases} \tag{5.68}$$

Note that in the finite case, we can always add a counterterm and define a renormalized operator,

$$\hat{\mathcal{O}}(\Sigma_p) \equiv \exp\left(-c \oint_{\Sigma_p} dV \right) \mathcal{O}(\Sigma_p), \tag{5.69}$$

such that

$$\langle \hat{\mathcal{O}} \rangle \neq 0. \tag{5.70}$$

Therefore, the vacuum is charged under the generalized symmetry. In the infinite case, there is no way to overcome this behavior by adding a local counterterm, and the vacuum is not charged under the generalized symmetry. These ideas are rather abstract, but we will now present an example using objects with which we are familiar.

5.2.2.1 Spontaneous symmetry breaking in Maxwell theory

As we have seen, the charged operators in Maxwell theory are the Wilson and 't Hooft lines. The Wilson loop is known to have an expectation value that typically depends on geometrical properties, such as the area enclosed by the curve C on which it is defined or the perimeter of the curve,

$$\langle W(C) \rangle \sim e^{-\mathrm{Area}(C)}, \tag{5.71}$$

$$\langle W(C) \rangle \sim e^{-\mathrm{Perimeter}(C)}. \tag{5.72}$$

The area grows faster than the perimeter of the curve,

$$\lim_{\mathrm{Perimeter}(C)\to\infty} \left(\frac{\mathrm{Area}(C)}{\mathrm{Perimeter}(C)} \right) \to \infty, \tag{5.73}$$

$$\lim_{\mathrm{Perimeter}(C)\to\infty} \left(\frac{\mathrm{Perimeter}(C)}{\mathrm{Perimeter}(C)} \right) \to \text{finite}. \tag{5.74}$$

In this case, $\mathrm{Vol}(\Sigma_p)$ corresponds to the perimeter of the curve C. Essentially, what we have is that

$$e^{-\mathrm{Area}(C)} \to 0, \tag{5.75}$$

$$e^{-\mathrm{Perimeter}(C)} \neq 0. \tag{5.76}$$

For the case of Maxwell theory, the expectation value of the Wilson loop decays with a scale invariant dependence on the parameters of the loop, since Maxwell is a conformal-invariant theory. Its decay is given by

$$\langle W(C) \rangle \sim e^{-\alpha \frac{T}{R} - \beta \frac{R}{T}}, \tag{5.77}$$

where α, β are dimensionless quantities and T and R are the sides of the loop (considering is as a rectangle for simplicity). Even thought this does not have the same behavior as are or perimeter, is decays slower than the perimeter law. So essentially it does not goes to zero as well.

Therefore, the Wilson loop is a suitable choice for an order parameter for SSB in Maxwell theory. We now proceed to analyze this for a 1-form symmetry and show a very interesting result.

We have shown that a Goldstone boson can be understood from the amplitude

$$\langle n, \boldsymbol{k} | j_{\mu\nu} | 0 \rangle, \tag{5.78}$$

where n is some quantum number (in our derivation it was the energy) and $\boldsymbol{k}$ is its momentum. We will compute this amplitude by considering the one-photon state $|\lambda, \boldsymbol{k}\rangle$ with polarization λ and momentum $\boldsymbol{k}$. The one-photon state is given by the action of the photon creation operator on the vacuum,

$$|\lambda, \boldsymbol{k}\rangle = A^{\dagger}_{\lambda}(k)|0\rangle. \tag{5.79}$$

When quantizing Maxwell theory, we note that the Lorentz gauge condition $\partial_\mu A^\mu$ does not hold as an operator equation, since this would lead to contradictions in the commutation relations. We must impose it as an equality acting on physical states, that is,

$$\partial_\mu A^\mu |\text{phys}\rangle = 0. \tag{5.80}$$

This implies that there are degrees of freedom that are not physical. These discussions are brief here, but they are standard and can be found in many QFT textbooks.

Suppose that only $\lambda = 1, 2$ are physical polarizations. The gauge field can be written as

$$A^\mu(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3\boldsymbol{p}}{2|\boldsymbol{p}|} \sum_{\lambda=1}^{2} {e^\mu}_\lambda(p) (A_\lambda(p) e^{-ipx} + A^\dagger_\lambda(p) e^{ipx}). \tag{5.81}$$

The commutation relations are

$$\left[A_\lambda(p), A^\dagger_{\lambda'}(p')\right] = 2|\boldsymbol{p}| \delta_{\lambda,\lambda'} \delta^3(\boldsymbol{p} - \boldsymbol{p}'). \tag{5.82}$$

Now we compute

$$\langle 0 | F_{\mu\nu} | \lambda, \boldsymbol{p} \rangle = \langle 0 | (\partial_\mu A_\nu - \partial_\nu A_\mu) A^\dagger_\lambda(p) | 0 \rangle. \tag{5.83}$$

One of the terms is

$$\partial_\mu A_\nu = \frac{i}{(2\pi)^{3/2}} \int \frac{d^3p}{2|\boldsymbol{p}|} \sum_{\lambda=1}^{2} p_\mu {\epsilon_\nu}^\lambda (A^\dagger_\lambda(p) e^{ipx} - A_\lambda(p) e^{-ipx}). \tag{5.84}$$

Thus,

$$\langle 0 | \partial_\mu A_\nu A^\dagger_\lambda | 0 \rangle = \frac{i}{(2\pi)^{3/2}} \left\langle 0 \middle| \int \frac{d^3p'}{2|\boldsymbol{p}'|} \sum_{\lambda'=1}^{2} p'_\mu {\epsilon_\nu}^{\lambda'} (A^\dagger_{\lambda'}(p') e^{ip'x} - A_{\lambda'}(p') e^{-ip'x}) A^\dagger_\lambda(p) \middle| 0 \right\rangle. \tag{5.85}$$

Using the commutation relations, we obtain

$$\begin{aligned}\langle 0|\partial_\mu A_\nu A_\lambda^\dagger|0\rangle = \frac{i}{(2\pi)^{3/2}}\langle 0|\int \frac{d^3p'}{2|\boldsymbol{p}'|}\sum_{\lambda'=1}^{2} p'_\mu {\epsilon_\nu}^{\lambda'} A^\dagger_{\lambda'}(p')A^\dagger_\lambda(p)e^{ip'x}|0\rangle \\ -\frac{i}{(2\pi)^{3/2}}e^{-ipx}{\epsilon_\nu}^{\lambda}p_\mu,\end{aligned} \tag{5.86}$$

where the first term vanishes because the creation operators act on the vacuum to their left. The other term gives the same contribution with the indices exchanged. Thus

$$\langle 0|F_{\mu\nu}|\lambda,\boldsymbol{p}\rangle = \frac{i}{(2\pi)^{3/2}}e^{-ipx}({\epsilon_\mu}^{\lambda}p_\nu - {\epsilon_\nu}^{\lambda}p_\mu). \tag{5.87}$$

This shows that the photon is the Goldstone boson arising from the spontaneous symmetry breaking of a 1-form symmetry in Maxwell theory.

There are natural questions to ask. Which symmetry is being broken here? Since Maxwell theory possesses two 1-form symmetries, do we obtain two Goldstone bosons?

Although the charges are globally independent, if we inspect locally, that is, the charge densities, F and $\star F$ are related by a Levi-Civita. Thus, these symmetries are not locally independent. This shows that in fact there is only one Goldstone boson arising from the spontaneous symmetry breaking in Maxwell theory.

Note that if we add electric matter, Eq.(5.30) is no longer a symmetry. Nevertheless, a Goldstone boson (the Photon) still arises due to Eq.(5.26). Conversely, if we add magnetic matter, Eq.(5.26) ceases to be a symmetry, while Eq.(5.30) remains intact. Therefore, unless both symmetries are explicitly broken, the Photon always appears as a Goldstone boson.

5.3 Discrete Gauge Theory

Discrete gauge transformations belong to discrete groups, rather than continuous ones. This can be better understood in terms of higher-form symmetries. Before explaining how this works, we need to introduce two concepts: the *Pontryagin dual* and the *screening* of operators.

5.3.1 Pontryagin Dual

For some p-form symmetry group $G^{(p)}$, we may define the homomorphisms

$$\phi : G^{(p)} \to U(1). \tag{5.88}$$

These maps form a group with composition given by

$$\phi(g)\phi'(g) = (\phi\phi')(g) \in U(1). \tag{5.89}$$

We denote this group by

$$\widehat{G}^{(p)} = \{\text{all homomorphisms } G^{(p)} \to U(1)\}. \tag{5.90}$$

Let us illustrate this idea with two examples.

Example: $G^{(p)} = U(1)$

Define the homomorphism as

$$\phi(g) = g^q = e^{i\alpha q},\ \alpha \in [0, 2\pi),\ q \in \mathbb{Z}. \tag{5.91}$$

Let

$$\phi_1(g) = g, \tag{5.92}$$
$$\phi_2(g) = g^2, \tag{5.93}$$
$$\phi_3(g) = g^3. \tag{5.94}$$

The composition gives

$$\phi_1(g)\phi_2(g) = \phi_3(g). \tag{5.95}$$

This is the composition rule of the $\mathbb{Z}$ group. Therefore, the Pontryagin dual of $U(1)$ is

$$\widehat{G}^{(p)} = \mathbb{Z}. \tag{5.96}$$

Example: $G^{(p)} = \mathbb{Z}_N$

In this case, $g = e^{\frac{2\pi i\alpha}{N}}$. Defining the homomorphism as

$$\phi = g^q = e^{\frac{2\pi i\alpha q}{N}}, \tag{5.97}$$

take $N = 3$ and let

$$\phi_1(g) = g, \tag{5.98}$$
$$\phi_2(g) = g^2. \tag{5.99}$$

The composition gives

$$\phi_1(g)\phi_2(g) = \phi_1(g). \tag{5.100}$$

This is the composition rule of $\mathbb{Z}_N$. Therefore,

$$\widehat{G}^{(p)} = \mathbb{Z}_N. \tag{5.101}$$

That is, the Pontryagin dual of $\mathbb{Z}_N$ is the group itself.

Double Pontryagin Duality

Since
$$\phi : G^{(p)} \to U(1), \tag{5.102}$$
we can define
$$\varphi : \widehat{G}^{(p)} \to U(1). \tag{5.103}$$
Note that $\phi(g) \in U(1)$ and $\varphi(\phi(g)) \in U(1)$. Then,
$$\widehat{\widehat{G}}^{(p)} = G^{(p)}. \tag{5.104}$$

5.3.2 Screening

Definition 5.1 *A $p \geq 1$-dimensional operator $\mathcal{O}_p$ can be screened into another operator $\mathcal{O}'_p$ if there exists a $(p-1)$-dimensional operator $\mathcal{O}_{p-1}$ that can be inserted between $\mathcal{O}_p$ and $\mathcal{O}'_p$.*

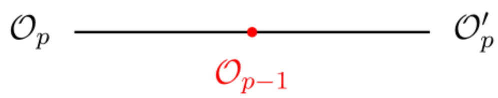

Figura 5.5 – Screening of operators. Taken from [21].

If $\mathcal{O}_p$ can be screened to the identity, we say that it is completely screened.

Screening implies equal charges. Note that if two operators $\mathcal{O}_p$ and $\mathcal{O}'_p$ are screened, we can deform the symmetry operator, say $U_g(S^{D-p-1})$, to measure the charge of either $\mathcal{O}_p$ or $\mathcal{O}'_p$. Since this deformation does not cross the regions where the operators are defined, the measured charges must be equal. The figure below represents this idea.

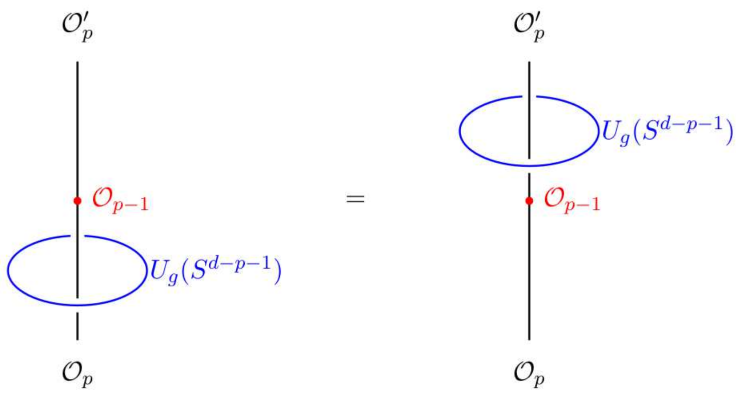

Figura 5.6 – Link with two screened operators. Taken from [21].

If two operators carry different charges, they cannot be screened into one another.

As we have shown, the charges are given by phases. Each phase corresponds to a different charge, that is,

$$Q : G^{(p)} \to U(1). \tag{5.105}$$

Therefore, the charge operators are elements of $\widehat{G}^{(p)}$. Thus, we may define an equivalence class of operators with the equivalence relation $\sim$, characterized by whether these operators can be screened into one another,

$$\mathcal{O}'_p \sim \mathcal{O}_p.$$

We define the set D_p, where each element of this set is an equivalence class of screened operators. Therefore, there must exist p-form symmetries that distinguish different elements of D_p. The elements of D_p are precisely the different charges that an operator may have. Then,

$$\widehat{G}^{(p)} = D_p,$$

which implies

$$G^{(p)} = \widehat{D}_p.$$

Hence,

$$Q(\mathcal{O}_p) = [\mathcal{O}_p] \in D_p. \tag{5.106}$$

Example: Maxwell Theory in the Presence of Matter Fields

We have seen that in Maxwell theory we have a 1-form symmetry and a $(D-3)$-form symmetry[3]. In order to have these symmetries, we evaluated the theory in the vacuum, that is,

$$d \star F = 0.$$

One may ask: what happens if we add matter? We can understand the behavior using the concept of screening.

In the vacuum case, the symmetry group is $U(1)_e \times U(1)_m$. The Pontryagin dual is

$$\widehat{G}^{(1)} = \mathbb{Z},$$
$$\widehat{G}^{(D-3)} = \mathbb{Z}.$$

Therefore,

$$D_1 = \mathbb{Z},$$
$$D_{D-3} = \mathbb{Z}.$$

Suppose now that we add a matter field $\phi(x)$ with charge $N \in \mathbb{Z}$. We can use the operator $\phi(x)$ to screen the Wilson lines.

[3] Actually, Maxwell theory is specifically in $D = 4$, in which case we have two 1-form symmetries. For simplicity, however, we keep this notation.

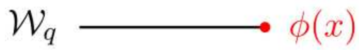

Figura 5.7 – Non-genuine operator. Taken from [21].

We then say that $\phi(x)$ is a *non-genuine* operator.

Definition 5.2 (Non-genuine operator) *A non-genuine q-dimensional operator is an operator that is attached to a collection of $p > q$-dimensional operators.*

We must introduce these non-genuine operators because $\phi(x)$ and $W_q(L)$ are not gauge invariant, since

$$\phi(x) \to e^{iq\Lambda(x)}\phi(x), \tag{5.107}$$

and

$$W_q(L) = \exp\left(2\pi i q \int_L A\right) \to \exp\left(2\pi i q \int_L A - \frac{d\Lambda}{2\pi}\right) = W_q \exp\left(-iq\int_{\partial L}\Lambda\right). \tag{5.108}$$

Assuming that $\partial L = x$,

$$W_q'(L) = e^{-iq\Lambda(x)}W_q(L). \tag{5.109}$$

Therefore, the product $\phi(x)W_q(L)$ is gauge invariant. With this, we can screen several Wilson lines by adding powers of ϕ.

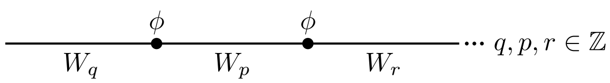

Figura 5.8 – Screening of multiple Wilson lines using powers of ϕ.

Thus, the group of all unscreened operators is given by the total number of Wilson lines modulo the number of screened Wilson lines,

$$D_1 = \frac{\mathbb{Z}}{N\mathbb{Z}} \equiv \mathbb{Z}_N. \tag{5.110}$$

Note that $N\mathbb{Z}$ counts the number of screened Wilson lines because, since they have the same charge, the charges add up N times. Taking the Pontryagin dual,

$$\widehat{D_1} = \widehat{\mathbb{Z}}_N = \mathbb{Z}_N = G^{(1)}. \tag{5.111}$$

We conclude that adding a matter field with charge N reduces the gauge group from $U(1)$ to $\mathbb{Z}_N$. The term "break" is not quite correct, since gauge symmetries are not physical symmetries. The appropriate term found in the literature is that the symmetry group is *Higgsed down* to $\mathbb{Z}_N$. Note that this does not occur for the $(D-3)$-form symmetry, since this symmetry is associated with a codimension-2 operator that does not have a trivial linking with the Wilson line.

5.3.2.1 BF Theory revisited

We studied in Chapter 4 the BF theory in 2+1 dimensions. We began by adding a matter field with charge N and then, upon condensation and taking the low-energy limit, we obtained

$$S_{BF} = \int \frac{iN}{2\pi} b \wedge da - a \wedge j - b \wedge \tilde{j} - \frac{1}{4}(f^a)^2 - \frac{1}{4}(f^b)^2. \tag{5.112}$$

Considering the pure BF theory,

$$S_{BF} = \frac{iN}{2\pi} \int b \wedge da, \tag{5.113}$$

the equation of motion for b leads to

$$da = 0 \Rightarrow d \star (\star a) = 0. \tag{5.114}$$

Then, the conserved charge is

$$Q = \int_C a, \tag{5.115}$$

where C is a curve. If instead we write

$$S_{BF} = -\frac{iN}{2\pi} \int a \wedge db, \tag{5.116}$$

and compute the equation of motion for a, we obtain

$$db = 0, \tag{5.117}$$

with the conserved current given by

$$Q = \int_{C'} b, \tag{5.118}$$

where C' is also a curve. The unitary operators (with unit charge) are

$$W_A(C) = e^{i \int a}, \tag{5.119}$$

$$W_B(C') = e^{i \int b}. \tag{5.120}$$

Therefore, the conserved quantities of the 2+1 BF theory are the Wilson lines for the a and b fields. Note that this theory, like Chern–Simons theory, is purely topological, i.e., it does not involve the metric. Now, the idea is the same as in the Chern–Simons case. From the action we can compute the canonical momenta (using the gauge $a_0 = b_0 = 0$),

$$\begin{aligned} \Pi_1^b &= \Pi_2^b = 0, \\ \Pi_1^a &= \frac{\partial \mathcal{L}}{\partial \dot{a}_1} = -\frac{iN}{2\pi} b_2, \\ \Pi_2^a &= \frac{\partial \mathcal{L}}{\partial \dot{a}_2} = \frac{iN}{2\pi} b_1. \end{aligned} \tag{5.121}$$

The commutation relations

$$\big[a_i(x), \Pi_j^a(y)\big] = i\delta_{ij}\delta(x-y), \tag{5.122}$$

lead to

$$[a_1(x), b_2(y)] = -\frac{2\pi}{N}\delta(x-y), \tag{5.123}$$

$$[a_2(x), b_1(y)] = \frac{2\pi}{N}\delta(x-y), \tag{5.124}$$

$$[a_1(x), \Pi_2^a(y)] = [a_2(x), \Pi_1^a(y)] = 0. \tag{5.125}$$

This implies that

$$W_1^a W_2^b = e^{-\frac{2\pi i}{N}} W_2^b W_1^a. \tag{5.126}$$

On the other hand, since W^A and W^B are symmetries, they commute with the Hamiltonian. Therefore, we reach the same conclusion as in the Chern–Simons case: the energy eigenstates are degenerate. It appears that the degeneracy is of order N, but we have only considered the commutator (5.123). We must also consider the commutator (5.124),

$$W_2^a W_1^b = e^{-\frac{2\pi i}{N}} W_1^b W_2^a. \tag{5.127}$$

Hence, we obtain an N-fold degeneracy from Eq. (5.126) and another N-fold degeneracy from Eq. (5.127), leading to an N^2-fold degeneracy.

Recalling how we arrived at BF theory, we had to introduce a scalar field of charge N. As shown in the previous section, this reduces the symmetry group from $U(1)$ to $\mathbb{Z}_N$. This implies that the gauge field a has discrete gauge transformations. But what about the field b?

Note that if we start from the BF theory

$$\mathcal{L} = \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}\partial^\mu b^\nu a^\rho \tag{5.128}$$

and add an auxiliary field $\hat{a}$ that enforces the constraint $f = da$,

$$\mathcal{L} = \frac{iN}{2\pi}\epsilon_{\mu\nu\rho}\partial^\mu b^\nu a^\rho + \frac{i}{2\pi}\partial_\mu \hat{a}_\nu f^{\mu\nu}, \tag{5.129}$$

we can rewrite the Lagrangian as

$$\mathcal{L} = -\frac{iN}{4\pi}\epsilon_{\mu\nu\rho} b^\mu f^{\nu\rho} - \frac{1}{4}(f_{\mu\nu})^2 + \frac{i}{2\pi}\partial_\mu \hat{a}_\nu f^{\mu\nu}. \tag{5.130}$$

We then integrate out $f_{\mu\nu}$, which leads to

$$\mathcal{L} = \frac{iN}{2\pi}(\partial_\mu \hat{a}_\nu - N\epsilon_{\mu\nu\rho} b^\rho)^2. \tag{5.131}$$

This Lagrangian has the same form as the Lagrangian (4.83). Hence, by analogy, we can interpret the dual field $\hat{a}$ as a matter field that is charged under b. Therefore, BF theory provides an example of a discrete gauge theory.

FINAL REMARKS

In this work, we have studied two examples of topological field theories: Chern–Simons theory and BF theory. In the case of Chern–Simons theory, we showed that the Chern–Simons term acts as a topological mass–generation mechanism when coupled to a gauge theory, giving mass to the gauge field without the need for a conventional Higgs field. Moreover, Chern–Simons theory correctly reproduces the Hall conductivities for both the integer and fractional quantum Hall effects, as well as the existence of quasiparticles with exotic statistics, known as anyons. For these reasons, Chern–Simons theory is regarded as the effective field theory describing the quantum Hall effect. We also showed that BF theory can be obtained as the low-energy effective theory of the Abelian Higgs model.

We further demonstrated that higher-form symmetries constitute a natural generalization of ordinary symmetries, in which the charged objects are extended operators rather than point-like ones. We analyzed how these symmetries arise in Maxwell theory and obtained the remarkable result that the photon can be interpreted as a Goldstone boson associated with the spontaneous breaking of a higher-form symmetry. Within this framework, we investigated additional aspects of Chern–Simons and BF theories, such as their ground-state degeneracy. We explicitly constructed the symmetry operators in these theories and showed that they do not commute with each other, leading to a degeneracy of order k for Chern–Simons theory and of order N^2 for BF theory. Furthermore, by employing the language of higher-form symmetries, we showed that BF theory describes a discrete gauge theory: introducing a matter field of charge N reduces the gauge group from $U(1)$ to $\mathbb{Z}_N$.

Several topics were left beyond the scope of this work, notably the K-matrix formalism, which provides a systematic generalization of Chern–Simons theory to multiple gauge fields. From this perspective, BF theory can be viewed as a particular example of a Chern–Simons theory with two gauge fields. The K-matrix formalism offers a powerful tool to analyze the spectrum and topological properties of more general theories and will play an important role in future investigations.

The next steps of this research include developing a deeper understanding of discrete gauge theories formulated intrinsically, rather than as low-energy limits of continuous gauge theories. In addition, we aim to study the gauging of discrete global symmetries in a systematic way, as well as the relationship between level-$4k$ and level-k Chern–Simons theories, which are connected through the gauging of a discrete global 1-form symmetry [27].

REFERÊNCIAS

[1] K. V. Klitzing, G. Dorda, and M. Pepper, “New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance,” *Physical Review Letters*, vol. 45, pp. 494–497, Aug. 1980. 12, 66

[2] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, “Generalized global symmetries,” *Journal of High Energy Physics*, vol. 2015, p. 172, Feb. 2015. 12, 81

[3] A. Kapustin and N. Seiberg, “Coupling a qft to a tqft and duality,” *Journal of High Energy Physics*, vol. 2014, Apr. 2014. 12, 81

[4] L. Bhardwaj and S. Schäfer-Nameki, “Generalized charges, part i: Invertible symmetries and higher representations,” *SciPost Physics*, vol. 16, Apr. 2024. 12, 86

[5] P. R. S. Gomes, “An Introduction to Higher-Form Symmetries,” 2023. Publisher: arXiv Version Number: 2. 12, 81, 82, 91

[6] X. Chen, “Essay: Generalized landau paradigm for quantum phases and phase transitions,” *Physical Review Letters*, vol. 135, Dec. 2025. 12

[7] L. Bhardwaj, L. E. Bottini, D. Pajer, and S. Schäfer-Nameki, “Categorical landau paradigm for gapped phases,” *Phys. Rev. Lett.*, vol. 133, p. 161601, Oct 2024. 12

[8] L. H. Ryder, *Quantum field theory.* Cambridge ; New York: Cambridge University Press, 2nd ed ed., 1996. 13, 17, 20, 21, 26, 27, 53, 55

[9] M. D. Schwartz, *Quantum field theory and the standard model.* New York: Cambridge University Press, 2014. 13, 53, 82

[10] A. Beekman, L. Rademaker, and J. Van Wezel, “An introduction to spontaneous symmetry breaking,” *SciPost Physics Lecture Notes*, p. 11, Dec. 2019. 13, 53

[11] C. N. Yang and R. L. Mills, “Conservation of Isotopic Spin and Isotopic Gauge Invariance,” *Physical Review*, vol. 96, pp. 191–195, Oct. 1954. 21

[12] R. A. Bertlmann, *Anomalies in quantum field theory.* No. 91 in Oxford science publications, Oxford ; New York: Clarendon Press, 1996. 31, 32, 33, 34, 35, 36, 38, 39, 49, 51

[13] M. Nakahara, *Geometry, topology and physics.* Graduate student series in physics, Bristol Philadelphia: Institute of physics publ, 2nd ed ed., 2003. 31, 37, 73

[14] P. Dirac, “Quantised singularities in the electromagnetic field„” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 133, pp. 60–72, Sept. 1931. 48

[15] D. Tong, “Lectures on the Quantum Hall Effect,” Sept. 2016. arXiv:1606.06687 [hep-th]. 64, 66, 67

[16] P. A. M. Dirac, “Generalized hamiltonian dynamics,” *Canadian Journal of Mathematics*, vol. 2, p. 129–148, 1950. 73

[17] G. V. Dunne, “Aspects of Chern-Simons Theory,” 1999. Version Number: 1. 74, 78

[18] G. V. Dunne, R. Jackiw, and C. A. Trugenberger, “"Topological"(Chern-Simons) quantum mechanics,” *Physical Review D*, vol. 41, pp. 661–666, Jan. 1990. 74

[19] S. M. Girvin, “The Quantum Hall Effect: Novel Excitations and Broken Symmetries,” 1999. Publisher: arXiv Version Number: 1. 74

[20] R. B. Laughlin, “Anomalous quantum hall effect: An incompressible quantum fluid with fractionally charged excitations,” *Phys. Rev. Lett.*, vol. 50, pp. 1395–1398, May 1983. 75

[21] L. Bhardwaj, L. E. Bottini, L. Fraser-Taliente, L. Gladden, D. S. Gould, A. Platschorre, and H. Tillim, “Lectures on generalized symmetries,” *Physics Reports*, vol. 1051, pp. 1–87, Feb. 2024. 81, 83, 96, 98

[22] D. Tong, “Gauge Theory,” 2018. 81

[23] T. D. Brennan and S. Hong, “Introduction to Generalized Global Symmetries in QFT and Particle Physics,” June 2023. arXiv:2306.00912 [hep-ph]. 81

[24] T. Bartsch, M. Bullimore, and A. Grigoletto, “Higher representations for extended operators,” 2023. 86

[25] H. Watanabe and T. Brauner, “Number of nambu-goldstone bosons and its relation to charge densities,” *Physical Review D*, vol. 84, Dec. 2011.

[26] S. Schäfer-Nameki, “ICTP lectures on (non-)invertible generalized symmetries,” *Physics Reports*, vol. 1063, pp. 1–55, Apr. 2024.

[27] N. Seiberg and E. Witten, “Gapped boundary phases of topological insulators via weak coupling,” 2016. 101

[28] N. Seiberg, T. Senthil, C. Wang, and E. Witten, “A duality web in 2+1 dimensions and condensed matter physics,” *Annals of Physics*, vol. 374, p. 395–433, Nov. 2016.

[29] T. Hansson, V. Oganesyan, and S. Sondhi, "Superconductors are topologically ordered," *Annals of Physics*, vol. 313, pp. 497–538, Oct. 2004.

[30] A. Zee, *Quantum field theory in a nutshell.* Princeton Oxford: Princeton University Press, second edition ed., 2010.

[31] D. Tong, "Tasi lectures on solitons," 2005.