Topological Quantum Field Theory, Anyons and Filling Anomalies

Andres Perez Fadon

In partial fulfillment of the requirements for the degree
MSc Quantum Fields and Fundamental Forces
at Imperial College London

London, England

September 2023

Supervisor Prof. Daniel Waldram

Abstract

This work presents a review of Topological Quantum Field Theory (TQFT) in the context of condensed matter. The focus is on Chern-Simons theory as the low-energy leading contribution to 2+1 dimensional actions and how, when coupled to matter, results in fractional exchange statistics. Further, examples of how TQFT is used in condensed matter are discussed, including the Fractional Quantum Hall Effect and the Filling Anomaly in Higher Order Topological Insulators.

Acknowledgements

I would like to acknowledge my supervisor, Prof. Daniel Waldram, for providing advice and resources towards the completion of this dissertation. I would also like to acknowledge Prof. Frank Schindler for his kind and helpful discussion on Higher Order Topological Insulators, and suggesting the topic of this dissertation. I would also like to thank my family and friends for their support throughout the course, who have all made this a memorable year. Special thanks to my girlfriend, Drishti, for her care and support during the duration of this dissertation and the rest of the academic year.

Contents

Chapter 1

Introduction

In the late 1800s, Lord Kelvin famously claimed that "there is nothing new to be discovered in physics now. All that remains is more and more precise measurements". Three years later, Max Planck argued that the black body ultraviolet catastrophe could be solved if light could only have discrete energies, from which the quantum revolution emerged. However, topological aspects of quantum mechanics, which is one of the most active fields of research in condensed matter today, is a much more recent concept. For instance, the geometric phase was first discovered by Shivaramakrishnan Pancharatnam in 1956 [1], but was only generalized in 1984 by Michael Berry, 84 years after the birth of quantum mechanics [2].

In this work, topological concepts in Quantum Field Theory (QFT) will be discussed and reviewed. Particular attention will be given to the Chern-Simons theory, which acts as the basis of many Topological Quantum Field Theories (TQFTs). This is a particularly important theory, as it is often the leading contribution in 2+1 dimensional systems, such as in the Fractional Quantum Hall Effect (FQHE) or magnetic Higher Order Topological Insulators (HOTIs). To see this, consider the theory of an $U(1)$ gauge 1-form valued field $A(x) = A_\mu(x)dx^\mu$ (analogous to electromagnetism) on a manifold $\mathcal{M} = \mathbb{R} \times \Gamma$, where Γ is a 2-dimensional manifold and $\mathbb{R}$ represents time. For now, pretend that Γ is closed, so that $\partial\Gamma = 0$. We may then consider what the leading term (in powers of A), of the most general action on $\mathcal{M}$ is. The Lagrangian is integrated over $\mathcal{M}$, so it must be a 3-form. The only candidate for a first-order term is $d * A$, however, this is a total derivative and integrates to 0. At second order, there are 2 candidates, $A \wedge *A$ (which is not gauge invariant), and $A \wedge dA$. For the second term, if $\partial\Gamma = 0$, meaning $\partial\mathcal{M} = 0$, then

$$S[A] \to S[A + d\chi] = \int_{\mathcal{M}} (A \wedge dA + d\chi \wedge dA), \tag{1.1}$$

$$= S[A] + \int_{\mathcal{M}} d(\chi \wedge dA) = S[A] + \int_{\partial\mathcal{M}} \chi \wedge dA = S[A], \tag{1.2}$$

where we have used $d^2 = 0$ and $\partial\mathcal{M} = 0$. Therefore, the leading term in powers of the gauge field is $S[A] = \int_{\mathcal{M}} A \wedge dA$, known as the Chern-Simons action (up to a constant prefactor). Note that the action is metric-independent, making it a topological invariant. Therefore, observables obtained from this action must also be topological invariants - hence the Chern-Simons action is a TQFT. Given that we are often interested in the low energy theory in condensed matter, the Chern-Simons action is often the dominant term in this regime. However, the system should violate parity and time reversal for this observation to be valid, as the action transforms non-trivially under those transformations. As we will see in section 3.1, the Chern-Simons theory by itself is quite unremarkable, as it predicts a vanishing field stress tensor, energy-momentum tensor, and an empty phase space. However, when coupled to matter it gives rise to interesting fractional exchange statistics, which will be discussed in detail. In fact, the Chern-Simons action will be introduced in chapter 3, not as a leading order low-energy theory, but as an action that supports charge-flux composites, which will obey fractional exchange statistics due to the Aharonov-Bohm effect.

We will start in section 2 describing anyon theories, in particular, we will consider fusion, braiding, and twisting of anyons, as well as two examples that will be relevant in our discussion of the Chern-Simons theory and the FQHE. Then, the Chern-Simons action is introduced in chapter 3. It will be shown that classically the theory is forbidden, but quantum mechanically the action must have a quantized prefactor known as the "level". The theory will then be quantized in the path integral formulation of QFT, for Abelian and non-Abelian gauge groups. We will end our discussion of the theory with observations about the Hilbert space structure. In the remainder of the thesis, two examples of topological quantum systems will be reviewed. In section 4, a gentle introduction to the Fractional Quantum Hall Effect (FQHE) will be provided, which inevitably (for Abelian states) will be described by a hierarchy of Chern-Simons terms. Then, some comments will be provided about non-Abelian states. Finally, in section 5, we will end our discussion with the filling anomaly in HOTIs, which will also be explained in terms of an effective TQFT known as the Wen-Zee action.

Chapter 2

Anyons

In this chapter, we will discuss particles with fractional exchange statistics. When two particles are exchanged, the wavefunction describing the system picks up a phase $e^{i\theta}$, where θ is known as the exchange angle. If θ is 0, then the particles are called bosons, if θ is π then the particles are called fermions, and if θ is neither, then they are called anyons. Let's begin with a false proof that anyons cannot exist in nature. Suppose $\psi(\vec{r_1}, \vec{r_2})$ is the wavefunction describing two particles. Then, let $\hat{O}$ be an operator that swaps the positions of the particles. Then,

$$\hat{O}^2 \psi(\vec{r_1}, \vec{r_2}) = \hat{O}\psi(\vec{r_2}, \vec{r_1}) = \psi(\vec{r_1}, \vec{r_2}), \tag{2.1}$$

from which we conclude that $\hat{O}^2 = \mathbb{1}$. Therefore, $\hat{O}$ has 2 distinct eigenvalues, ± 1, so only $\theta = 0$ or π are allowed. While the result is true in 3+1 dimensions, it is not a valid argument and the result is not true in 2+1 dimensions. The reason is that we cannot instantaneously switch the two particles; one must drag one into the position of the other and vice-versa, and the way this is done matters. Suppose we are in 2+1 dimensions. Then, there are 2 topologically distinct ways to switch the particles directly (without any unnecessary loops around each other) (fig. 2.1). Note that the two paths are topologically distinct, if we tried to smoothly deform one path into the other, the worldlines would intersect. This would imply that at some point the two particles are at the same position. To be more precise, assume we have translation invariance so that we are only interested in $\vec{r_2} - \vec{r_1}$. Then, the manifold of the state space of the particles is $\mathbb{R} \times \Gamma_2$, where $\Gamma_2 = \mathbb{R}^2 \setminus \{\vec{0}\}$ is the space manifold and $\mathbb{R}$ represents time. The key point is that Γ_2 is not simply connected, meaning that there are topologically non-equivalent loops. Likewise, if we followed the same analysis in 3+1 dimensions, then, the manifold would be $\mathbb{R} \times \Gamma_3$ where $\Gamma_3 = \mathbb{R}^3 \setminus \{\vec{0}\}$. The difference here is that Γ_3 is simply connected, so all paths are topologically equal, and the way particles are braided around each other does not matter. In this case, the argument above does hold and we only have bosons and fermions.

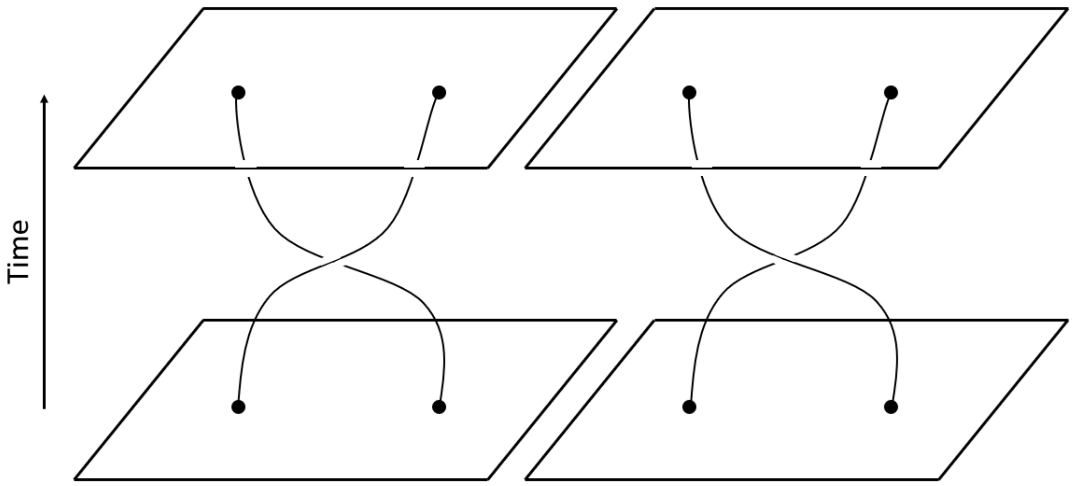


Figure 2.1: Two particles being exchanged in 2+1 dimensions (time going up). On the left, the particles are being switched by an anticlockwise rotation, while on the right the particles are exchanged by a clockwise rotation.

Considering 2+1 dimensions, we can construct an operator that switches the two particles anticlockwise, $\hat{U}$. As $\hat{U}$ is a form of time evolution, it must be unitary $\hat{U}\hat{U}^\dagger = \mathbb{1}$, so the eigenvalues are complex phases $e^{i\theta}$ where we do not have any further restrictions. Therefore, anyons are allowed in 2+1 dimensions. Nevertheless, we do live in a 3+1 dimensional universe, so fundamental particles can only be bosons or fermions. Hence, particles obeying fractional exchange statistics must be quasiparticles: collections of multiple fundamental particles, collectively behaving as one. This is analogous to how atoms in lattices collectively behave in a way that produces phonons. A further restriction is that the low energy dynamics must be constrained in one dimension. A common way to do this is simulating an infinite square well of a small length in the z-direction, so that the energy gap to the first excited eigenstate is much larger than the energy scales in the x and y directions. This, effectively, causes the dynamics to be on the xy plane, as those on the z axis are suppressed.

2.1 Anyons as Flux-Charge Composites

In this section we will construct a simple model of Abelian anyons, charge-flux composites, following closely [3, 4]. Suppose we are in 2+1 dimensions and are able to attach a flux Φ out of the plane, to a particle of charge q. For convenience, we will use the notation to refer to such a composite as (q, Φ). Now, we may consider what happens if two such particles are braided around each other as in figure 2.2. In order to discuss this we must first consider the Aharanov-Bohm effect. If a charge q moves around a loop with a flux Φ enclosed, the electron wavefunction picks up a phase $\exp[iq\Phi/\hbar]$ [5]. The particle on the left has a charge q and is moved around a loop enclosing a flux Φ. Hence, due to the Aharanov-Bohm effect, this will add a phase of $e^{iq\Phi/\hbar}$ to the wavefunction. Note that this is a topological phase, as smooth deformations of the path

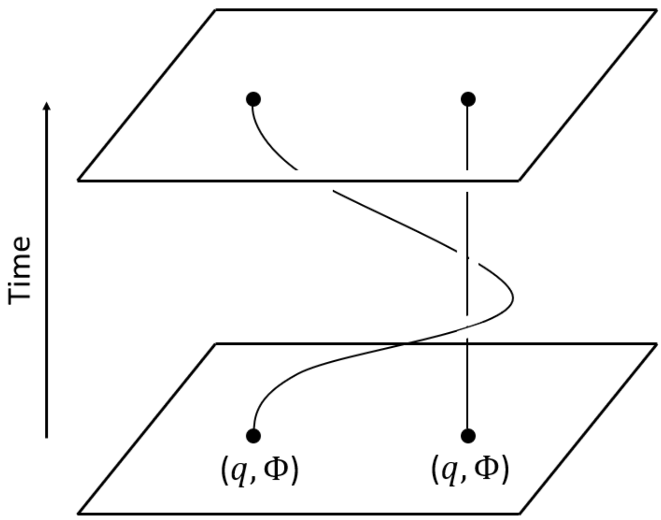


Figure 2.2: Two charge-flux composites being exchanged anticlockwise around each other.

do not change the phase. All that matters is the "winding number", the number of times the particle has gone around the flux. Naively, one would think that if a flux Φ is moved around a charge q, then the wavefunction also picks up a phase of $e^{iq\Phi/\hbar}$, which is what is happening to the composite on the right. Therefore, the total phase obtained from braiding the two charge-flux composites would be $e^{2iq\Phi/\hbar}$. However, this is famously not the case [6]; the phase obtained for a complete rotation is $e^{iq\Phi/\hbar}$. This loop is equivalent to exchanging the position of the composites twice, meaning that the exchange angle of flux-charge composites is $\theta = q\Phi/(2\hbar)$. Note that this is not restricted to either 0 or π: hence, these composites exhibit fractional exchange statistics.

An interesting property of these anyons is that they can fuse into each other. Consider a (q, Φ) and a $(-q, -\Phi)$ particle close together. The overall charge and flux are zero, making a $(0, 0)$ composite. However, this obeys trivial exchange statistics and is not composed of anything; it is the vacuum. Therefore, we conclude that a (q, Φ) and a $(-q, -\Phi)$ particle can annihilate, so we may refer to them as each other's antiparticle. Similarly, two (q, Φ) particles can fuse to a single $(2q, 2\Phi)$ particle, with an exchange angle of $\theta = (2q)(2\Phi)/(2\hbar)$. Likewise, a $(2q, 2\Phi)$ particle may "decay" into two (q, Φ) particles. In this section, we will discuss the dynamics of anyon theories, following closely [3, 7]. We will expand, and formalize this idea, as well as consider braiding of the quasi-particles.

2.2 Fusion

We can start by going back to the familiar examples of charge-flux composites. Consider two (q, Φ) particles close to each other. If the position of these two particles coincides, the system is indistinguishable from that of a single $(2q, 2\Phi)$ particle. Even if the two particles are not in the same place, a distant observer braiding around them cannot tell the difference. This will be a common theme in all anyon theories, the possibility of fusion into other particle types. However, we will have a further restriction, having a finite number of particle types. The reason for this will become apparent in section 3.7.5. In the case of charge-flux composites, this is satisfied when the double exchange angle $2\theta = q\Phi = 2\pi p/q$ (p and q coprime) is rational, which gives q particle types. This is because all we can do is take q particles of one type and fuse them to a single particle with an exchange angle $\theta = 0$, equivalent to the identity. Hence, the only particle type that exists in this theory is the particle we started with, fused with itself i times, where $i \in \{1, ..., q = 0\}$. Fusion can be represented diagrammatically as in figure 2.3. This

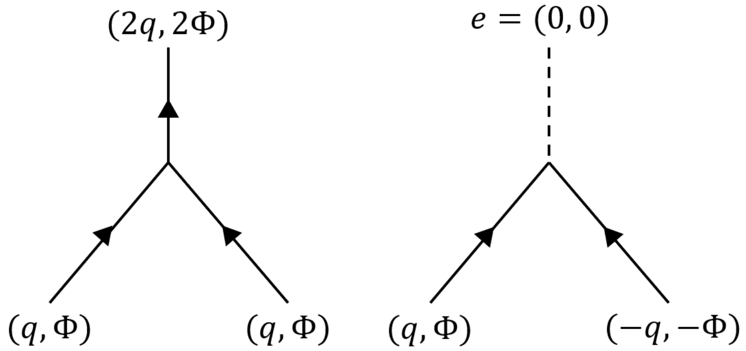


Figure 2.3: Left: Fusion of two (q, Φ) composites to a single $(2q, 2\Phi)$ composite. Right: A (q, Φ) and a $(-q, -\Phi)$ annihilating the the vacuum $\mathbf{e}$.

leads us to another feature of anyonic theories, the existence of the "vacuum" or identity. Sometimes we will draw this as a dashed line, and sometimes not at all. We can completely specify the fusion rules by defining the particle $\mathbf{n} = (nq, n\Phi)$, and the fusion rule is

$$\mathbf{n} \times \mathbf{m} = \mathbf{m} \times \mathbf{n} = \mathbf{a}, \tag{2.2}$$

$$a = (n + m) \mod q, \tag{2.3}$$

where $\times$ denotes fusion. The identity will always obey $\mathbf{e} \times \mathbf{a} = \mathbf{a} \times \mathbf{e} = \mathbf{a}$. In addition, if $\mathbf{n}$ and $\mathbf{m}$ can fuse to $\mathbf{p}$, then it must be possible for $\mathbf{p}$ to decay into $\mathbf{n}$ and $\mathbf{m}$. We can also define the notion of antiparticles. A particle of type $\mathbf{a}$ is the antiparticle of $\overline{\mathbf{a}}$ if $\mathbf{a} \times \overline{\mathbf{a}} = \mathbf{e}$. In non-Abelian theories, we will see that it is possible for two anyons to fuse by several channels. If particles $\mathbf{n}$ and $\mathbf{m}$ can fuse into two different particles, $\mathbf{p}$ and $\mathbf{q}$, we will use the notation

$$\mathbf{n} \times \mathbf{m} = \mathbf{p} + \mathbf{q}. \tag{2.4}$$

This being the case, one might wonder if antiparticles are still well-defined, as a particle anti-particle pair could fuse into something other than the identity. However, if two particles can fuse to the identity, then we will call them particle anti-particle pairs, with the restriction that each particle only has a single anti-particle. Therefore, a particle of type $\mathbf{a}$ can only annihilate to $\mathbf{e}$ with one single particle type, $\overline{\mathbf{a}}$. Particles can be their own anti-particle ($\mathbf{e}$ always is), in which case we need not draw arrows in the fusion diagram, in complete analogy to Feynman diagrams. Also in analogy to Feynman diagrams, we can think of an $\bar{\mathbf{a}}$ forward arrow as an $\mathbf{a}$ backward arrow.

A further restriction of fusion rules is the so-called "no transmutation principle", which states that if we start with a single anyon, of type $\mathbf{a}$ say, and after some time a single particle comes out, then it must be an anyon of type $\mathbf{a}$. This statement is equivalent to local anyon charge conservation.

2.3 The N-matrices and Quantum Dimensions

Let us be more general with our notation. We can describe all possible fusion channels with the N matrices,

$$\mathbf{a} \times \mathbf{b} = \sum_c N_{ab}^c \mathbf{c}, \tag{2.5}$$

where N_{ab}^c is the number of distinct ways $\mathbf{a}$ and $\mathbf{b}$ can fuse into $\mathbf{c}$. If $N_{ab}^c > 1$, then the fusion diagram should also have an index μ representing which channel happened. Consider now repeated fusion of 4 anyons. We are interested in the number of possible fusion channels, as each channel is one possible "state" of the system, so the number of fusion channels is the dimension of the Hilbert space. This will become clearer in section 3.7.5. The way to tackle this is to fuse the particles one by one,

$$\mathbf{a} \times \mathbf{b} \times \mathbf{c} \times \mathbf{d} = \sum_{\mathbf{m}_1} N_{ab}^{m_1} \mathbf{m}_1 \times \mathbf{c} \times \mathbf{d}, \tag{2.6}$$

$$= \sum_{m_1 m_2} N_{ab}^{m_1} N_{m_1 c}^{m_2} \mathbf{m}_2 \times \mathbf{d}, \tag{2.7}$$

$$= \sum_{m_1 \ldots m_3} N_{ab}^{m_1} N_{m_1 c}^{m_2} N_{m_2 d}^{m_3} \mathbf{m}_3, \tag{2.8}$$

from which we conclude that the dimension of fusing 4 anyons is

$$\dim(\mathcal{H}) = \sum_{m_1 \ldots m_3} N_{ab}^{m_1} N_{m_1 c}^{m_2} N_{m_2 d}^{m_3}. \tag{2.9}$$

Suppose that the anyons were instead created from the vacuum, $\mathbf{e}$, and we are re-fusing all particles. Then, by the no transmutation principle, the only particle that can come out is $m_3 = e$. This makes the dimension of n particles in the system

$$\dim(\mathcal{H}) = \sum_{m_1 \ldots m_{n-2}} N_{ab}^{m_1} N_{m_1 c}^{m_2} \ldots N_{m_{n-2} z}^{e}, \tag{2.10}$$

$$= \sum_{m_1 \ldots m_{n-2}} N_{ab}^{m_1} N_{m_1 c}^{m_2} \ldots \delta_{m_{n-2} \bar{z}}, \tag{2.11}$$

$$= \sum_{m_1 \ldots m_{n-3}} N_{ab}^{m_1} N_{m_1 c}^{m_2} \ldots N_{m_{n-3} y}^{\bar{z}}, \tag{2.12}$$

where we have used the fact that two anyons can fuse to the vacuum only if they are antiparticles. We will see this expression for the dimension of the Hilbert space later in section 3.7.5.

As an aside, consider a half-spin system. On its own, the Hilbert space is 2-dimensional. However, when a second half-integer spin is added to the system, the dimension of the Hilbert space doubles to 4. In general, if there are N spins, the dimension of the Hilbert space is d^N, where d is the quantum dimension, which is 2 in this case. We can define the quantum dimension d_a of particle **a** in a similar way,

$$\dim(\mathcal{H}_a^{(n)}) = d_a^n, \tag{2.13}$$

in the limit as $n \to \infty$. Writing the fusion matrix N_{ab}^c as $[N_a]_{bc}$, and noting that $N_{ab}^c = N_{ba}^c$, we can write eq. 2.9 for the fusion of multiple a particles as

$$\dim(\mathcal{H}_a^{(n)}) = \sum_b \left([N_a]_{ab}\right)^n, \tag{2.14}$$

where the product is dominated by the largest eigenvalue of N_a. Therefore, the quantum dimension d_a is the largest eigenvalue of N_a. We can also consider the eigenvector corresponding to the largest eigenvalue, $\vec{e}$,

$$N_{ab}^c e_c = d_a e_b. \tag{2.15}$$

For a matrix with non-negative entries, there is only a unique eigenvector with all positive elements, and it corresponds to the largest eigenvalue (this is known as the Perron-Frobenius theorem [8]). Therefore, we know $e_b > 0$ and, in order for eq. 2.15 to be symmetric under exchange of a and b, we must have $e_a = d_a$. Substituting this into eq. 2.15, we obtain the relationship between quantum dimensions of different anyons,

$$d_a d_b = N_{ab}^c d_c. \tag{2.16}$$

Before continuing, it is worth giving a couple of examples that will come up later.

2.3.1 Fibonacci Anyons

Here we have two particle types, the vacuum **e** and another particle denoted τ. The fusion rules are

$$\tau \times \tau = \mathbf{e} + \tau, \tag{2.17}$$

with other rules being fixed by the properties of the vacuum. As two τ particles can fuse to the identity, it is its own antiparticle. The fusion matrix for τ is given by

$$N_\tau = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \tag{2.18}$$

whose largest eigenvalue is the golden ratio,

$$d_\tau = \frac{1+\sqrt{5}}{2}. \tag{2.19}$$

2.3.2 Ising anyons

In this case, we have 3 particles, the vacuum and two non-trivial ones, often called ψ and σ. The non-trivial fusion rules are

$$\sigma \times \sigma = \mathbf{1} + \psi, \tag{2.20}$$

$$\psi \times \psi = \mathbf{1}, \tag{2.21}$$

$$\sigma \times \psi = \sigma. \tag{2.22}$$

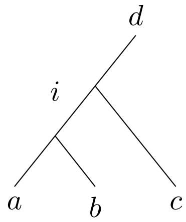
From the fusion rules, we can tell that both σ and ψ are their own antiparticles. Again, we can find the quantum dimensions, $d_\psi = 1$ and $d_\sigma = \sqrt{2}$.

2.4 The F symbols

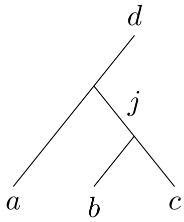
Throughout this section, we have assumed that the fusion operator $\times$ is associative. If we want to preserve this we will need a way to explain what happens as we change the order of fusions. The F symbols allow us to do just that; consider the two different ways to fuse 3 anyons,

(2.23)

Both fusion diagrams are allowed. The only difference is the order in which we fuse the particles, which we do not want to alter the end result. One can think of this as representing the same result on a different basis. As such, there should be a way to convert from one basis to another,



$$= \sum_j (F^d_{abc})_{ij}$$



(2.24)

The F symbol is called the "fusion matrix". In other words, the process on the left for a particular intermediate state i is a linear superposition of the left diagram with different intermediate states. We have also skipped a subtlety, as N^c_{ab} may be greater than 1. In this case, as explained earlier, the fusion diagram would have an index μ indicating which mode happened, and this index would also be summed over. This would give the F matrix another index, but it is not important for our analysis. Transforming a fusion diagram with an F matrix is called performing an F-move.

The F matrices must satisfy a consistency relation, creatively called the "pentagon equation", and comes from considering the fusion of 4 anyons, as shown in figure 2.4. Alge-

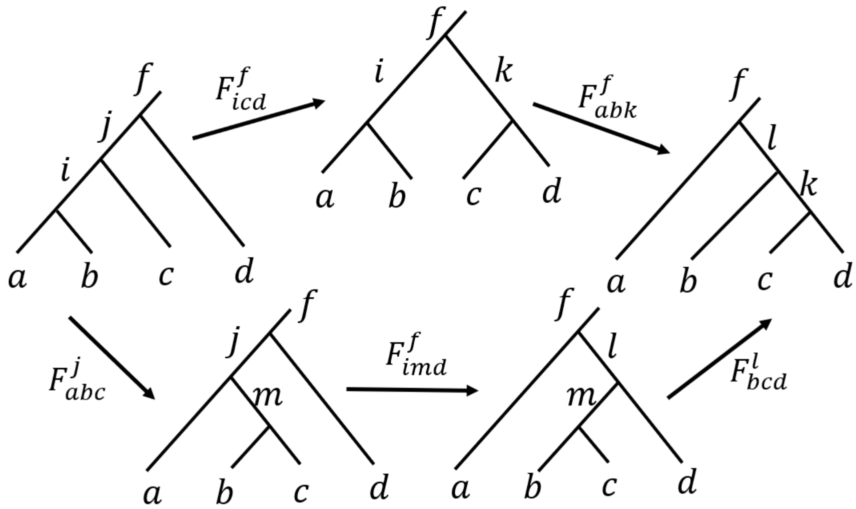


Figure 2.4: Performing various F-moves on the initial diagram (top left), we can reach another diagram (top right) in two different ways. Both ways should be equivalent, giving rise to the pentagon equation.

braically, the pentagon equation can be written as

$$(F^f_{abk})_{il}(F^f_{icd})_{jk} = \sum_m (F^l_{bcd})_{mk}(F^f_{amd})_{jl}(F^j_{abc})_{im}. \tag{2.25}$$

It turns out that if we consider more anyons, we get an equation of which the pentagon equation is a special case, so they do not cause any further restrictions. This is a consequence of MacLane's coherence theorem [9, 3]. For some theories, the F matrices are uniquely determined from the fusion rules, whereas in other theories there are multiple solutions.

2.5 Braiding, Twisting and the R matrices

We started our discussion of anyons by considering their exchange statistics in the Abelian case. In this section, we develop the theory of the braiding of non-Abelian anyons. One might ask what is non-commutative about non-Abelian anyons, the answer is that non-Abelian anyons are associated with non-Abelian gauge groups, such that the order in which particles are exchanged matters. For instance, if we have 3 non-Abelian anyons, **a**, **b** and **c**, we can braid **a** around **b** and then **c** around **b**, or **c** around **b** first and then **a** around **b**. These two processes will have distinct amplitudes in non-Abelian theories.

We will start by considering the R matrices. Consider two anyons **a** and **b**, which will eventually fuse to **c**. However, we have the option to swap the positions of the two anyons before fusion, or to fuse them straight away. We are interested in the relationship between these two processes. We can write this diagrammatically as figure 2.5 Here, whether the

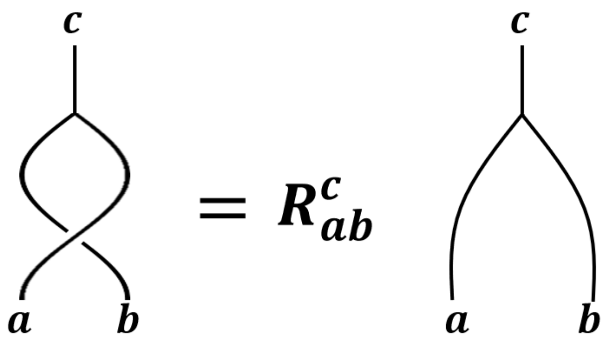


Figure 2.5: Definition of the R matrices.

anyon is Abelian or non-Abelian matters. If the multiplicity of the fusion is $N^c_{ab} = 1$ (ie the anyons are Abelian), then $R^c_{ab} \in U(1)$. We may verify this with two reasons. Firstly, $U(1)$ is an Abelian group, so if we do this multiple times, the order will not matter, which we expect from Abelian anyons. Secondly, this is something that can be physically performed, so it should be a unitary operator. Consider now the case where the multiplicity is $N^c_{ab} = 2$. Then, labelling the fusion channels without braiding $|\mu\rangle$, and with braiding $|\mu'\rangle$,

$$\begin{pmatrix} |0'\rangle \\ |1'\rangle \end{pmatrix} = R^c_{ab} \begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix}, \tag{2.26}$$

where now $R^c_{ab} \in U(2)$. As $U(2)$ is not Abelian, the order of braiding does matter. In conclusion, R^c_{ab} is an N^c_{ab} by N^c_{ab} matrix. Similar to the F matrices, there is a consistency relation that the R matrices must satisfy. This is, again, creatively known as the "hexagon equation", see figure 2.6. Again, we can mathematically express the hexagon equation as

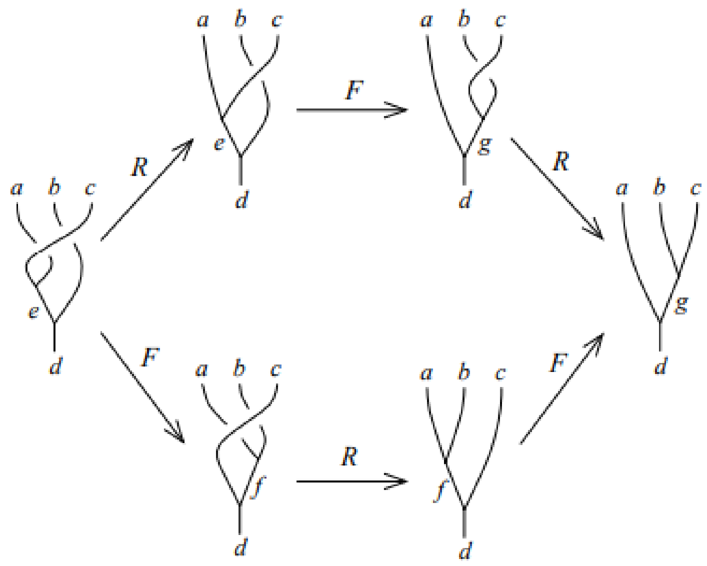


Figure 2.6: Performing various R-moves and F-moves on the initial diagram (the leftmost one), we can reach another diagram (the rightmost one) in two different ways. Both ways should give the same diagram, giving rise to the pentagon equation. Diagram obtained from [10].

$$R^k_{ac}(F^d_{bac})_{ki}R^i_{ab} = \sum_j (F^d_{bca})_{kj}R^d_{ja}(F^d_{abc})_{ji}. \tag{2.27}$$

Again, for certain theories, there might only be a single set of solutions allowed by the fusion rules, multiple solutions, or none at all. In fact, it can be proven that for any set of fusion rules, there only exists a finite number of solutions to the pentagon and hexagon equations. In other words, if we have found a set of solutions for F and R, no small deformation of them will give another solution. This is a principle known as Ocneanu rigidity [11]. It becomes apparent that the constraints given by the pentagon and hexagon equations are so strong that it makes most anyon fusion rules inconsistent. On the other hand, this very limiting property of anyon theories allows the construction of "periodic tables" for anyon theories. A list of 5 to 6 particle types can be found at [12]. So far, we have not found a modular anyon theory (those which do not have a fermion with exchange angle π) that cannot be constructed from a Chern-Simons action, or something closely related [3]. As such, the chapter 3 will be an in-depth introduction to Chern-Simons theories.

Chapter 3

Chern-Simons Theory

As discussed in the introduction, the Chern-Simons action is often a good description of the low-energy dynamics in condensed matter physics. Moreover, it provides a way for charge-flux binding to occur. This section is adapted from [13, 7, 14, 15, 16, 17]. We will start by defining the Chern-Simons action, coupling it to matter, and observing that the magnetic field associated with the vector potential looks like Dirac δ functions located at the positions of the particles. We will then quantize the theory for both Abelian and non-Abelian gauge groups and will end with some comments on the Hilbert space structure.

3.1 Abelian Chern-Simons Theory

The pure Chern-Simons action of a vector potential A on a spacetime manifold $\mathcal{M}$ is given by

$$S_{CS}[A] = \frac{k}{4\pi}\int_{\mathcal{M}} A \wedge dA = \frac{k}{4\pi}\int_{\mathcal{M}} d^3x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho, \tag{3.1}$$

where k is known as the "level" of the theory. On its own, this is quite a simple theory, as seen by the equations of motion.

$$\frac{\delta S_{CS}}{\delta A_\mu} = \frac{k}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho = \frac{k}{4\pi}\epsilon^{\mu\nu\rho}F_{\nu\rho} = 0, \tag{3.2}$$

where F is the usual field strength tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{3.3}$$

Therefore, there are no propagating degrees of freedom and k does not even affect the equations of motion. Additionally, the Hamiltonian vanishes, and by the Hamilton equations, this represents a static configuration and hence an empty phase space. We also note that the action breaks parity (under parity, the integrand picks up a negative sign) and time reversal. Hence we only expect this action to come up in systems that break parity, such as systems with a non-vanishing Hall conductivity, as we will show in section 4.2. We can couple the Chern-Simons action to matter with a current J^μ and charge q as

$$S = \int_{\mathcal{M}} d^3x \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho - qA_\mu J^\mu. \tag{3.4}$$

If there are N particles, we can write the 0th component of the 4-current J^μ as

$$j^0(\vec{x}) = \sum_{n=1}^{N}\delta(\vec{x} - \vec{x_n}), \tag{3.5}$$

where $\vec{x_n}$ is the position of the n$^{\text{th}}$ particle. The equations of motion given by the Euler-Lagrange equations are

$$\frac{\delta S}{\delta A_\mu} = \frac{k}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho - qJ^\mu = 0. \tag{3.6}$$

The 0th component of the equation reads

$$qJ^0 = q\sum_{n=1}^{N}\delta(\vec{x} - \vec{x_n}) = \frac{k}{2\pi}(\nabla \times A) = \frac{k}{2\pi}B, \tag{3.7}$$

$$B(\vec{x}) = \sum_{n=1}^{N}\Phi\delta(\vec{x} - \vec{x_n}) \tag{3.8}$$

where B is the magnetic field due to the vector potential A and $\Phi = 2\pi q/k$. Hence, the magnetic field looks like an infinitely thin flux attached to each particle. We recognise this as charge-flux composites $(q, 2\pi q/k)$, which has a double exchange angle of $2\theta = 2\pi q^2/k$ (in units where $\hbar = 1$). Interestingly, this implies that B is sourced by charges, while the other components of the equations of motion indicate that the electric field is sourced by currents, which is the opposite of what happens in electromagnetism. In addition, when particles are braided around each other, the phase picked up by the wavefunction only depends on k and the topology of the path it takes, particularly its winding number (how many particles are enclosed by the loop followed). As we will see, it is convenient to think of the flux lines as loops closing at infinity, as this reveals that the Chern-Simons theory computes the "linking number" of these knots. This gives us intuition about why Chern-Simons is a topological theory; the action is a topological invariant of the spacetime manifold, so the only quantities that can be calculated from it are themselves topological invariants.

3.2 Gauge Invariance and Quantization of the Level

For the rest of the chapter, we will only consider the pure Chern-Simons action, without coupling to other fields. This section is adapted from [13]. It becomes apparent that even in this special case, the action is not gauge invariant. Although we seem to have shown that the action is invariant by considering transformations of the form $A \to A + d\chi$ in eq. 1.2, this was only under the assumption that we had a closed spatial manifold. If this is not the case, then the action changes by a boundary term. However, in the absence of magnetic monopoles, the boundary term vanishes by Gauss' law. Therefore, we shall now consider the simplest case in which we do not have gauge invariance. As we will see, this implies that the classical theory is not well defined, but at the quantum level, we just need the level k to be quantized.

We will start by compactifying the spatial manifold to S^2 and perform a Wick rotation into Euclidean signature $t \to \tau = it$. As the Chern-Simons action is of first order in derivatives, it transforms under a Wick rotation by $S_{CS} \to S_{CS}^E = -iS_{CS}$. Now we can choose a convenient gauge transformation $\chi = 2\pi\tau/\beta$. The Wick rotation turns the time axis $\mathbb{R}$ into a thermal circle [3, 14] S^1 with periodicity β, see appendix A for more information. The gauge transformation of this form is chosen because it cannot be continuously deformed to the identity map, as it winds around the thermal circle and is hence non-contractible. These kinds of gauge transformations are often called "large gauge transformations". Before performing this gauge transformation, let us rewrite the action as

$$S_{CS}^E = -i\frac{k}{4\pi}\int_{S^1\times S^2} d^3x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho, \tag{3.9}$$

$$= -i\frac{k}{4\pi}\int_{S^1\times S^2} d^3x (A_0F_{12} + A_1F_{20} + A_2F_{01}). \tag{3.10}$$

where x is now Euclidean. Now, it is important to be careful, as there is a famous factor of 2 which is easy to miss. Consider

$$\int_{S^1\times S^2} d^3x A_1\partial_2(\partial_0\chi). \tag{3.11}$$

One might be tempted to claim this integral vanishes, as $\partial_0\chi$ does not depend on the variable with respect to which we are differentiating. However, as χ is in a topologically non-trivial configuration, this may not hold. To see this, we can integrate by parts to obtain

$$-\int_{S^1\times S^2} d^3x (\partial_2 A_1)(\partial_0\chi), \tag{3.12}$$

which certainly does not vanish (note there is no boundary term after integrating by parts as $\partial(S^1 \times S^2) = 0)$. Hence, we must integrate by parts to obtain

$$\int_{S_1 \times S^2} d^3x A_1 F_{20} = \int_{S^1 \times S^2} d^3x (-\partial_2 A_1) A_0 - A_1(\partial_0 A_2), \tag{3.13}$$

$$\int_{S_1 \times S^2} d^3x A_2 F_{01} = \int_{S^1 \times S^2} d^3x (\partial_1 A_2) A_0 + A_2(\partial_0 A_1). \tag{3.14}$$

Note that the sum of these terms makes another factor of $A_0 F_{12}$, which lets us rewrite the Euclidean Chern-Simons action as

$$S^E_{CS} = -i\frac{k}{2\pi} \int_{S^1 \times S^2} d^3x (A_0 F_{12}) + \frac{1}{2}(A_2 \partial_0 A_1 - A_1 \partial_0 A_2). \tag{3.15}$$

Now we can perform the gauge transformation $A_0 \to A_0 + \frac{2\pi}{\beta}$,

$$S^E_{CS} \to S^{E\prime}_{CS} = S^E_{CS} - i\frac{k}{\beta} \int_{S^1 \times S^2} d^3x F_{12}, \tag{3.16}$$

$$= S^E_{CS} - i\frac{k}{\beta} \int_{S^1} d\tau \int_{S^2} d^2x F_{12}. \tag{3.17}$$

The integral over a closed surface of F_{12} is simply $2\pi n$ where n is an integer, due to the Dirac quantization condition, which we will see several times. This quantization is deeply tight with the fact that we have a compact gauge group, and is equivalent to placing n magnetic monopoles inside the integration domain. The key point is that n is an integer, this analysis is not based on the existence of magnetic monopoles, just on the fact that this integral must be quantized.

$$S^{E\prime}_{CS} = S^E_{CS} - i\frac{2\pi n k}{\beta} \int_{S^1} d\tau \tag{3.18}$$

Finally, the thermal circle has periodicity β, so the integral over S^1 is just β. Hence, going back to Minkowski signature,

$$\delta S_{CS} = 2\pi n k. \tag{3.19}$$

Therefore, we have shown that the Chern-Simons action is not gauge invariant. This means that classically, the theory is not consistent. However, at the quantum level, not all is lost. Observables are obtained from the path integral $\mathcal{Z}$, so as long as this is invariant, the theory is gauge invariant.

$$\mathcal{Z} = \int \mathcal{D}A e^{iS_{CS}[A]} \to \int \mathcal{D}A e^{iS_{CS}[A] + 2\pi n k} = e^{2\pi n k} \mathcal{Z}, \tag{3.20}$$

so as long as the prefactor is 1, the theory is saved. This happens when the exponent is a multiple of 2π, which forces the level k to be an integer. In conclusion, we have used gauge invariance to show that the Chern-Simons level is quantized.

3.3 Anyons and Wilson loops

We already saw in the introduction to this chapter how the Chern-Simons action allows for a simple model of charge-flux composites obeying fractional statistics. In this subsection, we will rediscover anyonic statistics from a more formal perspective [18]. Consider the Chern-Simons path integral with a source J^μ,

$$\mathcal{Z}[J] = \int \mathcal{D}A \exp\left[i \int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho + A_\mu J^\mu\right)\right]. \tag{3.21}$$

We are interested in the dynamics of two particles moving and braiding around each other. Let the spatial path in $\mathbb{R}^2$ of the particles be γ_1 and γ_2, both of which are closed. Let the curve γ_a be parameterized by $x_a^\mu(t)$ ($x_a^0(t) = t$). Then, we can write the source corresponding to particle a as

$$J_a^\mu(\vec{x}, t) = \dot{x}_a^\mu(t)\delta^{(2)}(x - x_a(t)), \tag{3.22}$$

$$J^\mu = J_1^\mu + J_2^\mu, \tag{3.23}$$

which makes the source term

$$\int_{\mathcal{M}} d^3x A_\mu J^\mu = \sum_a \int_{\mathcal{M}} d^3x A_\mu \dot{x}_a^\mu \delta^{(2)}(x - x_a(t)), \tag{3.24}$$

$$= \sum_a \int dt \frac{dx_a^\mu}{dt} A_\mu(x_a(t)), \tag{3.25}$$

$$= \sum_a \oint_{\gamma_a} dx_a^\mu A_\mu(x_a(t)). \tag{3.26}$$

Defining non-local observables known as Wilson loops,

$$W_a = \exp\left[i \oint_{\gamma_a} dx_a^\mu A_\mu\right], \tag{3.27}$$

we can write the sourced path integral as

$$\mathcal{Z}[J] = \int \mathcal{D}A W_1 W_2 e^{iS_{CS}[A]} \equiv \langle W_1 W_2 \rangle. \tag{3.28}$$

As the path integral is Gaussian, it can be solved exactly by a field redefinition $A_\mu \to A_\mu^{cl} + a_\mu$, such that $\mathcal{D}A = \mathcal{D}a$ and A_μ^{cl} obeys the classical equation of motion,

$$\frac{k}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho^{cl} + J^\mu = 0. \tag{3.29}$$

The path integral transforms as

$$\mathcal{Z}[J] = \int \mathcal{D}A \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}(A^{cl}_\mu + a_\mu)\partial_\nu(A^{cl}_\rho + a_\rho) + (A^{cl}_\mu + a_\mu)J^\mu\right)\right], \tag{3.30}$$

$$= \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu A^{cl}_\rho + A^{cl}_\mu J^\mu\right]\right) \tag{3.31}$$

$$\int \mathcal{D}a \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + a_\mu J^\mu + \frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu A^{cl}_\rho + \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu a_\rho\right)\right],$$

$$= \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu A^{cl}_\rho + A^{cl}_\mu J^\mu\right)\right] \tag{3.32}$$

$$\int \mathcal{D}a \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + a_\mu J^\mu - \frac{1}{2}a_\mu J^\mu - \frac{k}{4\pi}\epsilon^{\mu\nu\rho}\partial_\nu A^{cl}_\mu a_\rho\right)\right],$$

$$= \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu A^{cl}_\rho + A^{cl}_\mu J^\mu\right)\right] \tag{3.33}$$

$$\int \mathcal{D}a \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + \frac{1}{2}a_\mu J^\mu + \frac{k}{4\pi}\epsilon^{\rho\nu\mu}a_\rho\partial_\nu A^{cl}_\mu\right)\right],$$

$$= \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu A^{cl}_\rho + A^{cl}_\mu J^\mu\right)\right] \tag{3.34}$$

$$\int \mathcal{D}a \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho\right)\right].$$

However, we are only interested in

$$\frac{\mathcal{Z}[J]}{\mathcal{Z}[0]} = \exp\left[i\int_{\mathcal{M}} d^3x \left(\frac{k}{4\pi}\epsilon^{\mu\nu\rho}A^{cl}_\mu\partial_\nu A^{cl}_\rho + A^{cl}_\mu J^\mu\right)\right]. \tag{3.35}$$

Therefore, the generating functional takes the form $e^{iS[A^{cl}]+iA^{cl}_\mu J^\mu}$, as we expect from a quadratic action. Substituting in the form of A^{cl}_μ in Lorenz gauge, we obtain after a lot of algebra,

$$A^{cl}_\mu(x) = \frac{1}{2k}\int_{\mathcal{M}} d^3y\epsilon_{\mu\nu\rho}\frac{\partial^\nu J^\rho}{|x-y|} = \frac{1}{2k}\sum_a \oint_{\gamma_a} dx^\nu_a\epsilon_{\mu\nu\rho}\frac{(x-x_a)^\rho}{|x-x_a|^3}, \tag{3.36}$$

$$\mathcal{Z}[J] = \langle W_1 W_2\rangle, \tag{3.37}$$

$$= \exp\left(\frac{i}{2k}\oint_{\gamma_1} dx^\mu_1 \oint_{\gamma_2} dx^\nu_2\epsilon_{\mu\nu\rho}\frac{(x_1-x_2)^\rho}{|x_1-x_2|^3}\right), \tag{3.38}$$

$$= \exp\left(\frac{2\pi i}{k}\Phi[\gamma_1,\gamma_2]\right), \tag{3.39}$$

Where

$$\Phi[\gamma_1,\gamma_2] = \frac{1}{4\pi}\oint_{\gamma_1} dx^\mu_1 \oint_{\gamma_2} dx^\nu_2\epsilon_{\mu\nu\rho}\frac{(x_1-x_2)^\rho}{|x_1-x_2|^3} \tag{3.40}$$

is the "linking number" of γ_1 and γ_2, an integer that counts the number of times one curve winds around the other [19]. In reality, there are two kinds of terms, integrals over γ_1 and γ_2, and integrals over γ_a twice. The latter are self-interaction terms that diverge but can be regularized, we will discuss the regularization procedure in section 3.4. When $\Phi[\gamma_1, \gamma_2] = 1$, i.e. the curves wind around each other once, $\langle W_1 W_2 \rangle = e^{2\pi i/k}$, the same anionic phase we obtained while arguing that the Chern-Simons action generates flux-charge composites. The main conclusion is that since the action is a topological invariant, the observables obtained from it are also topological quantities, in this case being knot invariants of the worldlines of two sources.

The analysis above can be generalized to r point sources braiding around each other. Let particle a follow a closed path γ_a, and define a "link" to be the disjoint union of all paths. Before we continue, recall that if A is a G gauge field, then we can decompose its components in terms of the generators T^C of the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, $A_\mu = A^c_\mu T^C$. Taking the gauge group to be $G = U(1)$, the only generator of the Lie algebra $\mathfrak{u}(1)$ is a single integer n. Hence, we can write $A_\mu = A^1_\mu n$, but we will drop the group index as we only have one generator. Using this, we can define the "Wilson Link" as

$$W[L] = \prod_{a=1}^{r} W_a[\gamma_a] = \prod_{a=1}^{r} \exp\left(n_a \oint_{\gamma_a} dx^\mu_a A_\mu \right). \tag{3.41}$$

Note that in the non-Abelian case, we will need a path-ordering operator, as will be discussed in section 3.6. Using the Wilson Link, we can rewrite the path integral as

$$\mathcal{Z}[J] = \langle W[L] \rangle = \int \mathcal{D}A W[L] e^{iS_{CS}[A]}, \tag{3.42}$$

$$= \exp\left(\frac{i}{2k} \sum_{a,b=1}^{r} n_a n_b \oint_{\gamma_a} dx^\mu_a \oint_{\gamma_b} dx^\nu_b \epsilon_{\mu\nu\rho} \frac{(x_a - x_b)^\rho}{|x_a - x_b|^3} \right), \tag{3.43}$$

$$= \exp\left(\frac{2\pi i}{k} \sum_{a,b=1}^{r} n_a n_b \Phi[\gamma_a, \gamma_b] \right). \tag{3.44}$$

However, we note again that $\mathcal{Z}$ contains terms involving $\Phi_s[\gamma_a] = \Phi[\gamma_a, \gamma_a]$ (known as the "self-linking number"), which diverges. While earlier we simply ignored these terms, it is possible to regulate them through a process called framing.

3.4 Framing

For any curve $\gamma_a \subset L$, we can choose a vector field perpendicular to γ_a everywhere, a choice known as framing (see figure 3.1). We can define a closed loop γ'_a by moving every point in γ_a in the direction of the vector field. As the curves γ_a and γ'_a don't intersect by construction (as long as the distance by which every point is moved is small and γ_a is smooth), the self-linking number can be redefined to be $\Phi_s[\gamma_a] = \Phi[\gamma_a, \gamma'_a]$, which is finite and well defined.

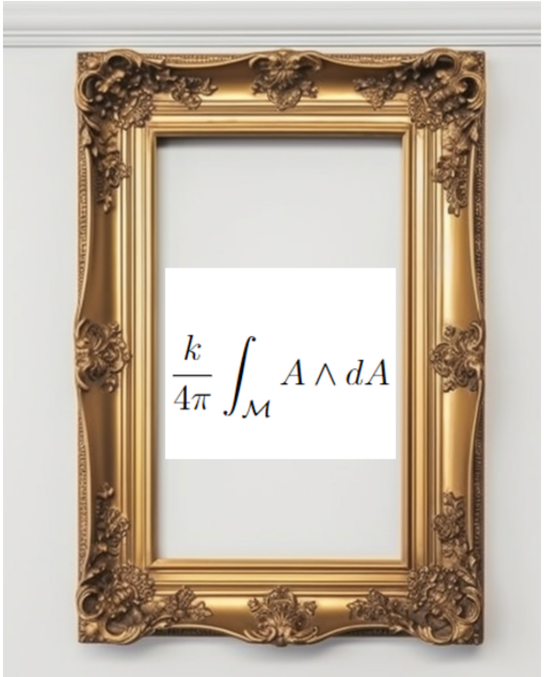


Figure 3.1: Framing the Chern-Simons theory.

However, we note that $\Phi_s[\gamma_a]$ depends on the topology of the vector field being used to extend γ_a into γ'_a. This is problematic, so we need a description of how $\Phi_s[\gamma_a]$ transforms under a change of framing. As Φ can only take integer values, the difference between two choices of framing must also be an integer t. This integer is the difference in twists that the two vector fields have around γ_a, see figure 3.2. From this, we conclude that under a

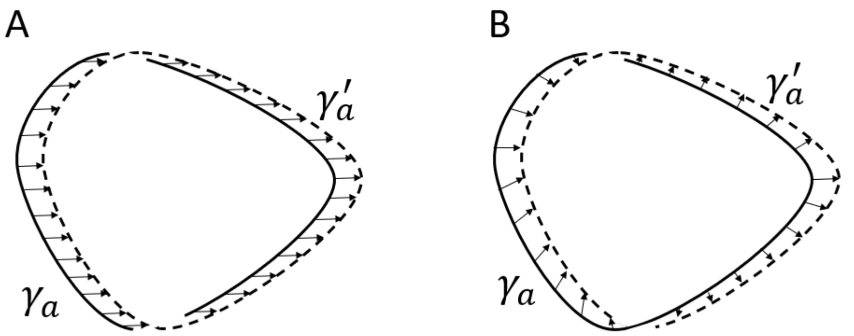


Figure 3.2: Two different choices of framing. The solid line represents γ_a, the dashed line γ'_a, and the arrows the vector field used to generate γ'_a. A: γ'_a does not wind around γ_a, so $\Phi_s[\gamma_a] = 0$. B: γ'_a winds around γ_a once, so $\Phi_s[\gamma_a] = 1$. Therefore, if we changed framing from A to B, the integer $t = 1$.

change of framing,

$$\Phi_s[\gamma_a] \to \Phi_s[\gamma_a] + t, \tag{3.45}$$

$$\langle W[L]\rangle \to \exp\left(\frac{2\pi i t n_a^2}{k}\right). \tag{3.46}$$

To be precise, we say that $\mathcal{Z} = \langle W[L] \rangle$ is a topological invariant of framed manifolds. Framing is an important concept, as it makes the theory anomalous, and hence certain Noether currents will not be conserved.

3.5 Aside: Chern-Simons Theory in a Torus

As the Chern-Simons action is a topological invariant, we expect the phase-space to only depend on the topology of $\mathcal{M}$. This happens to be true: the Chern-Simons phase space is completely characterized by the class of topologically non-trivial Wilson loops. For example, in $\mathcal{M} = \mathbb{R} \times \mathbb{R}^2$, there are no non-trivial Wilson loops, as $\mathcal{M}$ is simply connected. Therefore, the space of allowed non-trivial Wilson loops is 0 dimensional. This corresponds to the fact that the phase space is empty in this spacetime manifold. We already saw that the equation of motion predicts no propagating degrees of freedom, but it is easier to observe that the Hamiltonian vanishes. Therefore, through Hamilton's equations, we reach the conclusion that $\dot{q} = \dot{p} = 0$. In contrast, consider a Chern-Simons field on the surface of a torus, $\mathcal{M} = \mathbb{R} \times T^2$. Then, the space of non-trivial Wilson loops is 2-dimensional. One can define a Wilson loop around the handle (call this configuration $\mathbf{h}$), or around the azimuthal direction ($\mathbf{a}$). Then, a general non-trivial Wilson loop can be written as $W = h\mathbf{h} + a\mathbf{a}$, $(h, a \in \mathbb{Z})$, meaning h loops around the handle and a loops around the azimuthal direction. Therefore, we conclude $\pi_1(\mathcal{M}) = \pi_1(T^2) = \mathbb{Z} \oplus \mathbb{Z}$. The two Wilson loops are $U(1)$-valued, as seen by eq. 3.27. Therefore, the Chern-Simons phase space is $\mathcal{P} = U(1) \times U(1) = T^2$. Another important property of Chern-Simons in topologically non-trivial spaces is the ground state degeneracy. To see this, consider the equal time commutation relation (working in Coulomb gauge where A_0 vanishes),

$$[A_1(\vec{x}), A_2(\vec{y})] = \frac{2\pi i}{k} \delta^{(2)}(\vec{x} - \vec{y}). \tag{3.47}$$

Defining the two non-contractible loops as γ_1 and γ_2, and the two non-trivial Wilson loops as $w_i = \oint_{\gamma_i} dx^j A_j$, $W_i = \exp(i w_i)$, we obtain the following commutation relation,

$$[w_1, w_2] = \frac{2\pi i}{k}, \tag{3.48}$$

$$W_1 W_2 = e^{iw_1} e^{iw_2}, \tag{3.49}$$

$$= e^{[w_1, w_2]} e^{iw_2} e^{iw_1}, \tag{3.50}$$

$$= e^{2\pi i/k} W_2 W_1, \tag{3.51}$$

where eq. 3.50 comes from the Baker–Campbell–Hausdorff formula. It turns out that the smallest representation of the algebra followed by the Wilson loops, which we just derived, has dimension k. This is the ground-state degeneracy on the torus. In general, in a genus g surface, the ground state degeneracy is k^g.

3.6 Non-Abelian Chern-Simons

In this section, we will attempt to quantize the Chern-Simons action for an arbitrary gauge group G and ultimately fail. The reason is that the classical theory has a symmetry, diffeomorphism invariance (being independent of the metric), which is broken at the quantum level when we try to fix a gauge through the Faddeev–Popov procedure. This is due to the newly introduced fields being inevitably coupled with a metric. The solution is to add a metric-dependent term, the gravitational Chern-Simons action, which will cancel the metric dependence in the path integral. This section is adapted from a combination of [13] and [14].

3.6.1 Gauge Invariance and the Quantization of the Level

In order to generalize the concepts seen in the previous section to non-Abelian gauge groups, we must use the language of differential geometry and bundles. For an introduction to bundles, refer to appendix B. As before, we need a 3-manifold $\mathcal{M}$ representing spacetime, and a compact Lie group G, with Lie algebra $\mathfrak{g}$. Consider a principal G-bundle with projection $\pi : E \to \mathcal{M}$, where E is the total space and $\mathcal{M}$ is the root space. Sections of E are maps $g : \mathcal{M} \to E$ which smoothly assign to each $x \in \mathcal{M}$ a point in E. We call the g maps gauge transformations, which can be thought of as assigning a group element to every point in spacetime. We can also consider infinitesimal generators of $g(x)$, which are in $\mathfrak{g}$, and so we think of $g(x)$ as a $\mathfrak{g}$-valued function. Principal connections on E are $\mathfrak{g}$-valued 1-forms on $\mathcal{M}$, $A(x) = A(x)_\mu dx^\mu$, which is precisely our gauge field. As $A_\mu \in \mathfrak{g}$, we can expand A_μ in terms of a basis in $\mathfrak{g}$, $A_\mu = A_\mu^a T^a$, where T^a are the generators of the Lie algebra and a is the gauge group index. Now, we can ask what happens to A_μ under gauge transformations, or in other words, what happens to the connection under G-actions? We proceed in the usual way, by requiring the covariant derivative $D_\mu = \partial_\mu + [A_\mu, \cdot\,]$ to be invariant (similar to the process in Yang-Mills theory), and we obtain

$$A_\mu \to A'_\mu = g^{-1} A_\mu g + g^{-1} \partial_\mu g. \tag{3.52}$$

We already saw for the Abelian case that gauge invariance is problematic for the Chern-Simons action. Hence, we should check how the non-Abelian action transforms under gauge transformations,

$$\begin{aligned} S_{CS}[A] &= \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{Tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right), &(3.53)\\ &= \frac{k}{4\pi} \int_{\mathcal{M}} d^3x \epsilon^{\mu\nu\rho}\, \mathrm{Tr} \left(A_\mu (\partial_\nu A_\rho - \partial_\rho A_\nu) + \frac{2}{3} A_\mu [A_\nu, A_\rho] \right), &(3.54) \end{aligned}$$

which after a lot of algebra simplifies to

$$\delta S_{CS}[A] = \frac{k}{4\pi}\int_{\mathcal{M}} d^3x\epsilon^{\mu\nu\rho}\left(\partial_\mu \operatorname{Tr}\left[(\partial_\nu g)(g^{-1}A_\rho)\right] + \frac{1}{3}\operatorname{Tr}\left[(g^{-1}\partial_\mu g)(g^{-1}\partial_\nu g)(g^{-1}\partial_\rho g)\right]\right), \tag{3.55}$$

$$= \frac{k}{4\pi}\int_{\mathcal{M}} d^3x\epsilon^{\mu\nu\rho}\partial_\mu \operatorname{Tr}\left[(\partial_\nu g)(g^{-1}A_\rho)\right] + 2\pi k w(g) \tag{3.56}$$

The first term is the total derivative we saw in the Abelian case, and

$$w(g) = \frac{1}{24\pi^2}\int_{\mathcal{M}} d^3x\epsilon^{\mu\nu\rho}\operatorname{Tr}\left[(g^{-1}\partial_\mu g)(g^{-1}\partial_\nu g)(g^{-1}\partial_\rho g)\right] \tag{3.57}$$

is the "winding number" of g, which is quantized [20]. In particular, it is an integer up to a constant prefactor, which we can absorb into the normalization of the trace (here, by Tr we mean a multiple of the Killing form on $\mathfrak{g}$, which can be fixed to make $w(g) \in \mathbb{Z}$). In conclusion, $S_{SC} \to S_{CS} + 2\pi k w(g)$, and by the same argument as in the Abelian case, we must preserve the path integral, and this forces $k \in \mathbb{Z}$. When $w(g) \neq 0$, we call g a "large gauge transformation", as it wraps non-trivially around the gauge group.

3.6.2 Wilson Loops and the Path Integral

We have seen that the path integral can be rewritten as the expectation value of Wilson loops. However, Wilson loops for non-Abelian gauge groups are slightly more complicated. First, fix a representation R of G, and let γ be a closed curve in $\mathcal{M}$. We can then define Wilson loops as

$$W_R[\gamma] = \operatorname{Tr}_R\left[\mathcal{P}\exp\left(i\oint_\gamma dx^\mu A_\mu\right)\right], \tag{3.58}$$

where $\mathcal{P}$ is the path-ordering operator. This expression is analogous to the propagator $T\exp\big(-i\int dtH\big)$, T being the time-ordering operator. The path-ordering operator is needed as the generators of the group at different points of the path do not commute. Hence, eq 3.58 is just the trace of

$$(1 + iA^a_{\mu_1}(x_1)T^a dx_1^{\mu_1})(1 + iA^b_{\mu_2}(x_2)T^b dx_2^{\mu_2}) \ldots (1 + iA^z_{\mu_n}(x_n)T^z dx_n^{\mu_n}), \tag{3.59}$$

where we have divided the path γ into small steps of length dx_i. From this, we can define a link as the disjoint union of several paths, $L = \cup_{a=1}^r \gamma_a$, and a Wilson link as

$$W[L] = \prod_{a=1}^r W_{R_a}[\gamma_a]. \tag{3.60}$$

Finally, we can write the path integral for a general simple Lie gauge group G as

$$\mathcal{Z}[A, L] = \langle W[L] \rangle = \int \mathcal{D}A e^{iS_{CS}[A]} W[L]. \tag{3.61}$$

Witten proved that the expectation value of Wilson loops is a knot invariant, something we expect from a topological theory. In particular, for $SU(N)$ it can be shown to be the HOMFLY polynomial, which in the special case of $SU(2)$ becomes the Jones polynomial [14, 21]. Note that the partition function is no longer Gaussian, as there are terms of order 3 in powers of the gauge field. Therefore, the path integral cannot be evaluated explicitly and we estimated it through a saddle point approximation. We will consider the large k limit, as the exchange statistics go like $1/k$ and become small in this limit. The path integral can be estimated as the sum of contributions near the stationary phase ($\delta S_{CS} = 0$). Here we already run into trouble, as it is hard to solve the equations of motion in the first place. Nevertheless, in order to progress, let A^α be the αth solution to the classical equations of motion. For convenience, also define $S_{CS}[A] = kI_{CS}[A]$, which makes our first estimate of the path integral

$$\mathcal{Z} \approx \sum_\alpha \mu[A^\alpha] = \sum_a e^{ikI[A^\alpha]}, \tag{3.62}$$

where $\mu[A^\alpha]$ is the saddle point contribution of A^α. Next, we can improve the calculation of $\mathcal{Z}$ to include fluctuations around the classical solutions. As usual, perform a field redefinition $A_\mu = A_\mu^\alpha + a_\mu$, where $\mathcal{D}A = \mathcal{D}a$. This makes the action

$$S_{CS}[A] = kI[A^\alpha] + \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{Tr}(a \wedge Da) + ..., \tag{3.63}$$

$$= kI[A^\alpha] + \frac{k}{4\pi} \int_{\mathcal{M}} d^3x \epsilon^{\mu\nu\rho} \,\mathrm{Tr}(a_\mu D_\nu a_\rho) + ..., \tag{3.64}$$

where $D_\mu = \partial_\mu + [A_\mu^\alpha, \cdot]$ is the covariant derivative. Now, the action is approximately quadratic, so the path integral is Gaussian and can be solved. However, first we have to fix the gauge on a, which we will do through the Faddeev–Popov gauge fixing procedure. We will work in Lorenz gauge $D_\mu a^\mu = 0$, which we can force in the path integral by adding Faddeev-Popov-DeWitt ghosts (see appendix C for the derivation),

$$S[a] \to S[a] + S_{gf}[\phi, a] + iS_{gh}[\bar{c}, C], \tag{3.65}$$

$$= \int_{\mathcal{M}} \mathrm{Tr} \left(\frac{k}{4\pi} a \wedge Da + \phi * D * a + i\bar{c} D * Dc \right), \tag{3.66}$$

$$= \int_{\mathcal{M}} d^3x \,\mathrm{Tr} \left(\frac{k}{4\pi} \epsilon^{\mu\nu\rho} a_\mu D_\nu a_\rho + \phi D_\mu a^\mu + i\bar{c} D_\mu D^\mu c \right), \tag{3.67}$$

where ϕ, $\bar{c}$, c are ghost fields, and c and $\bar{c}$ anticommute. Most importantly, note that the ghost fields are coupled to a metric. As we will see later, this breaks diffeomorphism invariance, which will come back to haunt us. The saddle point contribution becomes

$$\mu[A^\alpha] = e^{ikI[A^\alpha]} \int \mathcal{D}a\mathcal{D}\phi\mathcal{D}\bar{c}\mathcal{D}c \exp\left[i\int_{\mathcal{M}} \mathrm{Tr}\left(\frac{k}{4\pi}a\wedge Da + \phi * D * a + i\bar{c}D * Dc\right)\right]. \tag{3.68}$$

Note that the $\bar{c}$ and c fields are decoupled from the others and can be directly evaluated,

$$\int \mathcal{D}\bar{c}\mathcal{D}c \exp\left[-\int_{\mathcal{M}} \mathrm{Tr}\left(\bar{c}D_\mu D^\mu c\right)\right] = \det(D_\mu D^\mu), \tag{3.69}$$

where we note that $\bar{c}$ and c are Grassmann variables. The rest of the path integral is more complicated. It turns out that the first two terms can be evaluated to $1/\sqrt{\det(L_-)}$, where $L = *D + D*$ is an adjoint operator which maps odd forms into even forms and vice-versa, and L_- is the restriction of L which maps odd forms to even forms. To see this, consider $H = (a, \phi)$, taking a to be a 1-form and ϕ a 3-form. Then, consider the following

$$H\wedge *L_-H = H\wedge(D + *D*)H = H\wedge DH + H\wedge *D*H \tag{3.70}$$

$$= a\wedge Da + \phi\wedge *D*a, \tag{3.71}$$

where we have used $** = 1$ in an odd-dimensional manifold, and in that the result must be a 3-form in order to be integrated. The terms in eq. 3.71 are the only ones that make 3-forms. Noting that eq. 3.71 is the remaining of the action up to rescaling of a and ϕ, (and that $*D*a$ is a 0-form, so the $\wedge$ is not needed), we can write the rest of the action as

$$\int \mathcal{D}H \exp\left[\frac{i}{2}\int_{\mathcal{M}} \mathrm{Tr}\left(H\wedge *(L_-H)\right)\right] = \frac{1}{\sqrt{\det(L_-)}}. \tag{3.72}$$

Bringing this together,

$$\mu[A^\alpha] = e^{ikI[A^\alpha]}\frac{\det(D_\mu D^\mu)}{\sqrt{\det(L_-)}}. \tag{3.73}$$

However, now a few problems arise. Firstly, $\mu[A^\alpha]$ depends on the metric. This should not be surprising as $\mathcal{Z}$ itself, once the ghosts are introduced, depends on the metric. Secondly, the ratio of determinants is complicated. It turns out that the absolute value is the so-called “Ray–Singer torsion” [22], T_α of the connection A^α. However, the phase will have to be regularized. Note that as $D_\mu D^\mu$ is positive and self-adjoint, $\det(D_\mu D^\mu) \in \mathbb{R}_+$, but the phase of $\det(L_-)$ is more complicated, as it diverges [14]. However, it can be regularized, after which the phase depends on the signature of L_-, the difference between positive and negative eigenvalues. Unfortunately, this is also not well defined, but it can be regularized once again, making the phase $e^{i\pi\eta[A^\alpha]/2}$, where the “eta invariant” is given by

$$\eta[A^\alpha] = \frac{1}{2} \lim_{s\to 0} \sum_{\lambda\neq 0} \exp\left(\frac{i\pi}{4}\sum_j \mathrm{sign}(\lambda_j)\right). \tag{3.74}$$

where λ_j are the eigenvalues of the L_- operator. Using the Atiyah–Patodi–Singer index theorem, we can rewrite the eta invariant in terms of other quantities,

$$\frac{1}{2}\left(\eta[A^\alpha] - \eta[0]\right) = \frac{c_2(G)}{2\pi} I[A^\alpha], \tag{3.75}$$

where $c_2(G)$ is a topological invariant known as the quadratic Casimir of G. The path integral becomes

$$\mathcal{Z} \approx e^{i\pi\eta[0]/2} \sum_\alpha \exp\left[i\left(k + \frac{c_2(G)}{2}\right) I[A^\alpha]\right] T_\alpha. \tag{3.76}$$

Nevertheless, the form of the path integral is not particularly important. The thing to note is that it depends on $\eta[0]$, which is not a topological invariant and depends on the metric. Therefore, the path integral is not diffeomorphism invariant, meaning that a classical symmetry is broken at the quantum level, so the theory is anomalous. This anomaly is known as the "framing anomaly" but note this is not the same framing as the one we discussed for regularizing $\Phi_s[\gamma]$.

3.6.3 The Framing Anomaly

In order to obtain a theory whose path integral is a topological invariant, we have to add a counterterm to the original Chern-Simons action. As we do not want to change the dynamics of the vector potential, the counterterm should not contain A itself. We also require the contribution of this counterterm to be metric-dependent and to exactly cancel the metric dependence on $\mathcal{Z}$ to obtain a topological invariant. We shall call this term the "gravitational counterterm" as it depends on the metric. The term we are looking for is known as the "gravitational Chern-Simons action", as it has the same form, but is made up of the Levi-Civita connection ω instead of the vector potential,

$$I[g] = \frac{1}{4\pi}\int_{\mathcal{M}} \mathrm{Tr}\left(\omega\wedge d\omega + \frac{2}{3}\omega\wedge\omega\wedge\omega\right). \tag{3.77}$$

By the Atiyah–Singer index theorem again, we can write

$$\Xi = \frac{1}{2}\frac{\eta[0]}{\dim(G)} + \frac{1}{12}\frac{I[g]}{2\pi}, \tag{3.78}$$

which is a topological invariant. Finally, this contribution makes the path integral well-defined and a topological invariant of the spacetime manifold $\mathcal{M}$,

$$\mathcal{Z} = e^{i\pi \dim(G)\Xi} \sum_{\alpha} \exp\left[i\left(k + \frac{c_2(G)}{2}\right) I[A^{\alpha}]\right] T_{\alpha}. \tag{3.79}$$

However, the addition of this counterterm comes at a cost. $I[g]$ is not entirely well defined, analogously to how S_{CS} is not entirely well defined, as gauge transformations can add $2\pi kn$ ($n \in \mathbb{N}$) to its value. Similarly, "gauge transformations" in ω, known as framing (again, this is a different framing to the one discussed when regularizing $\Phi_s[\gamma]$), can change the value of $I[g] \to I[g] + 2\pi s$ ($s \in \mathbb{N}$). This can be rephrased in the language of bundles. Every oriented 3-manifold can have its tangent bundle trivialized. Such trivialization is what we call framing. As there is no unambiguous value for $I[g]$, one must describe what happens to the path integral under a change of framing,

$$\mathcal{Z} \to \exp\left(\frac{2\pi i s \dim(G)}{24}\right) \mathcal{Z}. \tag{3.80}$$

As a side note, if the gauge group happens to be $G = SU(5)$, then $\dim(G) = 24$, and the phase obtained under a change of framing becomes unity. Therefore, there is no framing anomaly in $SU(5)$ Chern-Simons theory.

3.7 Hilbert Space Structure

In this section, we give insight into the structure of the Hilbert space obtained from the non-Abelian Chern-Simons action, without reference to the path integral, a theory developed by Witten and Atiyah. This section will mainly take the gauge group to be $SU(N)$, unless otherwise stated, and will work on a 3-manifold $\mathcal{M} = \mathbb{R} \times \Sigma$. The main source used for this section is [7].

3.7.1 The Phase Space and Hilbert Space Dimension

Let us start with the non-Abelian Chern-Simons action in the Coulomb gauge,

$$S_{CS} = \frac{k}{4\pi} \int_{\mathcal{M}} d^3x \epsilon^{ij} \operatorname{Tr}\Big(A_i \dot{A}_j\Big). \tag{3.81}$$

The corresponding equal time commutation relation obeyed by the gauge field is

$$[A_i^a(\vec{x}), A_j^b(\vec{y})] = \frac{2\pi i}{k} \epsilon_{ij} \delta^{ab} \delta^{(2)}(\vec{x} - \vec{y}), \tag{3.82}$$

constrained to the requirement of A begin a flat connection ($F = dA + A \wedge A = 0$). The resulting phase space after implementing this constraint, M, is called the "moduli space of flat connections", which looks like a compact manifold with some singularities. As we will see, M will inherit a symplectic structure from the unconstrained phase space.

It turns out, that finding the flat connections on Σ is something we have already done for the special case $\Sigma = T^2$ in section 3.5. The solutions were parameterized by the class of non-trivial Wilson loops. Here the same thing applies. For the gauge group $G = SU(N)$, there are N^2-1 generators. In other words, there are N^2-1 holonomies for each non-contractible cycle. The added complication here is that one must identify flat connections that differ by a gauge transformation. When this is done, one arrives at the conclusion that the moduli space M had dimension $(2g-2)(N^2-1)$, where g is the genus of Σ. Note that as the Chern-Simons action is first order in derivatives, the A_i describe both the positions and momenta of the system, meaning that the phase space $\mathcal{P}$ is just the moduli space M. Roughly speaking, by Heisenberg's uncertainty principle, we expect a quantum degree of freedom per unit volume in phase space. As M is compact, it has a finite symplectic area, so we expect the Hilbert space on $\mathcal{M}$ to be finite-dimensional.

The question one might ask now is, what is the dimension of the Hilbert space $\mathcal{H}$ when quantizing the $SU(N)$ Chern-Simons action on a space manifold Σ? Let's start with a simple example. When $\Sigma = S^2$, there are no non-contractible loops (or flat connections), so there is one single state, $\dim(\mathcal{H}) = 1$. It can be shown that when $G = SU(2)$, the dimension of the Hilbert space corresponding to a space manifold with $g > 1$ is [23],

$$\dim(\mathcal{H}) = \left(\frac{k+2}{2}\right)^{g-1} \sum_{j=0}^{k} \left[\sin\left(\frac{(j+1)\pi}{k+2}\right)\right]^{2(g-1)}. \tag{3.83}$$

For the case $\Sigma = T^2$ and $g = 1$, $\dim(\mathcal{H}) = k+1$. Finally, it is possible to compute the dimension of the Hilbert space by evaluating the path integral in the absence of sources and under a cyclic time manifold S^1 [24],

$$\mathcal{Z} = \int \mathcal{D}A \exp\left(\frac{ik}{4\pi} \int_{S^1\times\Sigma} d^3x S[A]\right) = \dim(\mathcal{H}). \tag{3.84}$$

3.7.2 Adding Particles

We are finally at the point where we can add sources to the theory, which couple to the vector potential as usual, $\mathrm{Tr}(A_\mu J^\mu)$. The difference now is that the current must transform under the Gauge group, as usual in non-Abelian Yang-Mills theory. Unfortunately for us, this means we can't just pick a static, non-dynamic background current, it carries degrees of freedom. We will call this degree of freedom color, in analogy to QCD. As usual, if we have a field that transforms under the fundamental representation of $SU(N)$, we can write the field as an N component vector, w_γ, $\gamma \in \{1, ..., N\}$. As the gauge group is $SU(N)$, the quantify $w^\dagger w = ww^\dagger$ is invariant. Further, we will define two field configurations to be equivalent if $w' = e^{i\theta} w$. Without this restriction, w_γ would span the space S^{2N-1},

$$w^\dagger w = \sum_{\gamma=1}^{N} |w_\gamma|^2 = \sum_{\gamma=1}^{N} (w_\gamma^{\text{Real}})^2 + (w_\gamma^{\text{Imag}})^2 = \kappa, \tag{3.85}$$

which is the equation of the surface of a $2N$ dimensional sphere. After associating field configurations that differ by a complex phase, the resulting manifold becomes $S^{2N-1}/U(1) = \mathbb{CP}^{N-1}$. Consider the following action for w,

$$S_w = \int dt(iw^\dagger D_t w - \kappa\alpha), \tag{3.86}$$

where $D_t = \partial_t - i\alpha$ is the covariant derivative, and α is a new gauge field. This action has a $U(1)$ symmetry which associates w and $e^{i\theta(t)}w$. Further, the equation of motion of α is $w^\dagger w = ww^\dagger = \kappa$, so the length constraint is satisfied. Therefore, the field configurations allowed by the equations of motion span $\mathbb{CP}^{N-1}$. We can now couple this color degree of freedom to the Chern-Simons field,

$$S_w = \int dt \left(iw^\dagger D_t w - \kappa\alpha - w^\dagger A_0(t) w \right), \tag{3.87}$$

where $A_0(t) = A_0(t, \vec{x} = \vec{x}_0)$ represents a stationary particle at $\vec{x}_0$. The corresponding equation of motion is

$$i\frac{dw}{dt} = A_0(t)w, \tag{3.88}$$

meaning that w precesses in a way dictated by the Chern-Simons field. It is now simple to quantize this system, starting with the color. The commutation relations read

$$[w_\gamma, w_{\gamma'}^\dagger] = \delta_{\gamma\gamma'}. \tag{3.89}$$

Following the usual canonical quantization procedure, define a vacuum state $|0\rangle$ as the one which satisfies $w_\gamma |0\rangle = 0 \; \forall\gamma$. From this, we can construct states in the Fock space as

$$|\gamma_1, ..., \gamma_n\rangle = w_{\gamma_1}^\dagger ... w_{\gamma_n}^\dagger |0\rangle . \tag{3.90}$$

Recall that classically we want the length $w^\dagger w = ww^\dagger = \kappa$. However, at the quantum level, $w^\dagger w \neq ww^\dagger$. A solution is to keep

$$Q = \frac{1}{2}(w^\dagger w + ww^\dagger) = w^\dagger w + \frac{N}{2} \tag{3.91}$$

fixed and equal to κ. Note that this symmetric choice shifts the spectrum by $N/2$, meaning that when N is even, Q is an integer, and when N is odd, Q is a half-integer.

For convenience, it is better to define

$$\kappa_{\text{eff}} = \kappa - \frac{N}{2}, \tag{3.92}$$

and the quantization condition becomes $\kappa_{\text{eff}} = w^\dagger w = \hat{n}_w \in \mathbb{N}$, where $\hat{n}_w$ is the number operator of the color degree of freedom. Hence, if $\kappa_{\text{eff}} = 0$, there are no particles and the only allowed state is $|0\rangle$. If $\kappa_{\text{eff}} = 1$, there is a single particle. As there are N types of color particles, the Hilbert space contains N states, $\{w_\gamma |0\rangle\}$. If $\kappa_{\text{eff}} = 2$, then there are two particles. However, not all pairs of w_γ, $w_{\gamma'}$ give rise to different states, by eq. 3.89, if $\gamma \neq \gamma'$ then $w_\gamma w_{\gamma'} |0\rangle = w_{\gamma'} w_\gamma |0\rangle$. Therefore, there are in total $\frac{1}{2}N(N+1)$ distinct states, which transform under the symmetric representation of $SU(N)$. This process generalizes, by considering higher κ_{eff} we can build all symmetric representations of $SU(N)$. A very similar analysis shows that if we quantize w_γ with anti-commutators instead of commutators, the resulting particles would transform under the anti-symmetric representations of $SU(N)$.

3.7.3 Wilson Lines

Let us calculate the probability amplitude of a particle starting and ending in the w_γ state. As we are dealing with a single particle, we are considering the $\kappa_{\text{eff}} = 1$ theory, and particles transform under the $\mathbf{N}$ representation of $SU(N)$.

$${}_{in}\langle w_\gamma | w_\gamma \rangle_{out} = \int \mathcal{D}\alpha \mathcal{D}w \mathcal{D}w^\dagger e^{iS_w[w,\, \alpha;\, A_0]} w_\gamma(t=\infty) w_\gamma^\dagger(t=-\infty) = W[A_0], \tag{3.93}$$

which can be evaluated to,

$$W[A_0] = \text{Tr}\left[\mathcal{P}\exp\left(i\int dt A_0\right)\right], \tag{3.94}$$

where the trace is evaluated in the defining representation, and again $\mathcal{P}$ is the path ordering operator. This object is known as a Wilson line. If one considers any time slice Σ of the manifold, it could be the case that a Wilson line passes through it. Now, we can see what happens to the Chern-Simons theory if we insert n Wilson lines. Let the representation of each line be R_i, $i \in \{1, ..., n\}$, and be located at $\vec{x}_i$. As an example, let $\mathcal{M} = \mathbb{R} \times S^2$. We already saw that without Wilson lines there is a single state, as there are no non-contractible loops. However, now we will get a different Hilbert space, $\mathcal{H}_{i_1 \ldots i_n}$. Recall the constraint without sources is $j_{12} = 0$. In the presence of n Wilson lines $w_\gamma^{(i)}$, this condition reads

$$\frac{k}{2\pi} f_{12}^a(\vec{x}) = \sum_{i=1}^{n} \delta^{(2)}(\vec{x} - \vec{x}_i) w^{(i)\dagger} T^a w^{(i)}, \tag{3.95}$$

where T^a are the generators of $SU(N)$, and the right-hand side takes the role of a current with internal degrees of freedom w^i which transforms under the gauge group. One can see from eq. 3.95 that the Dirac quantization condition applies to each f^a_{12} individually, which we expect to be true for a compact gauge group.

3.7.4 Dimension of the Hilbert Space

As before, it is easier to consider the weakly coupled (large k) limit. Firstly, we seem to have a problem, as the integral over $\frac{1}{4\pi} f^a_{12} \propto \frac{1}{k} \in \mathbb{Z}$, which for very large k, non-trivial charges looks impossible. Therefore, we require the charges on the right-hand side of eq. 3.95 to cancel each other out. This is equivalent to color charge in QCD, where by confinement, we require the total color charge to vanish. This is only possible if baryons transform under the trivial representation of the QCD group $SU(3)$. In an analogous way, we can take tensor products of the representations of each Wilson loop, R_i, and this must have a decomposition that includes the trivial representation.

$$\bigotimes_{i=1}^{n} R_i = \mathbf{1}^p \oplus ..., \tag{3.96}$$

where $p \in \mathbb{N}$ is the number of times $\mathbf{1}$ appears in the decomposition. Each $\mathbf{1}$ gives a different allowed state in the Hilbert space. Therefore, we conclude that

$$\lim_{k\to\infty} \dim(\mathcal{H}_{i_1...i_n}) = p. \tag{3.97}$$

However, this is an upper bound for large k. For two Wilson lines, we require $R_1 = \overline{R}_2$ so that the tensor product of representations has the trivial representation in its decomposition, otherwise the Hilbert space is empty. Consider now three Wilson loops, with representations R_i, R_j, R_k. We can write the tensor products as

$$R_i \otimes R_j = \bigoplus_m N^m_{ij} R_m, \tag{3.98}$$

$$R_i \otimes R_j \otimes R_k = \bigoplus_{m_1 m_2} N^{m_1}_{ij} N^{m_2}_{m_1 k} R_{m_2}, \tag{3.99}$$

where N^k_{ij} is the number of times R_k appears in the decomposition of $R_i \otimes R_j$. We are interested in the number of times the trivial representation, call it R_1, appears in this expansion,

$$\dim(\mathcal{H}_{ijk}) = \sum_{m_1} N^{m_1}_{ij} N^1_{m_1 k}. \tag{3.100}$$

Using the fact that $N^1_{m_1 k}$ is non-zero only if $m_1 = \overline{k}$,

$$\dim(\mathcal{H}_{ijk}) = N^{\overline{k}}_{ij}. \tag{3.101}$$

Now consider 4 Wilson loops. We can calculate the tensor product of the representations as

$$R_i \otimes R_j \otimes R_k \otimes R_k = \bigoplus_{m_1 \ldots m_3} N_{ij}^{m_1} N_{m_1 k}^{m_2} N_{m_2 l}^{m_3} R_{m_3}, \tag{3.102}$$

from which we conclude that the dimension of the Hilbert space is

$$\dim(\mathcal{H}_{ijkl}) = \sum_{m_1} N_{ij}^{m_1} N_{m_1 k}^{\bar{l}}. \tag{3.103}$$

It is easy to see that the general case is [25]

$$\dim(\mathcal{H}_{i_1 \ldots i_n}) = \sum_{m_1 \ldots m_{n-3}} N_{i_1 i_2}^{m_1} N_{m_1 i_3}^{m_2} \ldots N_{m_{n-3} i_{n-1}}^{\bar{i}_n}, \tag{3.104}$$

where we note that this is the same expression as the number of fusion channels of n anyons to the vacuum (eq. 2.12)! In fact, the relationship between the dimension of the Hilbert space and the fusion rules is not a coincidence. To see the full derivation, refer to [26].

3.7.5 Fusion Rules

We shall start with a simpler example, $SU(2)_k$ ($SU(2)$ gauge group and level k), whose representations are labeled by its spin s. s takes half-integer values and corresponds to the representation $\mathbf{d}$, where $d = 2s + 1$. The tensor product of two representations is given by

$$r \otimes s = |r - s| \oplus |r - s| + 1 \oplus \ldots \oplus r + s - 1 \oplus r + s. \tag{3.105}$$

The fusion rules for anyon types r and s look very similar, but with some subtleties. Firstly, we require the spin of the representation, j, to be $j \leq k/2$. The fusion rules are obtained from throwing away any representations that do not obey this constraint and some further restrictions. The correct version is

$$r \times s = |r - s| + \ldots + \min(k - r - s, r + s). \tag{3.106}$$

As an example, the fusion rules for $k = 2$ read

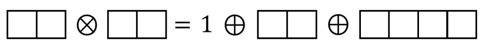
$$\mathbf{2} \times \mathbf{2} = \mathbf{1} + \mathbf{3}, \qquad \mathbf{2} \times \mathbf{3} = \mathbf{2}, \qquad \mathbf{3} \times \mathbf{3} = \mathbf{1}, \tag{3.107}$$

with all fusions of the form $\mathbf{d} \times \mathbf{1} = \mathbf{d}$. These are the same fusion rules as the Ising anyons we talked about in section 2.3.2, where $\mathbf{2} = \sigma$ and $\mathbf{3} = \psi$. It turns out that this is no coincidence, and the Ising anyons are related to the $SU(2)_2$ Chern-Simons theory, and play a role in the $\nu = 5/2$ FQHE state, which we will briefly discuss in section 4.4. In addition, the other example of non-Abelian anyons we saw, the Fibonacci anyons, are related to the $SU(2)_3$ Chern-Simons theory (there are more particles than the Fibonacci anyons present), and are thought to occur in the $\nu = 12/5$ FQHE state.

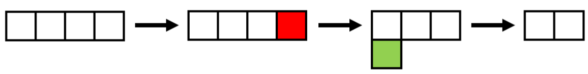
We are now going to explain the general fusion rules for $SU(N)_k$. This is mainly for completeness and we are just going to explain the way they are calculated, but we will not explain where they come from. We choose to do this since the process is quite convoluted and not particularly illuminating. We begin by stating that in the same way we were restricted to $j \leq k/2$ in $SU(2)_k$, here we are restricted to $l_1 \leq k$, where l_1 is the number of boxes in the first row of the Young tableau of the representation. In the case of $SU(2)_k$, $l_1 = 2s$, so our generalization is consistent with what we know so far. Then, define

$$t = l_1 - k - 1. \tag{3.108}$$

If $t < 0$, keep the diagram, and if $t = 0$, reject the diagram (as then $l_1 > k$). When $t > 0$ things get more interesting. Starting from the rightmost side of the first row, and moving downwards and left, remove t blocks from the Young tableau. Then, starting at the bottom of the first column and moving up and right, add t boxes to the Young tableau. Then, include a sign $(-1)^{r_- + r_+ + 1}$ to the diagram, where r_- is the number of columns that had boxed removed, and r_+ is the number of columns that had boxes added. Then, repeat the process until $t \leq 0$. Let us do an example, the third fusion rule for $SU(2)_2$ discussed above. The representations follow the decomposition



The first two terms on the right-hand side have $t < 0$, so we keep them. The third term is more interesting as it has $t = 1$. So, we remove 1 block from the right of the first (and only) row, shown in red, and we add a block on the bottom of the first column, shown in green. As we are dealing with $SU(2)$, a column of 2 blocks can be removed, giving us the



3 representation again. However, we have not yet talked about the sign. One column (the fourth one) got a box removed, so $r_- = 1$. Similarly, one row got a box added, so $r_+ = 1$. Therefore, this diagram is associated with a negative sign which cancels the second term in the decomposition. This leaves us the expected result for the fusion, $\mathbf{3} \times \mathbf{3} = \mathbf{1}$.

3.8 Asides

Chern-Simons theory happens to be closely related to several other fields of physics. We will greatly discuss its applications in condensed matter in the following chapters but also want to give some intuition about other places in which it comes up. These examples are adapted from [13].

3.8.1 θ Term in Yang-Mills Theory

The 3D Chern-Simons theory is related to the 4D Yang-Mills [27] theory in a manifold X with a boundary $\partial X = \mathcal{M}$, in which the Chern-Simons theory resides. Using Stokes' theorem,

$$S_{CS}^X[A] = \frac{k}{4\pi} \int_{\mathcal{M}=\partial X} A \wedge dA = \frac{k}{4\pi} \int_X d(A \wedge dA), \tag{3.109}$$

$$= \frac{k}{4\pi} \int_X dA \wedge dA = \frac{k}{4\pi} \int_X F \wedge F, \tag{3.110}$$

where we have used $d^2 = 0$. The θ term in Yang-Mills theory is

$$S_\theta[A] = \frac{\theta}{16\pi^2} \int_X F \wedge F, \tag{3.111}$$

which is the same as the bulk action equivalent to a boundary Chern-Simons term.

3.8.2 4D Gauge Theories and Spinors

The 4-dimensional action S_{CS}^X depends explicitly on F only, which is a gauge invariant quantity. However, S_{CS}^X is a topological invariant of the spacetime manifold X. For the theory to hold, for any two 4-manifolds X and X' that share the boundary, $\partial X = \partial X' = \mathcal{M}$, we must have the same path integral. This can be achieved if $S_{CS}^X - S_{CS}^{X'} \in 2\pi\mathbb{Z}$. It turns out, that this condition is equivalent to the following. If you perform surgery on X and X', by reversing the direction of one of the manifolds, and joining them along their common boundary, you obtain the manifold $Y = (X \cup X')/\mathcal{M}$. If Y is a "spin manifold" (a manifold that allows the existence of spinors), then the path integral for X and X' will be the same and the theory is consistent. In conclusion, 3D Chern-Simons theory is deeply tied with 4D gauge theories involving fermions.

3.8.3 3D Quantum Gravity

Recall that the Chern-Simons action can be expressed in a coordinate-free manner using 1-forms, $\mathcal{L} = \frac{k}{4\pi} A \wedge dA$. This is explicitly metric-independent, which forces the stress-energy tensor to vanish,

$$T_{CS}^{\mu\nu} = \frac{2}{\sqrt{|g|}} \frac{\delta \mathcal{L}}{\delta g_{\mu\nu}} = 0. \tag{3.112}$$

This has several consequences. Firstly, the Hamiltonian (the 0, 0 component) vanishes. Secondly, any diffeomorphism or transformation of coordinates leaves the action invariant. These two properties are common to that of gravity, in particular, 3D quantum gravity to first order is simply a Chern-Simons term [28].

Chapter 4

The Fractional Quantum Hall Effect

Perhaps one of the most active topological phases being researched is the Fractional Quantum Hall Effect (FQHE). In this section, we discuss effective field theoretic models of quantum Hall states. We will start with the classical Hall effect, then we will explain the integer quantum Hall effect from the filling of Landau levels, and we will end our discussion with a field-theoretic treatment. It will turn out that the Chern-Simons action accurately describes the integer quantum Hall effect. Lastly, we will add an additional quasi-hole field to our theory, which will explain the $1/m$ filling factor Abelian FQHE states. Adding a hierarchy of fields analogous to that which explains the $1/m$ states, we will explain any Abelian fractional state from a field-theoretic perspective. Note that this is not how the FQHE is usually introduced. The usual way is to argue from the filling of Landau levels what properties the ground state wavefunction should have, and then guess the answer [29]. Laughlin was famously good at this, he explained the $1/m$ states (known as the Laughlin states) in this fashion.

4.1 The Classical Effect

This section is mainly adapted from [30]. Suppose we have a 2+1 dimensional system, such as the surface of a metal, with a large magnetic flux through it. Then, if an electric field is applied and a current is induced, the charges will be deflected by the magnetic field and cause a current density in the perpendicular direction to the applied E field and external B field. To be more precise, consider the following setup (figure 4.1), The classical equation of motion for electrons is given by the Lorentz force

$$m\ddot{x} = -e\dot{y}B, \tag{4.1}$$

$$m\ddot{y} = -e(E_y - \dot{x}B). \tag{4.2}$$

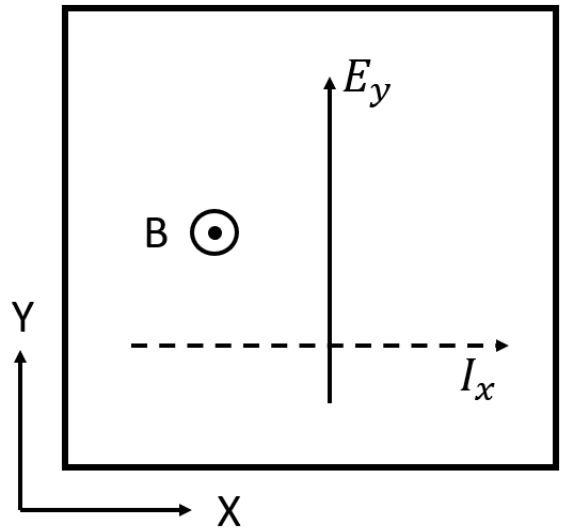


Figure 4.1: Setup for the classical quantum Hall effect. There is a B field perpendicular to the plane, and an E field is applied in the y direction. Consequently, this will induce a current in the x direction.

We can conveniently express both equations in terms of the complex representation of position $z = x + iy$,

$$m\ddot{z} - iBe\dot{z} = -ieE_y. \tag{4.3}$$

The solution to this differential equation is

$$z(t) = z_0 e^{i\omega_c t} + v_d t + z_1, \tag{4.4}$$

where $\omega_c = eB/m$ is the cyclotron frequency, $v_d = E_y/B$ is the drift velocity, and z_0, z_1 are integration constants. Note that the drift velocity term is real, meaning the electron's velocity is in the x direction, perpendicular to both the E and B fields. We can calculate the current density as

$$J_x = n(-e)v_d, \tag{4.5}$$

where n is the charge density. Finally, using Ohm's law,

$$\begin{pmatrix} J_x \\ J_y \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \tag{4.6}$$

$$-\sigma_{xy} = -\frac{J_x}{E_y} = \frac{ne}{B} = \sigma_H \tag{4.7}$$

where σ_H is the Hall conductivity.

4.2 The Integer Effect

It turns out that the analysis above is not the full story. The thermal de-Broglie wavelength is proportional to $T^{-1/2}$ (T being the temperature), so as the temperature drops, quantum effects become more important. In addition, under large magnetic fields, the energy levels of electrons become quantized into what are known as Landau levels. At low temperatures and large B fields, the Hall resistivity ($\rho_H = 1/\sigma_H$), is not proportional to the B field. Instead, it is quantized, see figure 4.2. The Hall resistivity has plateaus

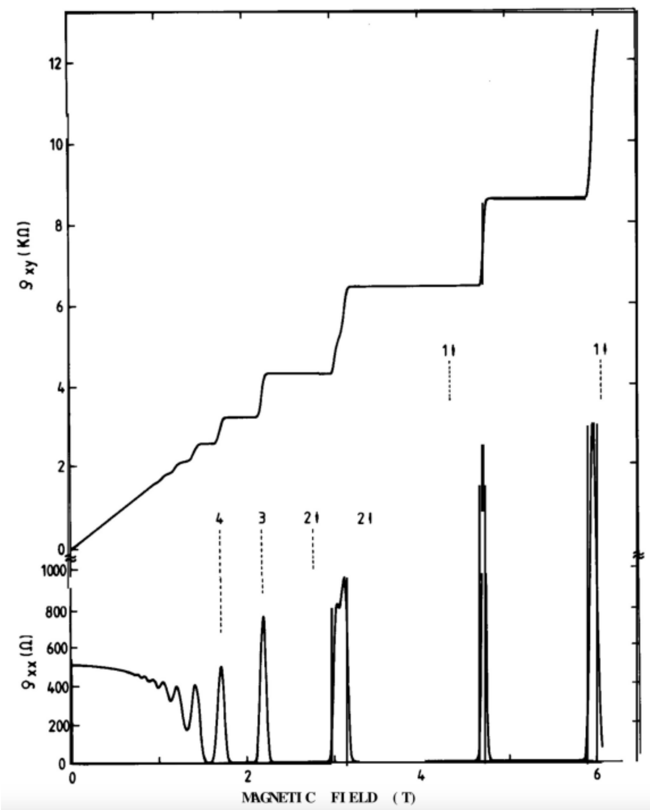


Figure 4.2: Relationship between the Hall resistivity and the applied B field at low temperature and high field. Instead of a linear relationship, the resistivity is quantized. Figure obtained from [31].

at values of $\frac{1}{\nu}\frac{h}{e^2}$, where $\nu \in \mathbb{N}$. As we will see, the integer ν happens to be the number of filled Landau levels. The magnetic fields at which the plateaus occur is when the "filling factor" $\nu = n\phi_0/B \in \mathbb{N}$, and $\phi_0 = h/e$ is the flux quantum. We can explain this effect as the filling of Landau levels. Consider the Hamiltonian

$$H = \frac{(-i\vec{\nabla} + e\vec{A})^2}{2m}, \tag{4.8}$$

which is obtained by considering a free electron Hamiltonian and using the conjugate momentum under a vector potential A, $p \to p + eA$. The easiest way to proceed is by picking the Landau gauge, $\vec{A} = Bx\hat{y}$ (representing a B field in the z direction), which makes the Hamiltonian

$$H = \frac{1}{2m}\left(-\partial_x^2 + (-i\partial_y + eBx)^2\right). \tag{4.9}$$

As the Hamiltonian has bulk translation invariance along the y direction, we may factorize the eigenstates as

$$\psi(x, y) = e^{ik_a y}\phi_{k_a}(x), \tag{4.10}$$

where $k_a = 2\pi a/L_y$ and L_y is the length of the material along the y axis. Plugging this ansatz into the time-independent Schrödinger equation,

$$(-i\partial_y + eBx)e^{ik_a y} = e^{ik_a y}(k_a + eBx), \tag{4.11}$$

$$(-i\partial_y + eBx)^2 e^{ik_a y} = (-i\partial_y + eBx)e^{ik_a y}(k_a + eBx) = e^{ik_a y}(k_a + eBx)^2, \tag{4.12}$$

$$He^{ik_a y}\phi_{k_a}(x) = e^{ik_a y}\frac{1}{2m}(\partial_x^2 + (k_a + eBx)^2)\phi_{k_a}(x) = Ee^{ik_a y}\phi_{k_a}(x), \tag{4.13}$$

$$E\phi_{k_a}(x) = \left(-\frac{\partial_x^2}{2m} + \frac{1}{2}m\omega_c^2(k_a l^2 + x)^2\right)\phi_{k_a}(x), \tag{4.14}$$

where $l^2 = \frac{\hbar}{eB}$. We note that the Schrödinger equation for ϕ_{k_a} is a simple harmonic oscillator with frequency ω_c, shifted to $x_0 = -k_a l^2$. Hence, the energy levels are

$$E_p = \hbar\omega_c\left(p + \frac{1}{2}\right). \tag{4.15}$$

These are known as Landau levels. The corresponding ground state eigenfunctions look like Gaussians centered at x_0 and oscillating complex exponentials on y. As we will explain the integer quantum Hall effect in terms of the filling of Landau levels, we must calculate the degeneracy of each level p. The range of possible x values is L_x, hence $0 \leq -k_a l^2 \leq L_x$. Noting that k_n is quantized in steps of $2\pi/L_y$, the number of possible k_a values for each Landau level p is

$$\text{degeneracy} = \frac{L_x L_y}{2\pi l^2} = \frac{\text{Area } B}{\phi_0} = \frac{\phi}{\phi_0}, \tag{4.16}$$

where ϕ is the total flux through the surface. Define the filling factor ν as the number of filled Landau levels,

$$\nu = \frac{N_e}{\text{degeneracy}} = \frac{N_e}{\phi/\phi_0} = \frac{N_e}{N_\phi} = \frac{nB}{\phi_0}, \tag{4.17}$$

where N_e is the number of electrons, n is the number density and N_ϕ is the number of flux quanta. We can now express the Hall resistivity in terms of the filling factor as

$$\rho_H = \frac{B}{ne} = \frac{1}{\nu}\frac{h}{e^2}. \tag{4.18}$$

Therefore, if ν can only take integer values, then we have found the right quantization condition. Note that if the chemical potential μ is anything except the exact energy of one of the Landau levels, $\hbar\omega_c(p+1/2)$, then exactly an integer number of Landau levels are filled. However, in experiments what can be controlled is the magnetic field, not the chemical potential. If the B field is anything except a particular discrete set of values $n\phi_0/\mathbb{Z}$, then the chemical potential will sit excatly on a Landau level. Laughlin showed that the Hall conductance must nevertheless be quantized [32]. Laughlin's argument uses the Byers-Yang theorem, which states that inserting a flux quantum ϕ_0 through a hole in the system leaves the wavefunction invariant [33]. The remainder of the subsection is based on [3]. Suppose we set up a quantum Hall state on a disk with a hole in the middle, such that the magnetic field is perpendicular to the disk and the electric field goes radially around the annulus (see fig 4.3). Consider adiabatically inserting a flux quantum

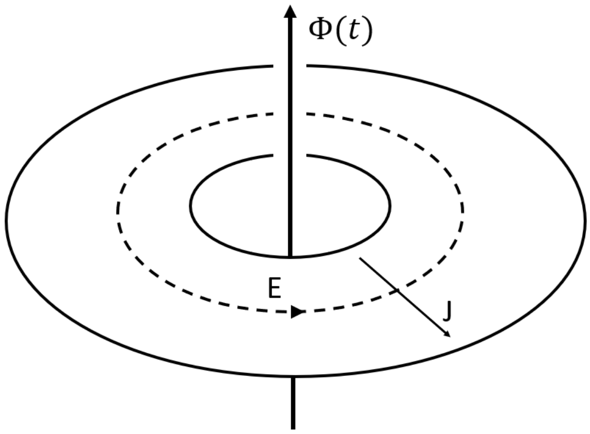


Figure 4.3: Setup for Laughlin's argument. A quantum Hall state on the surface of a disk with an annulus through which a time-varying flux is inserted. By the Hall effect, a radial current is induced.

through the annulus. By Faraday's law, an emf $\mathcal{E} = -\frac{d\Phi}{dt}$ is generated. By Ohm's law, $I = \sigma_H \mathcal{E}$ so

$$\Delta Q = \int dt I(t) = \sigma_H \int dt \mathcal{E}(t) = -\sigma_H \int dt \frac{d\Phi}{dt} = -\sigma_H \phi_0. \tag{4.19}$$

By the Byers-Yang theorem, after inserting the flux quantum we must return to the same state in the bulk. The only thing that could have changed is that an integer number of electrons r might have moved from the interior of the disk to the exterior, in which case $\Delta Q = -re$, from which we deduce that $\sigma_H = re^2/h$.

As an aside, consider the current in a system described by the Chern-Simons action,

$$S_{CS}[A] = \frac{k}{4\pi} \int d^3x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho, \tag{4.20}$$

$$J_i = \frac{\delta S_{CS}}{\delta A_i} = -\frac{k}{2\pi} \epsilon^{ij} E_j, \tag{4.21}$$

which directly implies by Ohm's law a Hall conductivity of $\sigma_H = \frac{k}{2\pi}$, whereas we previously had $\sigma_H = \frac{\nu}{2\pi}$ (now we are using units in which $e = \hbar = 1$). In addition, as we saw in section 3.2, the Chern-Simons level k must be an integer. Therefore, the Chern-Simons action describes a system with a quantized Hall conductance, which must be a multiple of $1/2\pi$. This link between Chern-Simons and the integer quantum Hall effect is quite suspicious, and indeed we can explain the integer quantum Hall effect as having a leading Chern-Simons effective action. This is not surprising as it is the lowest-order term in powers of A which can be written down and is gauge invariant, which has the greatest contributions for low-energy, long-wavelength dynamics. What is even better is that this analysis was remarkably simple. We did not have to consider Landau levels or magnetic fields or solve the Schrödinger equation. In the next section, we will attempt to describe the fractional quantum Hall effect from a field-theoretic perspective, in which Chern-Simons theory will play an important role.

4.3 The Abelian Fractional Effect

While we have shown that the Hall conductivity must be quantized, with each plateau being associated with the fact that the filling factor ν takes integer values, there also exists quantized Hall states associated with fractional values of the filling factor. In the last section, we saw how the integer quantum Hall effect could be explained if the leading term of the theory was a Chern-Simons term. However, this seems to be a problem, as it looks like it contradicts the existence of Hall states at fractional filling factors. To solve this issue, we have to violate some assumptions taken in the derivation that ν can only take integer values. The hidden assumption we took, was that there were no dynamical degrees of freedom that may affect the low-energy physics, which should be true for any system with an energy gap to excitations. It turns out that this is not entirely true, there can be degrees of freedom that are gapped and do affect the low-energy physics. These are topological degrees of freedom, which dictate the dynamics of the FQHE. In this section, we describe the topological degrees of freedom needed to explain the FQHE. We will first consider the Abelian Chern-Simons action, which will explain the FQHE states with filling factor $\nu = 1/m$, $m \in \mathbb{Z}$. Then, we will generalize to other filling factors.

4.3.1 $1/m$ Filling Factors

As explained in the introduction to this subsection, we will first consider an Abelian $U(1)$ gauge field, a_μ, which will explain the $1/m$ Laughlin states. However, note that a_μ is not the vector potential of electromagnetism. Instead, it is an emergent field, which comes about from the motion of the underlying electrons. It is completely analogous to how phonons arise from the collective behavior of atoms in lattice sites. While there is a relationship between a_μ and electronic degrees of freedom, this will become apparent later. This section is based on [7].

As seen in section 3.1, the Chern-Simons action by itself gives rise to trivial dynamics. A solution is to add a Maxwell term,

$$S[a] = \frac{1}{4g^2}\int d^3x f_{\mu\nu}f^{\mu\nu} + \frac{k}{4\pi}\int d^3x \epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho, \tag{4.22}$$

where $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$, and g is a coupling constant. The corresponding equations of motion are

$$\partial_\mu f^{\mu\nu} + \frac{kg^2}{4\pi}\epsilon^{\nu\rho\sigma}f_{\rho\sigma} = 0. \tag{4.23}$$

Firstly, we want to demonstrate that there is an energy gap to excitations (aka, the particles described are massive). Define

$$\widetilde{f}^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho}f_{\nu\rho}, \tag{4.24}$$

$$f_{\alpha\beta} = \epsilon_{\mu\alpha\beta}\widetilde{f}^\mu, \tag{4.25}$$

where the second equation is obtained by multiplying by $\epsilon_{\mu\alpha\beta}$ on both sides and using $\epsilon_{\mu\alpha\beta}\epsilon^{\mu\nu\rho} = (\delta^\nu_\alpha\delta^\rho_\beta - \delta^\rho_\alpha\delta^\nu_\beta)$. This lets us rewrite the equations of motion as

$$\partial_\mu\epsilon^{\sigma\mu\nu}\widetilde{f}_\sigma + \frac{kg^2}{2\pi}\widetilde{f}^\nu = 0, \tag{4.26}$$

$$\left(\partial_\mu\epsilon^{\sigma\mu\nu} + \frac{kg^2}{2\pi}\eta^{\nu\sigma}\right)\widetilde{f}_\sigma = 0, \tag{4.27}$$

$$\left(\partial^{\widetilde{\mu}}\epsilon_{\widetilde{\sigma}\widetilde{\mu}\nu} + \frac{kg^2}{2\pi}\eta_{\nu\widetilde{\sigma}}\right)\left(\partial_\mu\epsilon^{\sigma\mu\nu} + \frac{kg^2}{2\pi}\eta^{\nu\sigma}\right)\widetilde{f}_\sigma = 0, \tag{4.28}$$

$$\left(\partial^{\widetilde{\mu}}\partial_\mu(\delta^\sigma_{\widetilde{\sigma}}\delta^\mu_{\widetilde{\mu}} - \delta^\mu_{\widetilde{\sigma}}\delta^\sigma_{\widetilde{\mu}}) + \left(\frac{kg^2}{2\pi}\right)^2\delta^\sigma_{\widetilde{\sigma}}\right)\widetilde{f}_\sigma = 0, \tag{4.29}$$

$$\left(\partial^\mu\partial_\mu + \left(\frac{kg^2}{2\pi}\right)^2\right)\widetilde{f}_{\widetilde{\sigma}} = 0, \tag{4.30}$$

where we have used $\partial^\mu\widetilde{f}_\mu = 0$. We recognize the final result as a Klein-Gordon equation for a particle of mass $kg^2/2\pi$. This result makes sense, as in the limit as $g \to \infty$, the mass becomes unbounded and there are no propagating degrees of freedom, exactly what we would expect if the Maxwell term vanished.

For the remainder of this subsection, we will attempt to write down the path integral,

$$\mathcal{Z}[A_\mu] = \int \mathcal{D}a e^{iS_{\text{eff}}[a,A]}, \tag{4.31}$$

for the $\nu = 1/m$ states and a_μ is a dynamical $U(1)$ gauge field. Here, by A_μ, we mean the background electromagnetic field on top of the one required to create the Hall state in the first place. In other words, we have to find the effective action S_{eff}. We know that the electromagnetic field A_μ couples to the electron current J^μ. Hence, once we have a relationship between the electron current J^μ and the emergent field a_μ, we will have the coupling between the two fields. As we need the current to be conserved, pretty much the only term we can write down that is also gauge invariant is

$$J^\mu = \frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu a_\rho. \tag{4.32}$$

The normalization is chosen so that the Dirac quantization condition is satisfied, as already seen in section 3.2,

$$\int_{S^2} d^2x J^0 = \frac{1}{2\pi}\int_{S^2} d^2x f_{12} \in \mathbb{Z}, \tag{4.33}$$

or in other words, the smallest allowed non-zero charge is 1. We therefore guess the effective action to be

$$S_{\text{eff}} = \int d^3x A_\mu J^\mu - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho, \tag{4.34}$$

$$= \int d^3x \frac{1}{2\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho, \tag{4.35}$$

where $m \in \mathbb{Z}$ takes the role of the Chern-Simons level. We could also add an $\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho$ term but this just adds an integer contribution to σ_H and will be ignored, as we are only interested in the fractional part. We can now calculate the Hall conductivity predicted from this action. This is done by integrating out a_μ (replacing it with its equation of motion),

$$\frac{\delta S_{\text{eff}}}{\delta a_\mu} = -\frac{m}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu a_\rho + \frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho = 0, \tag{4.36}$$

$$a_\mu = \frac{1}{m}A_\mu. \tag{4.37}$$

This gives an effective action for A_μ as[1]

$$S_{\text{eff}} = \int d^3x \frac{1}{4\pi m}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho, \tag{4.38}$$

which is the usual Chern-Simons action with a "fractional level" of $1/m$. We have already seen that the corresponding Hall conductance is $\sigma_H = 1/2\pi m$. Therefore, we have successfully described $1/m$ Hall states from a field-theoretic perspective.

We still haven't considered the dynamics of a_μ. To do this, we can couple the emergent field with its own particle current, j^μ. It will turn out that j^μ describes quasi-holes and quasi-electrons in the system. For simplicity, set the electromagnetic field to 0. The

[1]Please note [7] had a typo, an extra $1/2\pi$ as a prefactor.

equation of motion for a_μ reads

$$\frac{1}{2\pi}f_{\mu\nu} = \frac{1}{m}\epsilon_{\mu\nu\rho}j^\rho. \tag{4.39}$$

Consider the simplest case, in which we have a static unit charge at the origin, $j^0 = \delta^{(2)}(\vec{0})$, $j^1 = j^2 = 0$. The equation of motion then becomes

$$\frac{1}{2\pi}f_{12} = \frac{1}{2\pi}B = \frac{1}{m}\delta^{(2)}(\vec{0}). \tag{4.40}$$

From this, we can conclude that the particles described by the current j^μ are charged under $U(1)_{\mathrm{EM}}$. This can be seen by eq. 4.32,

$$J^0 = \frac{1}{2\pi}f_{12} = \frac{1}{m}\delta^{(2)}(\vec{0}), \tag{4.41}$$

$$q = \int d^2x J^0 = \frac{1}{m}. \tag{4.42}$$

We also see directly from eq. 4.40 that a flux of ϕ_0/m is attached to every particle of unit charge (in units where $e = \hbar = 1$, $\phi_0 = h/e = 2\pi$). Hence, as the particles described by j^μ have flux attached to them and they are charged, they are flux-charge composites and obey fractional exchange statistics. Recall the exchange angle for flux-charge composites of charge q and flux Φ is $\theta = q\Phi/2$. For the charge density $j^0 = \delta^{(2)}(\vec{0})$, $q = 1$ and $\Phi = 2\pi/m$, so the exchange angle is $\theta = \pi/m$.

4.3.2 Other Filling Factors

Before generalizing the effective field approach to other filling fractions, let us summarize what we have done so far. We started with the electron current J^μ in terms of the emergent field a_μ, eq. 4.32. Then, we coupled the EM field to its current and gave a_μ a Chern-Simons term. We can now define further emergent fields, like $\widetilde{a}_\mu$, which is related to the quasi-particle current by

$$j^\mu = \frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu\widetilde{a}_\rho, \tag{4.43}$$

and give it a Chern-Simons term in the action,

$$S_{\text{eff}}[a,\widetilde{a},A] = \int d^3x A_\mu J^\mu - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + a_\mu j^\mu - \frac{\widetilde{m}}{4\pi}\epsilon^{\mu\nu\rho}\widetilde{a}_\mu\partial_\nu\widetilde{a}_\rho, \tag{4.44}$$

$$= \int d^3x \frac{1}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + \frac{1}{2\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu\widetilde{a}_\rho - \frac{\widetilde{m}}{4\pi}\epsilon^{\mu\nu\rho}\widetilde{a}_\mu\partial_\nu\widetilde{a}_\rho. \tag{4.45}$$

We can now compute the Hall conductivity in the same way, integrating out $\widetilde{a}_\mu$,

$$S_{\text{eff}}[a,A] = \int d^3x \frac{1}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{m}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + \frac{1}{4\pi\widetilde{m}}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho, \tag{4.46}$$

$$= \int d^3x \frac{1}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu a_\rho - \frac{m - 1/\widetilde{m}}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho, \tag{4.47}$$

and then a_μ,

$$S_{\text{eff}}[A] = \int d^3x \frac{1}{4\pi(m - 1/\widetilde{m})} \epsilon^{\mu\nu\rho} A_\mu \partial_\nu a_\rho. \tag{4.48}$$

From this, we conclude that the Hall conductivity is

$$\sigma_H = \frac{\nu}{2\pi}, \tag{4.49}$$

$$\nu = \frac{1}{m - \frac{1}{\widetilde{m}}}. \tag{4.50}$$

We can also use the arguments developed for the $1/m$ states to calculate the charge and exchange angle. For a static quasi-hole at the origin that couples to a, the equations of motion are

$$m f_{12} - \widetilde{f}_{12} = 2\pi\delta^{(2)}(\vec{0}), \qquad \widetilde{m}\widetilde{f}_{12} - f_{12} = 0, \qquad \Longrightarrow f_{12} = \frac{2\pi}{m - 1/\widetilde{m}} \delta^{(2)}(\vec{0}). \tag{4.51}$$

while if the current couples to $\widetilde{a}$,

$$m f_{12} - \widetilde{f}_{12} = 0, \qquad \widetilde{m}\widetilde{f}_{12} - f_{12} = 2\pi\delta^{(2)}(\vec{0}), \qquad \Longrightarrow f_{12} = \frac{2\pi}{m\widetilde{m} - 1} \delta^{(2)}(\vec{0}). \tag{4.52}$$

Hence, the fractional charges of the emergent quasiparticles are

$$q(a) = \frac{1}{m - 1/\widetilde{m}}, \tag{4.53}$$

$$q(\widetilde{a}) = \frac{1}{m\widetilde{m} - 1}. \tag{4.54}$$

For example, in the $\nu = 5/2$ FQHE state, $m = 3$ and $\widetilde{m} = 2$, from which we obtain the charges $q(a) = 2/5$ and $q(\widetilde{a}) = 1/5$, which has been measured experimentally.

4.3.3 Generalization

This procedure of coupling new fields to the currents of previous ones recursively can be easily generalized. Consider N emergent fields a^i_μ and the action[2]

$$S_{\text{eff}}[a^i_\mu, A] = \int d^3x \frac{1}{2\pi} t_i \epsilon^{\mu\nu\rho} A_\mu \partial_\nu a^i_\rho - \frac{1}{4\pi} K_{ij} \epsilon^{\mu\nu\rho} a^i_\mu \partial_\nu a^j_\rho, \tag{4.55}$$

where K_{ij} encodes all the couplings between emergent fields and t_i specifies which combination of fields represents the electron current. For example, the analysis above with 2 emergent fields used $t_i = (1, 0)$, and

$$K_{ij} = \begin{pmatrix} m & -1 \\ -1 & \widetilde{m} \end{pmatrix}. \tag{4.56}$$

[2]Please note that [7] had a typo, they did not negate the second term, which means their following conclusions are wrong by a sign

We can perform the same analysis by integrating our emergent fields with their equations of motion, which are given by

$$\frac{\delta S_{\text{eff}}}{\delta a^i_\mu} = -\frac{1}{4\pi}K_{ij}\epsilon^{\mu\nu\rho}\partial_\nu a^j_\rho - \partial_\sigma\left(\frac{1}{2\pi}t_i\epsilon^{\nu\sigma\mu}A_\nu - \frac{1}{4\pi}K_{ji}\epsilon^{\nu\sigma\mu}a^j_\nu\right), \tag{4.57}$$

$$= \frac{1}{2\pi}t_i\epsilon^{\mu\sigma\nu}\partial_\sigma A_\nu - \frac{1}{2\pi}K_{ij}\epsilon^{\mu\nu\rho}\partial_\nu a^j_\rho = \frac{1}{2\pi}\epsilon^{\mu\nu\rho}(t_i\partial_\nu A_\rho - K_{ij}\partial_\nu a^j_\rho), \tag{4.58}$$

$$\implies t_i A_\rho = K_{ij}a^j_\rho, \tag{4.59}$$

$$a^j_\rho = (K^{-1})^{ji}t_i A_\rho. \tag{4.60}$$

This makes the effective action for the EM field

$$S_{\text{eff}}[A] = \int d^3x \frac{1}{4\pi}t_i(K^{-1})^{ij}t_j\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho. \tag{4.61}$$

As usual, the filling factor is given by the "level" of the effective Chern-Simons term for A,

$$\nu = t_i(K^{-1})^{ij}t_j, \tag{4.62}$$

$$\sigma_H = \frac{t_i(K^{-1})^{ij}t_j}{2\pi}. \tag{4.63}$$

In addition, we can work out the charge of each quasi-particle by coupling a current $j^\mu = (\delta^{(2)}(\vec{0}), 0, 0)$ to a^k_μ and setting the background field $A_\mu = 0$. The resulting equation of motion is

$$\frac{\delta S_{\text{eff}}}{\delta a^i_\mu} = -\frac{1}{2\pi}K_{ij}\epsilon^{\mu\nu\rho}\partial_\nu a^j_\rho + \delta^k_i j^\mu = 0, \tag{4.64}$$

$$\frac{1}{4\pi}K_{ij}\epsilon^{\mu\nu\rho}f^j_{\nu\rho} = \delta^k_i j^\mu, \tag{4.65}$$

$$\frac{1}{2\pi}f^j_{\alpha\beta} = \epsilon_{\alpha\beta\mu}(K^{-1})^{jk}j^\mu. \tag{4.66}$$

Using the definition of the electron current,

$$J^\mu = \frac{1}{2\pi}t_i\epsilon^{\mu\nu\rho}\partial_\nu a^i_\rho, = \frac{1}{4\pi}t_i\epsilon^{\mu\nu\rho}f^i_{\nu\rho}, \tag{4.67}$$

$$= \frac{1}{2}t_i\epsilon^{\mu\nu\rho}\epsilon_{\nu\rho\sigma}(K^{-1})^{ik}j^\sigma, \tag{4.68}$$

$$= \frac{1}{2}t_i\epsilon^{\mu\nu\rho}\epsilon_{\nu\rho 0}(K^{-1})^{ik}\delta^{(2)}(\vec{0}), \tag{4.69}$$

$$J^0 = t_i(K^{-1})^{ik}\delta^{(2)}(\vec{0}), \tag{4.70}$$

$$q(a^k) = \int d^2x J^0 = t_i(K^{-1})^{ik}. \tag{4.71}$$

Therefore, the quasi-hole charge associated with field a_μ^k is $t_i(K^{-1})^{ik}$. In terms of the exchange statistics, we can see from eq. 4.66,

$$\frac{1}{2\pi}f_{12}^j = (K^{-1})^{jk}\delta^{(2)}(\vec{0}), \tag{4.72}$$

so each quasi-hole of type j has an attached flux of size $2\pi(K^{-1})^{jk}$. Hence, each quasi-hole is a flux-charge composite. The exchange angle for switching the positions of a k particle with a j particle is therefore,

$$\theta^{jk} = \pi(K^{-1})^{jk}. \tag{4.73}$$

This analysis, obtained from the action in eq. 4.55, is almost enough to classify all Abelian FQHE states. All it misses is the "shift" (related to the degeneracy of the ground state) when we consider quantum Hall states in manifolds with a non-trivial topology. But, for quantum Hall effects in a flat disk, this is all there is.

4.4 The Non-Abelian Fractional Effect

As one might expect, Chern-Simons theory with non-Abelian gauge groups gives rise to non-Abelian exchange statistics. We already saw this in section 3.6, the $SU(2)_2$ Chern-Simons theory describes Ising anyons, and $SU(2)_3$ describes Fibonacci anyons plus some others. Both of which are non-Abelian, so one might naively expect the non-Abelian states to follow the same hierarchy but with a different gauge group. Unfortunately, this is far from the truth. Firstly, in our analysis above, we wrote the current to which the EM field couples as $\frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu a_\rho = \frac{1}{4\pi}\epsilon^{\mu\nu\rho}f_{\nu\rho}$. This is a reasonable guess as it is gauge invariant. However, if the gauge group is not $U(1)$, then this is no longer true. The usual way to make a gauge invariant quantity in Yang-Mills theories is to take the trace. However, this does not help us as while it is true that $\frac{1}{4\pi}\epsilon^{\mu\nu\rho}\,\mathrm{Tr}(f_{\nu\rho})$ is gauge invariant, it also vanishes. Nevertheless, advancements have been make by considering the gauge group $U(N) = U(1) \times SU(N)/\mathbb{Z}_N$ or similar [34]. The EM field can couple to the $U(1)$ factor, but coupling the $U(1)$ factor to the rest is more complicated [35]. Other attempts include mapping the interacting electrons under a magnetic field B, to one of interacting flux-charge composites under a magnetic field $B - 2\phi_0 n$, with each electron having a flux attached to them of size $2\phi_0$ [36]. This is known as the composite fermion theory. Other non-field theoretic approaches include guessing the answer, which after all worked for the Laughlin $\nu = 1/m$ states. In particular, the Pfaffian wavefunction is thought to describe the $\nu = 5/2$ non-Abelian state [37]. However, as of now, there lacks a complete description of the non-Abelian states.

Chapter 5

Filling Anomaly in Higher Order Topological Insulators

5.1 Introduction

As we have seen, the Chern-Simons action is not gauge invariant in a manifold with a boundary. To cancel this non-gauge invariance, one must add to the action a local boundary term. The addition of this term often creates a surplus charge in the manifold, known as a filling anomaly. Famously, the FQHE has a filling anomaly, which is to be expected as we have shown that it can be expressed as a series of Chern-Simons terms. More recently (May 2023), a similar filling anomaly in Higher-Order Topological Insulators (HOTIs) with C_{2n} symmetry was understood in terms of the low-energy effective field theory [38]. In this chapter, we review this result as an example of how TQFT is used in condensed matter. HOTIs are topological crystalline insulators (the energy spectrum in the bulk is gapped), with gapless topological defects. The defects in question are corners at the boundary and disclination defects. Disclination defects are geometrical defects in which the discrete local rotational invariance is broken. An example of a C_4 disclination defect is shown in figure 5.1. As these defects are the only gapless ones, they dominate the low-energy dynamics and often carry a fractional charge [41]. As such, when we consider the action describing a C_4 HOTI, we will consider the curvature (or rather, the spin connection) as one of the dynamic fields. However, even in the absence of corner and disclination defects, there remains an excess charge in the bulk, known as a filling anomaly. This filling anomaly will be explained in terms of the non-gauge invariance of a Wen-Zee (WZ) term in the action, which must be compensated by a boundary Gromov-Jensen-Abanov (GJA) term. Thanks to this addition, the theory does not require a gapless edge mode, but rather a coupling between electromagnetism and the extrinsic curvature in the gapped boundary theory. The GJA term induces a charge on the boundary, related to the integral over the extrinsic curvature, explaining the filling anomaly.

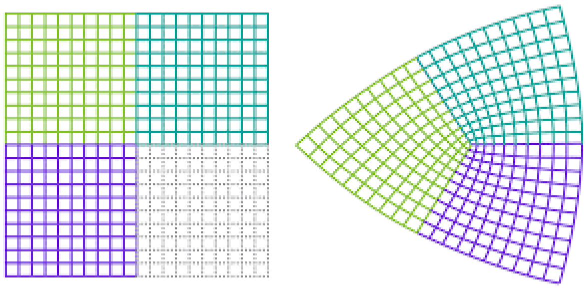


Figure 5.1: A C_4 disclination defect. On the left figure, a square lattice is shown which is missing one of the four sections (shown in grey), and the side right side of the purple, and bottom side of the blue sections, represent the same line. On the right, the lattice has been curved to join the purple and blue regions. Note that every point except the center has a local C_4 symmetry, while the center has a C_3 symmetry. In the limit as the lattice spacing vanishes, this space becomes isomorphic to a cone with deficit angle $\pi/2$, and the Riemann curvature becomes a Dirac δ function [39]. Image obtained from [40].

5.2 The Wen-Zee and Gromov-Jensen-Abanov Actions

Suppose we are on a 2+1 dimensional spacetime manifold $\mathcal{M} = \mathbb{R}\times\Gamma$, where $\mathbb{R}$ represents time and Γ the spatial manifold, which in general will have a boundary $\partial\Gamma$. As $\mathbb{R}$ has no boundary, $\partial\mathcal{M} = \mathbb{R}\times\partial\Gamma$. We will insist that the boundary is gapped (there is an energy gap to excitations), and coupled to an external electromagnetic field $A = A_\mu dx^\mu$ and the spin connection ω. For more information on the spin connection in 2+1 dimensions, see appendix D. For a gapped system, we expect the bulk action to be a local functional of the fields. This is because if there exists an energy gap ΔE, any correlations arise at lengths of order $1/\Delta E$, which is small compared to the long wavelength dynamics we are interested in. We can hence write the action as

$$S[A, e, \omega] = \int_{\mathcal{M}} \mathcal{L}_{\text{loc}}(e, F, R) + \mathcal{L}_{\text{top}}(A, e, \omega), \tag{5.1}$$

where $\mathcal{L}_{\text{loc}}$ is some general local contribution, which is invariant under electromagnetic $U(1)$ gauge symmetry and local $SO(2)$ rotations of the frame e^a_i, defined in appendix D. $F = dA$ is the electromagnetic field strength and $R = d\omega$ is the curvature. $\mathcal{L}_{\text{top}}$ is a topological term, which is gauge invariant up to a boundary term. Note that the action described above is the bulk action, so any non-gauge invariance by a boundary term can be compensated by local counter terms in the boundary theory.

We are interested in energy levels much lower than the energy gap, $\frac{1}{\Delta E}\int_A d^2x\rho \ll 1$, where ρ is the electromagnetic energy density and A is the area of a primitive unit cell.

This means that

$$\frac{A\rho}{\Delta E} = \frac{\epsilon_0}{2}\frac{AF_{i0}^2}{\Delta E} + \frac{1}{2\mu_0}\frac{AF_{12}^2}{\Delta E} \ll 1, \tag{5.2}$$

$$\frac{AF_{\mu\nu}^2}{\Delta E} \ll 1, \tag{5.3}$$

Given this condition, the leading terms in the action are those with the fewest powers of A. As explained in section 1, it is not possible to write a term with a single field that is gauge invariant up to a boundary term. However, with two fields, there are several terms we can write down, including

$$A \wedge dA, \tag{5.4}$$

$$\omega \wedge d\omega, \tag{5.5}$$

$$A \wedge d\omega, \tag{5.6}$$

the first two are individual Chern-Simons terms, and the last is known as the Wen-Zee action up to a constant prefactor. In order to isolate the C_4 HOTI filling anomaly, we can force the Hall conductivity to vanish (and hence not worry about the quantum Hall filling anomaly) by picking the level of the Chern-Simons term of A to be zero. Moreover, we are only interested in the charge distribution, found by taking functional derivatives with respect to A, so the term $\omega \wedge d\omega$ is irrelevant and will therefore be neglected. Hence, the leading contribution to $\mathcal{L}_{top}$ is

$$S_{\text{top}} = S_{\text{WZ}} = \frac{\bar{s}}{2\pi}\int_{\mathcal{M}} A \wedge d\omega, \tag{5.7}$$

$$= \frac{\bar{s}}{2\pi}\int_{\mathcal{M}} d^3x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu \omega_\rho, \tag{5.8}$$

where $\bar{s}$ is the "orbital spin per particle" [38], and is related to the Hall viscosity. Varying the Wen-Zee action with respect to the vector potential, we obtain an expression for the ground state 4-current,

$$J^\mu = \frac{\delta S_{\text{WZ}}}{\delta A_\mu} = \frac{\bar{s}}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu\omega_\rho, \tag{5.9}$$

$$\rho = J^0 = \frac{\bar{s}}{2\pi}\widetilde{R}, \tag{5.10}$$

where $\widetilde{R} = \epsilon^{0\nu\rho}\partial_\nu\omega_\rho = \partial_1\omega_2 - \partial_2\omega_1$ is the spatial curvature on Γ[1]. In particular, a disclination defect of deficit angle α at $\vec{x_0}$, causes a curvature $\widetilde{R} = \alpha\delta^2(\vec{x} - \vec{x_0})$ [39] in the limit as the lattice spacing vanishes (this can be understood as when a vector is parallel transported around a conical singularity, it rotates by the deficit angle). As we are interested in low-energy, long-wavelength dynamics, we will use this approximation. In a C_4 HOTI, the most common disclination defect has a deficit angle of $\alpha = \frac{\pi}{2}$ (although

[1] In ref [38], the authors incorrectly quote the charge density over neutrality to be $\rho = \frac{\bar{s}}{2\pi}d\omega$, and then call $d\omega = R$, which is an abuse of notation

π and $3\pi/2$ are also possible), in which case the charge of the defect is $Q = \frac{2}{2\pi} \times \frac{\pi}{2} = \frac{1}{2}$, where we have used $\bar{s} = 2$ as it has been shown that 2D magnetic, C_4 symmetric HOTIs are governed by a WZ action with $\bar{s} = 2$ [42]. Integrating the charge density over all space, we find the total charge above neutrality,

$$Q_{\text{bulk}} = \frac{\bar{s}}{2\pi} \int_{\Gamma} d^2 x \widetilde{R}. \tag{5.11}$$

However, in flat geometries where $\omega = 0$, Q_{bulk} vanishes. Therefore, the Wen-Zee action alone does not explain the filling anomaly. Now, consider the boundary theory. We require the overall action to be $U(1)$ gauge invariant. Hence, consider the Wen-Zee action under a gauge transformation[2],

$$\begin{aligned} S_{\text{WZ}}[A, \omega, \bar{s}] &\rightarrow S_{\text{WZ}}[A + df, \omega, \bar{s}], &(5.12)\\ &= S_{\text{WZ}}[A, \omega, \bar{s}] + \frac{\bar{s}}{2\pi} \int_{\mathcal{M}} df \wedge d\omega, &(5.13)\\ &= S_{\text{WZ}}[A, \omega, \bar{s}] + \frac{\bar{s}}{2\pi} \int_{\mathcal{M}} d(f \wedge d\omega), &(5.14)\\ &= S_{\text{WZ}}[A, \omega, \bar{s}] - \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} f \wedge d\omega, &(5.15) \end{aligned}$$

where we have used $d^2 = 0$ in eq. 5.14, and Stokes law in eq. 5.15. The negative sign arises from the change in orientation of $\partial\mathcal{M}$ used in Gauss' law to the convention described in [43]. The local boundary counter term that cancels this non-gauge invariance is the Gromov-Jensen-Abanov action. Let K be the extrinsic curvature of $\partial\mathcal{M}$, which depends on the embedding of $\partial\mathcal{M}$ in $\mathbb{R} \times \mathbb{R}^2$, and is defined in appendix E. From this, we can write the GJA term,

$$\begin{aligned} S_{\text{GJA}}[A, K, \bar{s}] &= \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} A \wedge K, &(5.16)\\ &= \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} d^2 x \epsilon^{\bar{\mu}\bar{\nu}} A_{\bar{\mu}} K_{\bar{\nu}}. &(5.17) \end{aligned}$$

The extrinsic curvature differs from the spin connection evaluated at the boundary by a closed form (for proof, see the supplementary information of [43]),

$$\partial_{\bar{\mu}} \alpha = (\omega|_{\partial\mathcal{M}})_{\bar{\mu}} + K_{\bar{\mu}}, \tag{5.18}$$

where by $(\omega|_{\partial\mathcal{M}})_{\bar{\mu}}$ we mean the following. Let $\sigma^{\bar{\mu}}$ be a coordinate on $\partial\mathcal{M}$. The embedding of $\partial\mathcal{M}$ can be characterized by embedding functions $X^{\mu}(\sigma^{\alpha})$. From this, we can define a tensor field $f^{\mu}_{\bar{\mu}} = \partial_{\bar{\mu}} X^{\mu}$ that allows us to project tensors in $\mathcal{M}$ to $\partial\mathcal{M}$ by $\omega_{\bar{\mu}} = f^{\mu}_{\bar{\mu}} \omega_{\mu}$. Please note that for convenience, we have used a different convention for the range of indices in the appendix. Using this, we can see how the GJA term transforms under a

[2]In ref [38], the authors miss the factor $\bar{s}/2\pi$ in eq. 5.13, but correct it on the next line

$U(1)$ gauge transformation,

$$S_{\mathrm{GJA}}[A, K, \bar{s}] \to S_{\mathrm{GJA}}[A + df, K, \bar{s}], \tag{5.19}$$

$$= S_{\mathrm{GJA}}[A, K, \bar{s}] + \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} df \wedge K, \tag{5.20}$$

$$= S_{\mathrm{GJA}}[A, K, \bar{s}] + \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} df \wedge (d\alpha - \omega), \tag{5.21}$$

$$= S_{\mathrm{GJA}}[A, K, \bar{s}] - \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} df \wedge \omega, \tag{5.22}$$

$$= S_{\mathrm{GJA}}[A, K, \bar{s}] + \frac{\bar{s}}{2\pi} \int_{\partial\mathcal{M}} f \wedge d\omega. \tag{5.23}$$

Here we have used the fact that $df \wedge d\alpha = d(f \wedge d\alpha)$ (as $d^2 = 0$) is a total derivative and over a closed manifold the integral vanishes, and that $d(f \wedge \omega) = df \wedge \omega + f \wedge d\omega$, from which the left hand side vanishes when integrated by Stokes law, as $\partial^2\mathcal{M} = 0$. The extra term is the same as the one obtained for the WZ bulk action but with the opposite sign.

5.3 The Filling Anomaly

Therefore, we are forced to conclude that if the boundary has a gap to excitation, then the leading order term in the action is the Wen-Zee action, which is not gauge invariant. This must be compensated by a GJA counterterm that induces a current on the boundary. In other words, the action describing the system is

$$S_{\mathrm{total}} = \frac{\bar{s}}{2\pi} \left(\int_{\mathcal{M}} A \wedge d\omega + \int_{\partial\mathcal{M}} A \wedge K \right), \tag{5.24}$$

is $U(1)$ gauge invariant. The inclusion of the GJA term to the action also gives rise to a non-vanishing 4-current in the boundary,

$$J^{\bar{\mu}}_{\mathrm{FA}} = \frac{\bar{s}}{2\pi} \epsilon^{\bar{\mu}\bar{\nu}} K_{\bar{\nu}}, \tag{5.25}$$

where FA stands for filling anomaly. We can work out the charge above neutrality in the boundary by integrating over the $\bar{\mu} = 0$ term,

$$Q_{\mathrm{boundary}} = \int_{\partial\Gamma} J^0_{\mathrm{FA}} = \frac{\bar{s}}{2\pi} \int_{\partial\Gamma} dx \widetilde{K}, \tag{5.26}$$

where $\widetilde{K} = K_1$[3] is the spatial curvature in $\partial\Gamma$. Therefore, the total charge over neutrality in a magnetic HOTI is given by

$$Q_{\mathrm{FA}} = \frac{\bar{s}}{2\pi} \left(\int_{\Gamma} d^2x \widetilde{R} + \int_{\partial\Gamma} dx \widetilde{K} \right) = \bar{s}\chi_d, \tag{5.27}$$

[3] The authors just have $\int_{\partial\Gamma} K$, which again is an abuse of notation.

where we have used the Gauss-Bonnet theorem [44, 38] and χ_d is the Euler characteristic of Γ ($V-E+F$, where V is the number of vertices, E is the number of edges and F is the number of faces in a graph inscribed within Γ), equivalent to how many times the function α winds around $\partial\Gamma$ [43]. In the absence of disclination defects, and assuming $\Gamma \cong D^2$ (the 2D disk), then $\chi_d = 1$, and we recover the known result of a charge above neutrality of $\bar{s}$. In the absence of disclinations, $\widetilde{R} = 0$, and all the charge is concentrated in the boundary, consistent with observations [45]. In conclusion, we have explained the filling anomaly in C_4-symmetric, magnetic HOTIs by insisting on having a $U(1)$ gauge invariant local theory and considering the leading terms in the bulk action. The only terms we can write down transform by a boundary term, which is canceled by a counter term in the boundary theory. This boundary term induces an anomalous current responsible for the filling anomaly.

Bibliography

1. Shapere A and Wilczek F. Geometric phases in physics. Vol. 5. World scientific, 1989
2. Berry MV. Quantal phase factors accompanying adiabatic changes. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 1984; 392:45–57
3. Simon S. Topological Quantum. 2016
4. Preskill J. Lecture notes for Physics 219: Quantum computation. Caltech Lecture Notes 1999; 7
5. Aharonov Y and Bohm D. Significance of electromagnetic potentials in the quantum theory. Physical Review 1959; 115:485
6. Zhu W and Lee Y. Charge-flux composites and fractional statistics. Physical Review B 1995; 51:10179
7. Tong D. Lectures on the quantum Hall effect. arXiv preprint arXiv:1606.06687 2016
8. Pillai SU, Suel T, and Cha S. The Perron-Frobenius theorem: some of its applications. IEEE Signal Processing Magazine 2005; 22:62–75
9. Mac Lane S. Categories for the working mathematician. Vol. 5. Springer Science & Business Media, 2013
10. Bonderson PH. Non-Abelian anyons and interferometry. California Institute of Technology, 2012
11. Etingof P, Nikshych D, and Ostrik V. On fusion categories. Annals of mathematics 2005 :581–642
12. Rowell E, Stong R, and Wang Z. On classification of modular tensor categories. Communications in Mathematical Physics 2009; 292:343–89
13. Grabovsky D. Chern-Simons theory in a knotshell. Accessed: May 19th 2022
14. Witten E. Quantum field theory and the Jones polynomial. Communications in Mathematical Physics 1989; 121:351–99
15. Dunne GV. Aspects of chern-simons theory. *Aspects topologiques de la physique en basse dimension. Topological aspects of low dimensional systems: Session LXIX. 7–31 July 1998.* Springer, 2002 :177–263
16. Moore GW. Introduction to chern-simons theories. Available on the internet 2019; 3:14

17. Witten E. Three lectures on topological phases of matter. La Rivista del Nuovo Cimento 2016; 39:313–70
18. Polyakov AM. Fermi-Bose transmutations induced by gauge fields. Modern Physics Letters A 1988; 3:325–8
19. Ricca RL and Nipoti B. GAUSS'LINKING NUMBER REVISITED. Journal of Knot Theory and Its Ramifications 2011; 20:1325–43
20. Coleman S. Aspects of symmetry: selected Erice lectures. Cambridge University Press, 1988
21. Astorino M. Kauffman knot invariant from SO (N) or Sp (N) Chern-Simons theory and the Potts model. Physical Review D 2010; 81:125026
22. Schwarz AS. The partition function of degenerate quadratic functional and Ray-Singer invariants. Letters in Mathematical Physics 1978; 2:247–52
23. Verlinde E. Fusion rules and modular transformations in 2D conformal field theory. Nuclear Physics B 1988; 300:360–76
24. Blau M and Thompson G. Derivation of the Verlinde formula from Chern-Simons theory and the G/G model. Nuclear Physics B 1993; 408:345–90
25. Pachos JK. Introduction to topological quantum computation. Cambridge University Press, 2012
26. Di Francesco P, Mathieu P, and Senechal D. Conformal Field Theory. Graduate Texts in Contemporary Physics. New York: Springer-Verlag, 1997. DOI: 10.1007/978-1-4612-2256-9
27. Tong D. Gauge theory. Lecture notes, DAMTP Cambridge 2018; 10
28. Achucarro A and Townsend PK. A Chern-Simons action for three-dimensional anti-de Sitter supergravity theories. Physics Letters B 1986; 180:89–92
29. Laughlin RB. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. Physical Review Letters 1983; 50:1395
30. Lee D. Quantum Theory of Matter. Chapter 7: Electrons in a Magnetic Field. Imperial College London, 2013
31. Klitzing K von. The Quantized Hall Effect. Nobel lecture. 1985
32. Laughlin RB. Quantized Hall conductivity in two dimensions. Physical Review B 1981; 23:5632
33. Byers N and Yang C. Theoretical considerations concerning quantized magnetic flux in superconducting cylinders. Physical review letters 1961; 7:46
34. Cabra DC, Fradkin E, Rossini GL, and Schaposnik FA. Non-Abelian fractional quantum Hall states and chiral coset conformal field theories. International Journal of Modern Physics A 2000; 15:4857–70
35. Seiberg N and Witten E. Gapped boundary phases of topological insulators via weak coupling. Progress of Theoretical and Experimental Physics 2016; 2016:12C101
36. Stern A. Non-Abelian states of matter. Nature 2010; 464:187–93
37. Moore G and Read N. Nonabelions in the fractional quantum Hall effect. Nuclear Physics B 1991; 360:362–96

38. Rao P and Bradlyn B. Effective action approach to the filling anomaly in crystalline topological matter. Physical Review B 2023; 107:195153
39. Tod K. Conical singularities and torsion. Classical and Quantum Gravity 1994; 11:1331
40. Slagle K. Crystal Defects Mimic Elusive Fractons. Physics 2018; 11:43
41. Peterson CW, Li T, Jiang W, Hughes TL, and Bahl G. Observation of trapped fractional charge and topological states at disclination defects in higher-order topological insulators. arXiv preprint arXiv:2004.11390 2020
42. May-Mann J and Hughes TL. Crystalline responses for rotation-invariant higher-order topological insulators. Physical Review B 2022; 106:L241113
43. Gromov A, Jensen K, and Abanov AG. Boundary effective action for quantum Hall states. Physical review letters 2016; 116:126802
44. Eguchi T, Gilkey PB, and Hanson AJ. Gravitation, gauge theories and differential geometry. Physics reports 1980; 66:213–393
45. Benalcazar WA, Li T, and Hughes TL. Quantization of fractional corner charge in C n-symmetric higher-order topological crystalline insulators. Physical Review B 2019; 99:245151
46. Schuller FP. Lectures on the geometric anatomy of theoretical physics. Institute for Quantum Gravity, Friedrich-Alexander Universität Erlangen-Nürnberg 2015
47. Tolley A. Advanced Quantum Field Theory. Imperial College London, 2016
48. Gribov VN. Quantization of non-Abelian gauge theories. Nuclear Physics B 1978; 139:1–19
49. Hull C. Differential Geometry. Imperial College London, 2022
50. Spivak M. A comprehensive introduction to differential geometry, Publish or Perish. Inc., Berkeley 1979; 2

Appendix A

Cyclic Time and Thermal Physics

In this section, we argue that a Wick rotation, $\tau = it$, turns the usual path integral into the thermal partition function. In doing so, the manifold associated with time goes from being isomorphic to $\mathbb{R}$ to S^1. A single particle with position q moving in a potential $V(q)$ has a corresponding thermal partition function

$$\mathcal{Z}[\beta] = \mathrm{Tr}\, e^{-\beta H}, \tag{A.1}$$

where H is the Hamiltonian and β is the inverse temperature. Let us now write a path integral for this expression. The path integral the amplitude for a given state $|q_i\rangle$ to evolve into some other state $|q_f\rangle$ is,

$$\langle q_f| \, e^{-itH} \, |q_i\rangle = \int_{q(0)=q_i}^{q(t)=q_f} \mathcal{D}q e^{iS}, \tag{A.2}$$

where S is the action. We can already appreciate a similarity between these two descriptions, one being ruled by $e^{-\beta H}$ and the other by e^{-itH}. To be more explicit, let us perform a Wick rotation $\tau = it$. The classical action transforms as

$$S = \int_0^t dt' \left[\frac{m}{2} \left(\frac{dq}{dt'} \right)^2 - V(q) \right], \tag{A.3}$$

$$= -i \int_0^{-i\tau} d\tau' \left[-\frac{m}{2} \left(\frac{dq}{d\tau'} \right)^2 - V(q) \right] = iS^E, \tag{A.4}$$

where S^E is the Euclidean action. Now, to make the connection with thermal field theory, suppose the particle evolves by an Euclidean time $\tau = \beta$.

$$\langle q_f| \, e^{-\beta H} \, |q_i\rangle = \int_{q(0)=q_i}^{q(\beta)=q_f} \mathcal{D}q e^{-S^E} \tag{A.5}$$

Now consider the trace over of $e^{-\beta H}$,

$$\mathrm{Tr}\, e^{-\beta H} = \int dq_i \left\langle q_i \right| e^{-\beta H} \left| q_i \right\rangle , \tag{A.6}$$

$$= \int dq_i \int_{q(0)=q_i}^{q(\beta)=q_i} \mathcal{D}q e^{-\beta H}, \tag{A.7}$$

$$= \int_{q(0)=q(\beta)} \mathcal{D}q e^{-\beta H}. \tag{A.8}$$

From this, we can conclude that a Wick rotation of the quantum path integrals allows us the write the thermal partition function in terms of a path integral subject to the constraint that all paths $q(\tau)$ are periodic, with periodicity β. This is analogous to setting the equivalence $\tau = \tau + \beta$, meaning that the manifold describing the possible values of τ is isomorphic to S^1.

Appendix B

Brief Introduction to Bundles

Throughout this work, it is assumed that the reader is comfortable with basic differential geometry, including topological spaces, manifolds, differential forms, connections, covariant derivatives and curvature. However, many aspects of section 3.6 are deeply tied to the idea of bundles. This appendix will mostly follow [46]. To define bundles, we need three ingredients, a topological space $\mathcal{M}$ known as the "base space", another topological space E, known as the total "space", and a continuous map

$$\pi : \; E \to \mathcal{M}. \tag{B.1}$$

The map π is known as the projection. Fixing a point $p \in \mathcal{M}$, the preimage of the projection at p is known as a fibre, F_p,

$$F_p = \text{preim}_\pi(\{p\}). \tag{B.2}$$

Therefore, one can express the total space E as the union of all the fibres. If all fibres are isomorphic to each other, $F_p \cong F \;\; \forall p \in \mathcal{M}$, then $(E, \pi, \mathcal{M})$ is a fibre bundle with typical fibre F. Next, one can define a section of a bundle as a map

$$\sigma : \; \mathcal{M} \to E, \tag{B.3}$$

such that $\pi \circ \sigma = \text{id}_\mathcal{M}$, the identity map on $\mathcal{M}$. Suppose $\mathcal{M}$ represents spacetime and E is a Lie group G. Then, sections of the bundle $(E, \pi, \mathcal{M})$ are maps from spacetime to an element of the group. In other words, a section of the bundle assigns to each point in spacetime an element of G. This is just what we call a local gauge transformation. Moreover, we define the principal G-connection on E as a $\mathfrak{g}$-valued 1-form. As a quick aside, we know that in non-Abelian Yang-Mills theory, the vector potential takes values in the Lie algebra of the gauge group. This is exactly the same as a G-connection. Therefore, the gauge field is what we call a principal connection in the principal bundle. Finally, we say that a fibre bundle with typical fibre F is trivial if $E \cong \mathcal{M} \times F$. The topic of bundles is a very big topic, and we have only scraped the surface. Nevertheless, this should be sufficient to understand any mention of bundles in this dissertation.

Appendix C

Faddeev-Popov-DeWitt Gauge Fixing Procedure

This working out is adapted from [47]. We would like to work in a modified Lorenz gauge,

$$F(a) = D_\mu a^\mu = 0. \tag{C.1}$$

The way to do that is to add a functional Dirac δ function to the path integral which ensures we have chosen the gauge that guarantees eq. C.1 is satisfied. Then, we will express the functional δ function in terms of new fields. We will end by dividing by the volume of the gauge orbit to ensure the path integral converges. As the first step is to add a δ function to make sure we are on the right gauge, let us find what gauge transformation that is. Suppose we are on a field configuration which does not satisfy eq. C.1, and want to perform a gauge transformation to obey the gauge condition,

$$a'_\mu = a_\mu + D_\mu \chi_*, \tag{C.2}$$

$$D^\mu a'_\mu = D^\mu a_\mu + D^\mu D_\mu \chi_* = 0, \tag{C.3}$$

$$\chi_* = -\frac{1}{D^\mu D_\mu} D^\nu a_\nu. \tag{C.4}$$

Here one has to worry whether $1/D^\mu D_\mu$ is defined, as $D^\mu D_\mu$ could have zero modes. It turns out that in perturbation theory it is straightforward to give meaning to this operator, and non-perturbatively it has to do with the Gribov ambiguity [48]. We can continue by defining a functional Dirac δ function in the way one might expect,

$$\int \mathcal{D}\chi \delta_F(\chi - \chi_*(a)) = 1. \tag{C.5}$$

Inserting this into the path integral,

$$\mathcal{Z}_{old} = \int \mathcal{D}a_\mu \int \mathcal{D}\chi \delta_F(\chi - \chi_*(a)) e^{iS_{CS}[a]}. \tag{C.6}$$

The functional δ function obeys

$$\delta_F(F(a)) = \frac{\delta_F(\chi - \chi_*(a))}{\left|\det\left(\frac{\delta F(a')}{\delta\chi}\right)|_{\chi_*}\right|}, \tag{C.7}$$

where a' solves $F(a') = 0$. Rearranging the equation for $\delta_F(\chi - \chi_*(a))$, the path integral becomes

$$\mathcal{Z}_{old} = \int \mathcal{D}a_\mu \int \mathcal{D}\chi \det\left(\frac{\delta F(a'(x))}{\delta\chi(y)}\right)\delta_F(F(a'))e^{iS_{CS}[a]}. \tag{C.8}$$

In this particular case,

$$F(a) = D_\mu a^\mu, \tag{C.9}$$

$$F(a') = D_\mu a^\mu + D_\mu D^\mu \chi, \tag{C.10}$$

$$\frac{\delta F(a'(x))}{\delta\chi(y)} = D_\mu D^\mu \delta^{(3)}(x-y). \tag{C.11}$$

So an explicit expression for the path integral is

$$\mathcal{Z}_{old} = \int \mathcal{D}a_\mu \int \mathcal{D}\chi \det(D_\mu D^\mu)\delta_F(D_\mu a^\mu + D_\mu D^\mu\chi)e^{iS_{CS}[a]}. \tag{C.12}$$

We can now perform a field redefinition to get rid of any χ in the path integral,

$$a^\mu \to a^\mu - D^\mu\chi, \tag{C.13}$$

$$\int \mathcal{D}a_\mu \to \int \mathcal{D}a_\mu, \tag{C.14}$$

$$e^{iS_{CS}[a]} \to e^{iS_{CS}[a]}, \tag{C.15}$$

the last of which is a consequence of the quantized level as already discussed. Implementing this field redefinition and dividing by the volume of the gauge orbit $\int \mathcal{D}\chi$,

$$\mathcal{Z}_{old} = \int \mathcal{D}a_\mu \int \mathcal{D}\chi \det(D_\mu D^\mu)\delta_F(D_\mu a^\mu)e^{iS_{CS}[a]}, \tag{C.16}$$

$$\mathcal{Z}_{new} = \frac{\mathcal{Z}_{old}}{\int \mathcal{D}\chi} = \int \mathcal{D}a_\mu \det(D_\mu D^\mu)\delta_F(D_\mu a^\mu)e^{iS_{CS}[a]}. \tag{C.17}$$

Finally, we may express the determinant and the functional δ function in terms of new fields

$$\delta_F(D_\mu a^\mu) = \int \mathcal{D}\phi e^{i\int d^3x \mathrm{Tr}(\phi D_\mu a^\mu)}, \tag{C.18}$$

$$\det(D_\mu D^\mu) = \int \mathcal{D}\bar{c} \int \mathcal{D}c e^{-\int d^3x\, \mathrm{Tr}(\bar{c}D_\mu D^\mu c)}, \tag{C.19}$$

which makes the final expression for the path integral

$$\mathcal{Z}_{new} = \int \mathcal{D}a_\mu \int \mathcal{D}\phi \int \mathcal{D}\bar{c} \int \mathcal{D}c e^{iS}, \tag{C.20}$$

where

$$S = \int_{\mathcal{M}} d^3x \, \mathrm{Tr}\left(\frac{k}{4\pi} \epsilon^{\mu\nu\rho} a_\mu D_\nu a_\rho + \phi D_\mu a^\mu + i\bar{c} D_\mu D^\mu c \right) \tag{C.21}$$

Appendix D

The Spin Connection

We will make heavy use of the spin connection in sections 3.6 and 5. Therefore, while we assume the reader is comfortable with differential geometry, we will recall some important results. In particular, we will point out that the spin connection has a single degree of freedom in 2+1 dimensions in a manifold of the form $\mathbb{R}\times\Gamma$, $\mathbb{R}$ signifying time and Γ being a spatial manifold. This section is based on [43, 49]. We will start by defining a frame β_a^μ and the corresponding co-frame $(\beta^{-1})_\mu^a$, where $a \in \{0,1,2\}$ represents the vector or co-vector, and $\mu \in \{0,1,2\}$ are the components. The definition of the spin connection ω, given a connection Γ_{ca}^b, and a vielbein $(\beta^{-1})_\mu^a$,

$$\omega^a{}_{b\mu} = (\beta^{-1})_\mu^c \Gamma_{cb}^a. \tag{D.1}$$

As we have chosen time to be different in the sense that the manifold Γ at any time slice is the same, let us separate the time and space sections,

$$v^\mu = \beta_0^\mu \qquad n_\mu = (\beta^{-1})_\mu^0, \tag{D.2}$$

$$E_A^\mu = \beta_A^\mu \qquad e_\mu^A = (\beta^{-1})_\mu^A, \tag{D.3}$$

where $A \in \{1,2\}$. For this manifold and for a metric connection, we can restrict spin connection to only have anti-symmetric spatial components,

$$\omega_{0\mu}^a = \omega_{b\mu}^0 = \omega_\mu^{(AB)} = 0, \tag{D.4}$$

where the parenthesis denotes symmetrization of indices, and we have raised the index in the last expression with δ^{AB}. We can see now that the metric connection is now a simple 1-form, as opposed to the general case in which it is a matrix-valued 1-form. As A, B only range over 2 values, and ω is anti-symmetric, we can write

$$\omega^A{}_{B\mu} = \epsilon^A{}_B \omega_\mu, \tag{D.5}$$

where $\epsilon^A{}_B$ is the Levi-Civita symbol and $\epsilon^1{}_2 = 1$. Using the frames, we can now define a metric for the 3-manifold as

$$\gamma_{\mu\nu} = e^a_\mu e^b_\nu \delta_{ab} \qquad\qquad \gamma^{\mu\nu} = E^\mu_a E^\nu_b \delta^{ab}. \tag{D.6}$$

If we are only interested in the spatial metric in Γ, this is given by $g_{ij} = e^A_i e^B_j \delta_{AB}$. Consider now Cartan's structure equations (we will call $\omega = \omega_\mu dx^\mu$ the one form, and $R^a{}_b = \frac{1}{2} R^a{}_{b\mu\nu} dx^\mu \wedge dx^\nu$ the curvature 2-form),

$$R^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b. \tag{D.7}$$

Using $\omega^A{}_B = \epsilon^A{}_B \omega$, $\epsilon^A{}_C \epsilon^C{}_B = -\delta^A_B$, and $\omega \wedge \omega = 0$ by contraction of symmetric and anti-symmetric indices,

$$R^a{}_b = \epsilon^A{}_B d\omega, \tag{D.8}$$

which means that $R^A{}_B$ is anti-symmetric with respect to the exchange of indices, and hence can be written as $R^A{}_B = \epsilon^A{}_B R$ (and only has spatial components), so

$$R = d\omega. \tag{D.9}$$

This is a remarkably simple result and explains how the spin connection depends on sources of curvature.

Appendix E

The Extrinsic Curvature

In this chapter, we develop the notion of the extrinsic curvature, used in section 5.2. Further, we will use the notation developed in Appendix D but will work at the level of the connection and embedding functions. As the extrinsic curvature depends on the particular embedding of $\partial\mathcal{M}$, we must characterize this mathematically. This is done through embedding functions $X^\mu(\sigma^\alpha)$, where $\alpha \in \{0,1\}$ are directions in $\partial\mathcal{M}$. While $X^\mu(\sigma^\alpha)$ is not a tensor, we can build one as

$$f^\mu_\alpha = \partial_\alpha X^\mu. \tag{E.1}$$

This tensor allows us to pullback covectors from $\mathcal{M}$ to $\partial\mathcal{M}$. For example,

$$n_\alpha = f^\mu_\alpha n_\mu. \tag{E.2}$$

Using the metric on $\mathcal{M}$, $\gamma_{\mu\nu}$, we can swap the indices on f^μ_α,

$$f^\alpha_\mu = \gamma^{\alpha\beta}\gamma_{\mu\nu} f^\nu_\beta. \tag{E.3}$$

As one might expect we can use this map to pullback vectors,

$$v^\alpha = f^\alpha_\mu v^\mu. \tag{E.4}$$

With this information, one can construct a covector normal to $\partial\mathcal{M}$,

$$N_\mu = \frac{1}{2}\epsilon_{\mu\nu_1\nu_2}\epsilon^{\alpha_1\alpha_2} f^{\nu_1}_{\alpha_1} f^{\nu_2}_{\alpha_2}, \tag{E.5}$$

which has the property

$$N_\alpha = f^\nu_\alpha N_\nu = \frac{1}{2}\epsilon_{\mu\nu_1\nu_2}\epsilon^{\alpha_1\alpha_2} f^\mu_\alpha f^{\nu_1}_{\alpha_1} f^{\nu_2}_{\alpha_2} = 0. \tag{E.6}$$

The last step comes from the fact that α runs from 1 to 2, and $\alpha_1 \neq \alpha_2$ because of the contraction with the Levi-Civita symbol. By the pigeonhole principle, α must be equal to either α_1 or α_2. Suppose $\alpha = \alpha_1$, then, the exchange of μ and ν_1 is antisymmetric on the Levi-Civita symbol but symmetric on the f tensors, meaning the contraction vanishes. As N_μ has a vanishing projection to ∂M, we conclude it is normal to the surface. N_μ is normalized,

$$N^\mu N_\mu = \frac{1}{2}(N^\mu \epsilon_{\mu\nu_1\nu_2} f^{\nu_1}_{\alpha_1} f^{\nu_2}_{\alpha_2})\epsilon^{\alpha_1\alpha_2} = \frac{1}{2}\epsilon_{\alpha_1\alpha_2}\epsilon^{\alpha_1\alpha_2} = 1, \tag{E.7}$$

where we have used $N^\mu \epsilon_{\mu\nu\rho} f^\nu_\alpha f^\rho_\beta = \epsilon_{\alpha\beta}$. Further, for a surface which does not vary in time, $n_\mu N^\mu = 0$ [43]. Now, we need to define the derivative on $\partial\mathcal{M}$, but this will need a connection. The one we use comes from requiring that the vielbein is covariantly conserved,

$$D_\mu \beta^\nu_a = \partial_\mu \beta^\nu_a + \Gamma^\nu{}_{\rho\mu}\beta^\rho_a - \beta^\nu_b \omega^b{}_{a\mu} = 0, \tag{E.8}$$

which has the solution

$$\Gamma^\mu{}_{\nu\rho} = \beta^\mu_a \partial_\rho (\beta^{-1})^a_\nu + \beta^\mu_a \omega^a{}_{b\rho}(\beta^{-1})^b_\nu. \tag{E.9}$$

Using the connection, we define the covariant derivative on $\partial\mathcal{M}$, $\mathring{D}_\alpha$. As an example, consider its action on a mixed field U^μ_β,

$$\mathring{D}_\alpha U^\mu_\beta = \partial_\alpha U^\mu_\beta + \Gamma^\mu{}_{\nu\alpha} U^\nu_\beta - \mathring{\Gamma}^\gamma{}_{\beta\alpha} U^\mu_\gamma, \tag{E.10}$$

where

$$\Gamma^\mu{}_{\nu\alpha} = \Gamma^\mu{}_{\nu\rho} f^\rho_\alpha, \tag{E.11}$$

$$\mathring{\Gamma}^\alpha{}_{\beta\gamma} = f^\alpha_\mu \partial_\gamma f^\mu_\beta + f^\alpha_\mu \Gamma^\mu{}_{\nu\gamma} f^\nu_\beta. \tag{E.12}$$

The covariant derivative can be used to define the second fundamental form [50],

$$\mathrm{I\!I}^\mu{}_{\alpha\beta} = \mathring{D}_\beta f^\mu_\alpha = \partial_\alpha f^\mu_\beta + \Gamma^\mu{}_{\nu\rho} f^\rho_\alpha f^\nu_\beta - \left(f^\gamma_\sigma \partial_\alpha f^\sigma_\beta + f^\gamma_\sigma \Gamma^\sigma{}_{\nu\rho} f^\rho_\alpha f^\nu_\beta\right) f^\mu_\gamma, \tag{E.13}$$

$$= \partial_\alpha f^\mu_\beta + \Gamma^\mu_{\nu\rho} f^\rho_\alpha f^\nu_\beta - f^\gamma_\sigma \partial_\alpha f^\sigma_\beta f^\mu_\gamma - f^\gamma_\sigma \Gamma^\sigma{}_{\nu\rho} f^\rho_\alpha f^\nu_\beta f^\mu_\gamma, \tag{E.14}$$

$$f^\delta_\mu \mathrm{I\!I}^\mu{}_{\alpha\beta} = f^\delta_\mu \partial_\alpha f^\mu_\beta + f^\delta_\mu \Gamma^\mu{}_{\nu\rho} f^\rho_\alpha f^\nu_\beta - f^\delta_\mu f^\gamma_\sigma \partial_\alpha f^\sigma_\beta f^\mu_\gamma - f^\delta_\mu f^\gamma_\sigma \Gamma^\sigma{}_{\nu\rho} f^\rho_\alpha f^\nu_\beta f^\mu_\gamma = 0, \tag{E.15}$$

where the last equality comes from $f^\delta_\mu f^\mu_\alpha = \delta^\delta_\alpha$. This makes the first and third terms cancel, as well as the second and fourth terms. As $f^\mu_\alpha N_\mu = 0$ by construction, we reach the conclusion that we can write the second fundamental form as

$$\mathrm{I\!I}^\mu{}_{\alpha\beta} = N^\mu \overline{K}_{\alpha\beta}, \tag{E.16}$$

for some tensor $\overline{K}_{\alpha\beta}$. We can easily find an explicit expression for $\overline{K}_{\alpha\beta}$ by multiplying both sides by N_μ (recall $N_\mu N^\mu = 1$),

$$\overline{K}_{\alpha\beta} = N_\mu \mathrm{II}^\mu{}_{\alpha\beta} = N_\mu \mathring{D}_\beta f^\mu_\alpha. \tag{E.17}$$

Consider

$$\mathring{D}_\beta (N_\mu f^\mu_\alpha) = (\mathring{D}_\beta N_\mu) f^\mu_\alpha + N_\mu \mathring{D}_\beta f^\mu_\alpha = 0, \tag{E.18}$$

where we have used the Leibniz rule explained in [43], and that $N_\mu f^\mu_\alpha = 0$. From this, we conclude that $\overline{K}_{\alpha\beta} = -f^\mu_a \mathring{D}_\beta N_\mu$. Next, we define the extrinsic curvature as

$$K_\alpha = \epsilon^{\beta\gamma} n_\beta \overline{K}_{\gamma\alpha} = -\epsilon^{\beta\gamma} n_\beta f^\mu_\gamma \mathring{D}_\alpha N_\mu, \tag{E.19}$$

$$= -t^\gamma f^\mu_\gamma \mathring{D}_\alpha N_\mu = N_\mu \mathring{D}_\alpha t^\mu, \tag{E.20}$$

where we have used the fact that $\epsilon^{\alpha\beta} n_\beta = t^\alpha$ is a spatial tangent vector, $t^\mu = f^\mu_\gamma t^\gamma$, and have flipped the sign and the action of the derivative in the same way as before.