REPRESENTATION THEORY AND ITS APPLICATIONS IN PHYSICS

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mathematics

by

Max Varverakis

June 2024

COMMITTEE MEMBERSHIP

TITLE:	Representation Theory and its Applications in Physics
AUTHOR:	Max Varverakis
DATE SUBMITTED:	June 2024
COMMITTEE CHAIR:	Sean Gasiorek Assistant Professor of Mathematics
COMMITTEE MEMBER:	Anton Kaul Professor of Mathematics
COMMITTEE MEMBER:	Eric Brussel Professor of Mathematics

ABSTRACT

Representation Theory and its Applications in Physics

Max Varverakis

Representation theory, which encodes the elements of a group as linear operators on a vector space, has far-reaching implications in physics. Fundamental results in quantum physics emerge directly from the representations describing physical symmetries. We first examine the connections between specific representations and the principles of quantum mechanics. Then, we shift our focus to the braid group, which describes the algebraic structure of braids. We apply representations of the braid group to physical systems in order to investigate quasiparticles known as anyons. Finally, we obtain governing equations of anyonic systems to highlight the differences between braiding statistics and conventional Bose-Einstein/Fermi-Dirac statistics.

ACKNOWLEDGMENTS

I am profoundly grateful to Dr. Sean Gasiorek for his exceptional support in helping me discover and explore a thesis topic that truly resonates with my interests. His willingness to invest time and effort into this endeavor has been invaluable.

I could not have undertaken this journey without the ongoing love and support from my grandparents, Diane and Larry Blau, as well as my parents, Cheryl Blau and Michael Varverakis. Finally, I extend my heartfelt gratitude to my partner, Alexandra Schwartz, for her unwavering support and encouragement throughout my academic pursuits and tribulations.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

The intimate connection between abstract mathematics and the physical world highlights the beautiful complexity of nature. In our efforts to understand the fundamental processes that govern the universe, we grow increasingly reliant on the principles of mathematics. One such tool that bridges the abstract and physical regimes is known as representation theory.

In this thesis, we examine representation theory and observe the physical consequences that emerge from the mathematics. First, we begin in Chapter 2 with a brief overview of representation theory. Then in Chapter 3, we elucidate specific applications of representation theory in the context of quantum physics, arriving at remarkably fundamental results. In Chapter 4, we introduce the braid group and explore some of its representations. Finally, we highlight physical applications of the braid group and its representations in Chapter 5.

It is assumed that the reader has a basic understanding of group theory and linear algebra. A familiarity with topology is also required, and the relevant definitions can be found in Appendix A. Notational conventions are chosen to align with the physics literature which at times differs from the standard mathematical notation. Discussions in the later chapters assume a general understanding of classical and quantum mechanics. To those less familiar with those topics and the notation, a supplementary overview is given in Appendix B.

Chapter 2

AN INTRODUCTION TO REPRESENTATION THEORY

Representation theory is a powerful tool that enables the algebraic properties of groups to be applied in contexts beyond group theory. In this chapter, we lay the foundation of representation theory for use in physical scenarios in later chapters. This is by no means a comprehensive treatment of the subject. There are a multitude of resources that cover representation theory in a variety of styles. One such resource that is particularly helpful in the context of mathematical physics is [26]. We start with the core definition.

Definition 2.1 (Representation of a group)**.** Let G be a group. A *representation* of G is a homomorphism from G to a group of operators on a linear vector space V. The dimension of V is the *dimension* or *degree* of the representation.

If X is a representation of G on V, then X is a map

$$g \in G \xrightarrow{X} X(g) \tag{2.1}$$

in which $X(g)$ is an operator on the vector space V. For a set of basis vectors $\{\mathbf{e}_i, i = 1, 2, \ldots, n\}$, we can realize each operator $X(g)$ as an $n \times n$ matrix:

$$X(g)\mathbf{e}_i = \sum_{j=1}^{n} X(g)^j{}_i \mathbf{e}_j \tag{2.2}$$

where the first index j is the row index and the second index i is the column index. Operator multiplication is defined as

$$X(g_1)X(g_2) = X(g_1g_2), \tag{2.3}$$

which satisfies the group multiplication rules. Since a representation is a group homomorphism, it follows that the operators $X(g)$ are invertible, and the inverse of $X(g)$ is $X(g^{-1})$.

Definition 2.2. If the homomorphism defining the representation is an isomorphism, then the representation is *faithful*. Otherwise, it is *degenerate*.

Example 2.1. The simplest representation of any group G is the *trivial* representation, in which every $g \in G$ is realized by $g \mapsto 1$. This representation is clearly degenerate.

Remark 2.1. *A representation must map the group identity element to the identity operator.*

Proof. Let X be a representation of a group G on a vector space V. Since X is a homomorphism, it must map the group identity element $e \in G$ to the identity operator I on V. Explicitly, $X(g) = X(eg) = X(e)X(g)$ for all $g \in G$ implies that $X(e)$ is the identity operator in V. □

Example 2.2. Consider the symmetric group S_n. The *defining* representation of S_n encodes each $\sigma \in S_n$ by placing a 1 in the j-th row and i-th column of the matrix $D(\sigma)$ if σ sends i to j, and 0 otherwise. For example, in S_3, the permutation (23) has

the matrix representation

$$D\big((23)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

whereas the permutation (123) is realized by the matrix

$$D\big((123)\big) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The above example involves a finite group. Infinite groups can also have representations, as demonstrated in the following example.

Example 2.3. Let G be the group of continuous rotations in the xy-plane about the origin. We can write $G = \{R(\phi), 0 \leq \phi < 2\pi\}$ with group operation $R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2)$. Consider the 2-dimensional Euclidean vector space V_2. Then we define a representation X of G on V_2 by the familiar rotation operation

$$\mathbf{e}_1' = X(\phi)\mathbf{e}_1 = \mathbf{e}_1 \cdot \cos\phi + \mathbf{e}_2 \cdot \sin\phi \tag{2.4}$$

$$\mathbf{e}_2' = X(\phi)\mathbf{e}_2 = -\mathbf{e}_1 \cdot \sin\phi + \mathbf{e}_2 \cdot \cos\phi, \tag{2.5}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthonormal basis vectors of V_2. This gives the familiar rotation matrix representation

$$R(\phi) \mapsto X(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}. \tag{2.6}$$

Definition 2.3 (Equivalence of Representations)**.** For a group G, two representations are *equivalent* if they are related by a similarity transformation. Equivalent representations form an equivalence class.

To determine whether two representations belong to the same equivalence class, we define the following.

Definition 2.4 (Character of a Representation)**.** The *character* $\chi(g)$ of an element $g \in G$ in a representation X is defined as the trace of the matrix representation: $\chi(g) = \mathrm{tr}\,(X(g))$.

Since trace is independent of basis, the character serves as a class label.

2.1 Irreducibility and Invariant Subspaces

The structure of a representation vector space of a group can offer insight into the representations themselves.

Definition 2.5 (Invariant Subspace)**.** Let $X(G)$ denote a representation of G on a vector space V, and let W be a subspace of V such that $X(g)\mathbf{x} \in W$ for all $\mathbf{x} \in W$ and $g \in G$. Then W is an *invariant subspace* of V with respect to $X(G)$. An invariant subspace is *minimal* or *proper* if it contains no non-trivial invariant subspaces with respect to $X(G)$.

The identification of invariant subspaces on a representation space leads to the following distinction.

Definition 2.6 (Irreducible Representation)**.** A representation $X(G)$ on V is *irreducible* if there is no non-trivial invariant subspace in V with respect to $X(G)$.

Otherwise, it is *reducible*. If $X(G)$ is reducible and its orthogonal complement to the invariant subspace is also invariant with respect to $X(G)$, then the representation is *fully reducible*.

A reducible representation can be decomposed into a direct sum of irreducible representations. This decomposition is unique up to similarity.

Example 2.4. Under the group of 2-dimensional rotations, consider the 1-dimensional subspace spanned by $\mathbf{e}_1$. This subspace is not invariant under 2-dimensional rotations because a rotation of $\mathbf{e}_1$ by $\pi/2$ results in the vector $\mathbf{e}_2$ that is clearly not in the subspace spanned by $\mathbf{e}_1$. A similar argument shows that the subspace spanned by $\mathbf{e}_2$ is not invariant under 2-dimensional rotations.

However, if there are eigenvectors of the rotation operator, we then have invariant subspaces formed by the span of the eigenvectors. Consider the linear combination of basis vectors

$$\mathbf{e}_\pm = \frac{1}{\sqrt{2}}\left(\mp\mathbf{e}_1 + i\mathbf{e}_2\right),$$

where $i = \sqrt{-1}$. Then a rotation by angle ϕ, denoted in operator form as $X(\phi)$, acts on $\mathbf{e}_\pm$ by

$$\begin{aligned}
X(\phi)\mathbf{e}_+ &= X(\phi)\frac{1}{\sqrt{2}}(-\mathbf{e}_1 + i\mathbf{e}_2) \\
&= \frac{1}{\sqrt{2}}(-X(\phi)\mathbf{e}_1 + iX(\phi)\mathbf{e}_2) \\
&= \frac{1}{\sqrt{2}}\left(-\mathbf{e}_1\cos\phi - \mathbf{e}_2\sin\phi - i\mathbf{e}_1\sin\phi + i\mathbf{e}_2\cos\phi\right) \\
&= \frac{1}{\sqrt{2}}\left(-\mathbf{e}_1(\cos\phi + i\sin\phi) + i\mathbf{e}_2(\cos\phi + i\sin\phi)\right) \\
&= e^{i\phi}\mathbf{e}_+,
\end{aligned}$$

and similarly $X(\phi)\mathbf{e}_- = e^{-i\phi}\mathbf{e}_-$.

Therefore, the span of each eigenvector of $X(\phi)$ is an invariant subspace of V_2 with respect to X. Since X is a 2-dimensional representation, we can decompose it into the $\mathbf{e}_\pm$ basis to obtain a direct sum of the two 1-dimensional irreducible representations:

$$X(\phi) = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix}.$$

Irreducible decompositions of a representation often simplify the application. In the following results, we assume that the representation space is finite-dimensional so that the operators can be represented by matrices. The following lemmas are needed to arrive at a powerful result regarding the dimension of irreducible representations that will be exploited in later chapters.

Lemma 2.2. *Let X and Y be irreducible representations of some group G such that $TX(g) = Y(g)T$ for all $g \in G$ and some fixed operator T. Then T is either zero or invertible.*

Proof. Let V be the representation vector space with which X,Y, and T operate. Consider $\mathbf{v} \in \ker(T)$. Then $TX(g)\mathbf{v} = Y(g)T\mathbf{v} = Y(g)\mathbf{0} = \mathbf{0}$ for all $g \in G$. Thus, $\ker(T)$ is invariant under $X(G)$. Since X is irreducible, there cannot exist non-trivial invariant subspaces, so either $\ker(T) = \{\mathbf{0}\}$ or $\ker(T) = V$.

Similarly, the image of T is invariant under $Y(G)$ since $Y(g)T\mathbf{v} = TX(g)\mathbf{v} \in \mathrm{im}(T)$ for all $g \in G$ and $\mathbf{v} \in V$. The irreducibility of Y implies that the Y-invariant subspace $\mathrm{im}(T) \subseteq V$ is either $\{\mathbf{0}\}$ or V.

By the rank-nullity theorem, T is either zero or square and invertible. □

Lemma 2.3 (Schur's Lemma)**.** *Let X be an irreducible representation of a group G defined on a vector space V. Consider some operator T on V that commutes with all operators $X(g)$ for $g \in G$. Then T is a scalar multiple of the identity matrix, I.*

Proof. Let λ be an eigenvalue of T. Then $T - \lambda I$ is not invertible. Since T commutes with $X(G)$, it follows that

$$(T - \lambda I)X(g) = X(g)(T - \lambda I),$$

for all $g \in G$. By Lemma 2.2, $T - \lambda I = 0$ since $X(G)$ is irreducible. Thus, $T = \lambda I$. □

Theorem 2.4. *Let G be a finite abelian group. Then the irreducible representations of G are one-dimensional.*

Proof. Suppose $X(G)$ is an irreducible representation of the abelian group G. Fix $h \in G$. Since X is a group homomorphism, we have $X(h)X(g) = X(hg) = X(gh) = X(g)X(h)$ for all $g \in G$. By Lemma 2.3, $X(h) = \lambda_h I$ for some scalar λ_h. The choice of $h \in G$ was arbitrary, so $X(g) = \lambda_g I$ for all $g \in G$. Thus, the irreducible representations of G are one-dimensional. □

The basic content covered in this chapter is sufficient to proceed to studying specific groups and their representations in the context of physics. As it turns out, there is a plethora of useful information encoded in the groups of rotations and translations that can be extracted to gain insight into physical systems. This is the focus of the following chapter.

Chapter 3

EXAMPLES IN PHYSICS

The goal of this chapter is to highlight some key results in physics through the lens of representation theory. The intimate connection between quantum mechanics and the representations discussed below offers a unique perspective on the emergence of the fundamental properties of quantum physics. To best illustrate the application of representation theory in quantum mechanics, notation is chosen to agree with the physics conventions. For those not as familiar with Dirac notation and the related quantum mechanical concepts, please refer to Appendix B.

3.1 Rotations in a plane and the group SO(2)

3.1.1 The rotation group

Consider the rotations of a 2-dimensional Euclidean vector space about the origin. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be orthonormal basis vectors of this space. Using geometry, we can determine how a rotation by some angle ϕ, written in operator form as $R(\phi)$, acts on the basis vectors:

$$R(\phi)\mathbf{e}_1 = \mathbf{e}_1 \cos\phi + \mathbf{e}_2 \sin\phi \tag{3.1}$$

$$R(\phi)\mathbf{e}_2 = -\mathbf{e}_1 \sin\phi + \mathbf{e}_2 \cos\phi. \tag{3.2}$$

In matrix form, we can write

$$R(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \tag{3.3}$$

which allows us to write Eqns. 3.1 and 3.2 in a condensed form

$$R(\phi)\mathbf{e}_i = R(\phi)^j{}_i \mathbf{e}_j, \tag{3.4}$$

with an implicit sum over the repeated index $j = 1, 2$.

Let $\mathbf{x}$ be an arbitrary vector in the plane. Then $\mathbf{x}$ has components x_i in the basis $\{\mathbf{e}_i\}$, where $i = 1, 2$. Equivalently, we can write $\mathbf{x} = \mathbf{e}_i x_i$, where again we implicitly sum over repeated indices. Then $\mathbf{x}$ transforms under rotations in accordance to the basis vectors

$$\begin{aligned} R(\phi)\mathbf{x} &= R(\phi)\mathbf{e}_i x_i \\ &= R(\phi)^j{}_i x_i \mathbf{e}_j \\ &= \left(R(\phi)^j{}_1 x_1 + R(\phi)^j{}_2 x_2 \right) \mathbf{e}_j \\ &= (x_1 \cos\phi - x_2 \sin\phi)\, \mathbf{e}_1 + (x_1 \sin\phi + x_2 \cos\phi)\, \mathbf{e}_2. \end{aligned} \tag{3.5}$$

Notice that $R(\phi)R^\top(\phi) = I$ where I is the identity matrix. This is precisely what defines *orthogonal matrices*. For 2-dimensional vectors in the plane, it is clear that these rotations do not change the length of said vectors. This can be verified by using

Eqn. 3.5:

$$
\begin{aligned}
|R(\phi)\mathbf{x}|^2 &= |R(\phi)^j{}_i x_i \mathbf{e}_j|^2 \\
&= \left|\left(x_1 \cos\phi - x_2 \sin\phi\right)\mathbf{e}_1 + \left(x_1 \sin\phi + x_2 \cos\phi\right)\mathbf{e}_2\right|^2 \\
&= \left(x_1 \cos\phi - x_2 \sin\phi\right)^2 + \left(x_1 \sin\phi + x_2 \cos\phi\right)^2 \\
&= \left(\cos^2\phi + \sin^2\phi\right) x_1^2 + \left(\sin^2\phi + \cos^2\phi\right) x_2^2 \\
&= x_1^2 + x_2^2 = |\mathbf{x}|^2.
\end{aligned}
$$

Similarly, notice that for any continuous rotation by angle ϕ, $\det R(\phi) = \cos^2\phi + \sin^2\phi = 1$. In general, orthogonal matrices have determinant equal to ± 1. However, the result of the above determinant of $R(\phi)$ implies that all continuous rotations in the 2-dimensional plane have determinant equal to $+1$. These are the *special orthogonal matrices of rank 2*. This family of matrices is denoted SO(2). Furthermore, there is a one-to-one correspondence with SO(2) matrices and rotations in a plane.

We define the group of continuous rotations in a plane by letting $R(0) = I$ be the identity element corresponding to no rotation (i.e., a rotation by angle $\phi = 0$), and defining the inverse of a rotation as $R^{-1}(\phi) = R(-\phi) = R(2\pi - \phi)$. Lastly, we define group multiplication as $R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2)$ and note that $R(\phi) = R(\phi \pm 2\pi)$, which can be verified geometrically. Although SO(2) is technically a 2-dimensional representation of a more abstract rotation group, it is often referred to as the rotation group due to the nature of the construction. Thus, group elements of SO(2) can be labelled by the angle of rotation $\phi \in [0, 2\pi)$.

3.1.2 Infinitesimal rotations

Consider an infinitesimal rotation labelled by some infinitesimal angle $d\phi$. This is equivalent to the identity plus some small rotation, which can be written as

$$R(d\phi) = I - id\phi J \tag{3.6}$$

where the scalar quantity $-i$ is introduced for later convenience and J is some quantity independent of the rotation angle. If we consider the rotation $R(\phi + d\phi)$, then there are two equivalent ways to interpret this rotation. Namely,

$$R(\phi + d\phi) = R(\phi)R(d\phi) = R(\phi)(I - id\phi J) = R(\phi) - id\phi R(\phi)J, \tag{3.7}$$

$$R(\phi + d\phi) = R(\phi) + d\phi \frac{dR(\phi)}{d\phi}, \tag{3.8}$$

where the Eqn. 3.8 can be thought of as a Taylor expansion of $R(\phi + d\phi)$ about ϕ. Equating the two expressions for $R(\phi + d\phi)$ yields

$$dR(\phi) = -id\phi R(\phi)J. \tag{3.9}$$

Solving this differential equation (with boundary condition $R(0) = I$) provides us with an equation for any group element involving J:

$$R(\phi) = e^{-i\phi J}, \tag{3.10}$$

where J is called the *generator* of the group.

The explicit form of J is found as follows. To first order in $d\phi$, we have

$$R(d\phi) = \begin{bmatrix} 1 & -d\phi \\ d\phi & 1 \end{bmatrix}.$$

Comparing to Eqn. 3.6,

$$I - id\phi J = \begin{bmatrix} 1 & -d\phi \\ d\phi & 1 \end{bmatrix} \implies J = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

Notice that $J^2 = I$, which implies that even powers of J equal the identity matrix and odd powers of J equal J. Taylor expanding $e^{-iJ\phi}$ gives

$$\begin{aligned} R(\phi) = e^{-iJ\phi} &= I - iJ\phi - I\frac{\phi^2}{2!} + iJ\frac{\phi^3}{3!} + \cdots \\ &= I\left(\sum_{n=0}^{\infty}(-1)^n\frac{\phi^{2n}}{(2n)!}\right) - iJ\left(\sum_{n=0}^{\infty}(-1)^n\frac{\phi^{2n+1}}{(2n+1)!}\right) \\ &= I\cos\phi - iJ\sin\phi \\ &= \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}. \end{aligned}$$

Therefore, the generator J can be used to recover the rotation matrix for an arbitrary angle ϕ.

3.1.3 Irreducible representations of SO(2)

Equipped with the generator J, we can construct the irreducible representations of SO(2). First, consider a representation U of SO(2) defined on a finite dimensional vector space V. Then $U(\phi)$ is the corresponding representation of $R(\phi)$. The same

argument as in Section 3.1.2 can be applied to an infinitesimal rotation to give

$$U(\phi) = e^{-iJ\phi},$$

which is an operator on V (for convenience, the same symbol J is used to denote the generator of the representation).

Since U is a representation of rotations, it preserves the length of vectors for all angles ϕ. Thus, for all vectors $\mathbf{a}$ (alternatively expressed as $|a\rangle$) in V and $\phi \in [0, 2\pi)$, we have

$$\begin{aligned}
|\mathbf{a}|^2 = |U(\phi)\mathbf{a}|^2 &\iff \langle a|a\rangle = \langle U(\phi)a|U(\phi)a\rangle = \langle a|U(\phi)^\dagger U(\phi)|a\rangle \\
&\iff U(\phi)^\dagger U(\phi) = I \\
&\iff e^{iJ^\dagger\phi}e^{-iJ\phi} = e^{-i(J-J^\dagger)\phi} = 1 \\
&\iff J = J^\dagger,
\end{aligned}$$

where $\dagger$ is the conjugate transpose. Note, the above calculations involve *Dirac notation*, which is described in more detail in Appendix B.

Thus, not only must U be unitary, but the generator J must be Hermitian. As a result, we know that V has an orthonormal basis of eigenvectors of J [2]. Moreover, note that Hermitian operators have real eigenvalues. This fact becomes especially important in the physical interpretation of the representations of 3-dimensional rotations in Section 3.5

According to Theorem 2.4, the abelian nature of SO(2) implies that all of its irreducible representations are one-dimensional. Then for any $|\alpha\rangle \in V$, the minimal

subspace containing $|\alpha\rangle$ that is invariant under SO(2) is one-dimensional. Hence,

$$J|\alpha\rangle = \alpha|\alpha\rangle,$$
$$U(\phi)|\alpha\rangle = e^{-iJ\phi}|\alpha\rangle = e^{-i\alpha\phi}|\alpha\rangle,$$

where the (real) number α is used as a label for the eigenvector of J with eigenvalue α. The periodicity conditions of SO(2) imply that $|\alpha\rangle = U(2\pi)|\alpha\rangle$, or equivalently, $e^{-i2\pi\alpha} = 1$. This implies that α must be an integer, as $e^{-i2\pi m} = 1$ for $m \in \mathbb{Z}$. Then U has a corresponding 1-dimensional representation for an integer m, defined by

$$J|m\rangle = m|m\rangle,$$
$$U^m(\phi)|m\rangle = e^{-im\phi}|m\rangle.$$

Though already true by Theorem 2.4, these representations are clearly irreducible, as there is no way to reduce the dimension of a 1-dimensional representation.

In general, the *single-valued irreducible representations of SO(2)* are defined as

$$U^m(\phi) = e^{-im\phi}, \tag{3.11}$$

for $m \in \mathbb{Z}$.

If $m = 0$, then $R(\phi) \mapsto U^0(\phi) = 1$, which corresponds to the trivial representation. If instead $m = 1$, then $R(\phi) \mapsto U^1(\phi) = e^{-i\phi}$, which maps rotations in SO(2) to distinct points on the unit circle in the complex plane. The $m = 1$ representation is faithful because each rotation by ϕ has a unique image under $U^1(\phi)$, which is clear when interpreting rotations of unit vectors geometrically. As ϕ ranges from 0 to 2π, U^1 traces over the unit circle in $\mathbb{C}$ in the clockwise direction. Similarly, U^{-1} traces over the unit circle in the counterclockwise direction because $U^{-1}(\phi) = e^{i\phi}$. The $m = -1$

case is therefore faithful as well. In general, U^n covers the unit circle $|n|$ times as ϕ ranges from 0 to 2π, and is not faithful for $n \neq \pm 1$.

The irreducible representations of SO(2) are orthonormal in the sense that

$$\frac{1}{2\pi}\int_0^{2\pi} \big(U^m(\phi)\big)^\dagger U^n(\phi)\, d\phi = \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)\phi}\, d\phi = \delta_{nm}, \tag{3.12}$$

where δ_{nm} is the Kronecker delta. Moreover, the irreducible representations satisfy the completeness relation:

$$\begin{aligned}\sum_{m\in\mathbb{Z}} U^m(\phi)(U^m(\phi'))^\dagger &= \sum_{m=-\infty}^{\infty} e^{-im(\phi-\phi')} \\ &= 1 + 2\sum_{m=1}^{\infty} \cos(m(\phi-\phi')) \\ &= 2\pi\delta(\phi-\phi'),\end{aligned} \tag{3.13}$$

where the last equality comes from the fact that the Fourier series of the Dirac delta is a sum over all frequencies of cosines (since it is an even distribution). The completeness relation implies that any appropriately defined function can be expressed in terms of the irreducible representations of SO(2), by integrating over ϕ. This is analogous to a Fourier series expansion. These relations will be exploited in Section 3.1.5.

3.1.4 Multivalued representations

If we relax the periodic condition on U to $U(2n\pi) = I$ for some $n \in \mathbb{Z}$, then the resulting 1-dimensional irreducible representations of SO(2) become multivalued. Consider the same construction of U^m in Section 3.1.3, but now with $m \in \mathbb{Q}$. For $m = \frac{1}{2}$, we

have

$$U^{1/2}(2\pi + \phi) = e^{-i\pi - i\frac{\phi}{2}} = -e^{-i\frac{\phi}{2}} = -U^{1/2}(\phi).$$

Hence, the rotation $R(\phi)$ is assigned to both $\pm e^{-i\frac{\phi}{2}}$ in the $U^{1/2}$ representation. For this reason, it can be said that $U^{1/2}$ is a *two-valued* representation of SO(2).

Despite this ambiguity in the realization of rotations in SO(2), the modified periodicity condition is still satisfied, as $U^{1/2}(4\pi) = e^{-i2\pi} = 1$. In other words, the double-valued representation of SO(2) traverses the unit circle twice before returning to the identity. In general, $U^{n/m}$ is an m-valued representation of SO(2) for $\frac{n}{m} \in \mathbb{Q}$ and $\gcd(n, m) = 1$.

The physical importance of these irreducible representations will be discussed when generalizing to rotations in 3-dimensional space in Section 3.5.

3.1.5 State vector decomposition

The concept of J generating 2-dimensional rotations is summarized in the following example. Consider a particle in a plane with polar coordinates (r, ϕ). The state vector of this particle is $|\phi\rangle$, where the coordinate r is suppressed in the vector notation, as the action of SO(2) preserves vector lengths. Note that the state vector $|\phi\rangle$ belongs to some Hilbert space V that is not necessarily the same as the physical space of the particle.

In this angle-basis, we have $U(\theta)\,|\phi\rangle = |\theta + \phi\rangle$. Then $|\phi\rangle$ can be decomposed as

$$|\phi\rangle = U(\phi)\,|\mathcal{O}\rangle = e^{-iJ\phi}\,|\mathcal{O}\rangle\,,$$

where $|\mathcal{O}\rangle$ is a "standard" state vector aligned with a pre-selected axis. The triviality of this result must not be overlooked, for it is important to note that any arbitrary state vector $|\phi'\rangle$ can be decomposed into $e^{-iJ\phi'}$ acting on $|\mathcal{O}\rangle$ [26]. This notion generalizes beyond the 2-dimensional case, and will be revisited for rotations in 3 spatial dimensions in Section 3.4.

The orthogonality and completeness relations given in Eqns. 3.12 and 3.13 imply that the state $|\phi\rangle$ can be decomposed into a linear combination of the eigenvectors of J:

$$|\phi\rangle = \left(\sum_m |m\rangle \langle m|\right) |\phi\rangle = \sum_m \langle m|\phi\rangle |m\rangle ,$$

where

$$\langle m|\phi\rangle = \langle m|U(\phi)|\mathcal{O}\rangle = \langle U^\dagger(\phi)m|\mathcal{O}\rangle = \langle e^{im\phi}m|\mathcal{O}\rangle = e^{-im\phi}\langle m|\mathcal{O}\rangle$$

is the projection of $|\phi\rangle$ onto the eigenvector $|m\rangle$ of J.

By construction, the eigenstates of J are invariant under rotations, so we are free to modify them up to a phase factor (i.e., pick different representatives from the eigenspaces). For example, we can choose the basis vector $|m\rangle$ to instead be $e^{i\phi_0 m}|m\rangle$ for some $\phi_0 \in [0, 2\pi)$. With this in mind, all eigenvectors $|m\rangle$ can be oriented along the direction of $|\mathcal{O}\rangle$ so that $\langle m|\mathcal{O}\rangle = 1$. Again, note that the inner product $\langle m|\mathcal{O}\rangle$ is a projection of the *state* $|m\rangle$ onto the *state* $|\mathcal{O}\rangle$, not to be confused with the projection of position vectors in the physical space of this system.

Thus, we can write the state vector $|\phi\rangle$ as

$$|\phi\rangle = \sum_m e^{-im\phi}|m\rangle . \tag{3.14}$$

As a result, the action of J on the state $|\phi\rangle$ can be written as

$$J\,|\phi\rangle = \sum_m e^{-im\phi} J\,|m\rangle = \sum_m m e^{-im\phi}\,|m\rangle = \sum_m i\frac{\partial}{\partial\phi}\left(e^{-im\phi}\,|m\rangle\right) = i\frac{\partial}{\partial\phi}\,|\phi\rangle\,.$$

For fixed m, multiplying Eqn. 3.14 by $\frac{1}{2\pi}e^{im\phi}$ and integrating over ϕ, we obtain

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} e^{im\phi}\,|\phi\rangle\;d\phi &= \sum_n \left(\frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)\phi}\,d\phi\right)|n\rangle \\ &= \sum_n \left(\frac{1}{2\pi}\int_0^{2\pi} (U^m(\phi))^\dagger U^n(\phi)\,d\phi\right)|n\rangle \\ &= \sum_n \delta_{nm}\,|n\rangle = |m\rangle\,.\end{aligned}$$

Then for an arbitrary state $|\psi\rangle \in V$, it follows that

$$\begin{aligned}|\psi\rangle = \sum_m \langle m|\psi\rangle\,|m\rangle &= \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_m e^{im\phi}\,\langle m|\psi\rangle\right)|\phi\rangle\;d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_m \langle\phi|m\rangle\,\langle m|\psi\rangle\right)|\phi\rangle\;d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\langle\phi|\underbrace{\left(\sum_m |m\rangle\,\langle m|\right)}_{\text{identity}}|\psi\rangle\,|\phi\rangle\;d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\langle\phi|\psi\rangle\,|\phi\rangle\;d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\psi(\phi)\,|\phi\rangle\;d\phi,\end{aligned}$$

where $\langle\phi|\psi\rangle$ is written as the *wavefunction* $\psi(\phi)$ since ϕ is a continuous parameter.

Then the action of J in the ϕ-basis generalizes to

$$\langle\phi|J|\psi\rangle = \langle J^\dagger\phi|\psi\rangle = -i\frac{\partial}{\partial\phi}\,\langle\phi|\psi\rangle = -i\frac{\partial}{\partial\phi}\psi(\phi).$$

If we let x and y be the Cartesian coordinates of the plane, then

$$\begin{aligned}\phi = \arctan\left(\frac{y}{x}\right) \implies \frac{\partial}{\partial\phi} &= \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} \\ &= \frac{\partial}{\partial\phi}\left(r\cos\phi\right)\frac{\partial}{\partial x} + \frac{\partial}{\partial\phi}\left(r\sin\phi\right)\frac{\partial}{\partial y} \\ &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \\ &= \left(\mathbf{r}\times\nabla\right)\cdot\mathbf{e}_z,\end{aligned}$$

where $\mathbf{r} = [x, y, z]^\top$ and $\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right]^\top$. Therefore, the general form of J is

$$J = -i\frac{\partial}{\partial\phi} = -i\left(\mathbf{r}\times\nabla\right)\cdot\mathbf{e}_z. \tag{3.15}$$

This result is of particular significance to quantum mechanics, as J has the same form as the orbital *angular momentum operator* $\hat{L}_z$ (normalized to $\hbar$ here), where z is the axis of rotation of the plane [12, 13].

3.2 Continuous 1-dimensional translations

Consider the group of continuous translations in one dimension, denoted by T_1, and let V be a 1-dimensional vector space with coordinate axis x. Then a vector $|x_0\rangle \in V$ is analogous to the point $x_0 \in \mathbb{R}$ on the real line. The translation of $|x_0\rangle$ by some amount x is described by the operator $T(x)$ in which

$$T(x)\,|x_0\rangle = |x + x_0\rangle\,.$$

The operator $T(x)$ has the expected group properties

$$T(0) = I, \tag{3.16}$$

$$T(x)^{-1} = T(-x), \tag{3.17}$$

$$T(x_1)T(x_2) = T(x_1 + x_2). \tag{3.18}$$

Consider an infinitesimal translation $T(dx)$. This derivation is identical to finding the generator J for SO(2) in Section 3.1.2. Thus, we rewrite

$$T(dx) = I - idxP,$$

where, for the moment, P is an arbitrary quantity. Eqns. 3.7 and 3.8 apply to $T(x)$ with P replacing J, T replacing R, and x replacing ϕ. This yields the familiar differential equation

$$\frac{dT(x)}{T(x)} = -iPdx, \tag{3.19}$$

along with the boundary condition Eqn. 3.16, which implies

$$T(x) = e^{-iPx}. \tag{3.20}$$

The exponential form of $T(x)$ satisfies the group properties of T_1 and is a valid formulation of the group. Therefore, P generates T_1. A similar decomposition of state vectors as in Section 3.1.5 can be done for T_1. Specifically, for $|x\rangle \in V$, we have

$$|x\rangle = T(x)\,|\mathcal{O}\rangle = e^{-iPx}\,|\mathcal{O}\rangle\,,$$

where $|\mathcal{O}\rangle$ is the standard state in V.

3.2.1 Irreducible representations of T_1

Consider a unitary representation U of T_1 on a finite dimensional vector space V. As before, U can be reduced to $U(x) = e^{-iPx}$, where P is the generator of the representation. The unitarity of U requires that P be Hermitian, as in the case of J for SO(2). It follows that the eigenvalues of P, labeled by p, are real. Since T_1 is abelian, Theorem 2.4 implies that the irreducible representations of T_1 are all 1-dimensional. Similar to Section 3.1.3, the irreducible representation $U^p(x)$ of $T(x)$ is given by

$$\begin{aligned} P\,|p\rangle &= p\,|p\rangle\,, \\ U^p(x)\,|p\rangle &= e^{-iPx}\,|p\rangle = e^{-ipx}\,|p\rangle\,. \end{aligned}$$

The above description satisfies Eqns. 3.16–3.18 with no further restrictions on p.

Notice that the eigenvalues of P are continuous, in contrast to the discrete eigenvalues of J for SO(2) which were a result of the periodicity condition. The orthonormality and completeness relations of the irreducible representations of T_1 are given by

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \big(U^p(x)\big)^\dagger U^{p'}(x)\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(p'-p)x}\,dx = \delta(p'-p), \tag{3.21}$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} U^p(x)\big(U^p(x')\big)^\dagger\,dp = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ip(x-x')}\,dp = \delta(x-x'), \tag{3.22}$$

where the normalization by 2π is chosen by convention.

3.2.2 Explicit form of P

Armed with Eqns. 3.21 and 3.22, we can perform the same arguments as in Section 3.1.5 for T_1 to expand a localized state $|x\rangle$ in terms of the eigenvectors of P:

$$|x\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \langle p|x\rangle\, |p\rangle\ dp = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ipx}\, |p\rangle\ dp,$$

where the sums from Section 3.1.5 are replaced by integrals due to the continuous and unbounded nature of p. Multiplying by e^{ipx} for some fixed p and integrating over x, we obtain an expression of $|p\rangle$ in terms of $|x\rangle$:

$$\begin{aligned}\int_{-\infty}^{\infty} e^{ipx}\, |x\rangle\ dx &= \int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(p'-p)x}\, |p'\rangle\ dx\right) dp' \\ &= \int_{-\infty}^{\infty} \delta(p'-p)\, |p'\rangle\ dp' = |p\rangle\,.\end{aligned}$$

The relationship between $|p\rangle$ and $|x\rangle$ is the familiar Fourier transform, where the state $|p\rangle$ is the momentum space representation of the state $|x\rangle$, which corresponds to position space.

The action of P on $|x\rangle$ can then be written as

$$P\,|x\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ipx} P\,|p\rangle\ dp = \frac{1}{2\pi}\int_{-\infty}^{\infty} pe^{-ipx}\, |p\rangle\ dp = i\frac{\partial}{\partial x}\,|x\rangle\,.$$

Therefore, an arbitrary state $|\psi\rangle$ can be expressed in either the position or momentum basis:

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)\, |x\rangle\ dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(p)\, |p\rangle\ dp,$$

where again $\psi(\cdot) = \langle\cdot|\psi\rangle$ is the wavefunction of the state $|\psi\rangle$ projected onto the relevant basis.

Lastly, we obtain the explicit form of P by viewing its action on $|\psi\rangle$ with respect to the position basis:

$$\langle x|P|\psi\rangle = \langle P^\dagger x|\psi\rangle = -i\frac{\partial}{\partial x}\langle x|\psi\rangle = -i\frac{\partial}{\partial x}\psi(x).$$

The above form of P agrees with the (normalized) quantum mechanical linear momentum operator $\hat{p}$ [12, 13].

3.2.3 Generalization to 3-dimensional space

The derivation in Section 3.2.2 generalizes to 3-dimensional space, where the group of 3-dimensional translations T_3 is defined by

$$\begin{aligned} T(\mathbf{r})\,|\mathbf{r}_0\rangle &= T(x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z)\,|x_0\mathbf{e}_x + y_0\mathbf{e}_y + z_0\mathbf{e}_z\rangle \\ &= |(x+x_0)\mathbf{e}_x + (y+y_0)\mathbf{e}_y + (z+z_0)\mathbf{e}_z\rangle \\ &= |\mathbf{r}+\mathbf{r}_0\rangle\,, \end{aligned}$$

subject to the equivalent group properties of T_1 in Eqns. 3.16–3.18 with $\mathbf{r}$ replacing x.

Notice that $T_3 \simeq T_1 \oplus T_1 \oplus T_1$, where the group operation is defined as $T(x_1, y_1, z_1)T(x_2, y_2, z_2) = T(x_1 + x_2, y_1 + y_2, z_1 + z_2)$. In other words, T_3 can be decomposed into independent 1-dimensional translations along each axis (or more generally along the span of each basis vector in 3-space). Thus, following the same procedure as in Section 3.1.2, an infinitesimal translation

$$T(d\mathbf{r}) = I - idxP_x\mathbf{e}_x - idyP_y\mathbf{e}_y - idzP_z\mathbf{e}_z$$

produces the following relations:

$$dT(x_j) = -idx_j T(x_j) P_j,$$

for $j = 1, 2, 3$ and $(x_1, x_2, x_3) = (x, y, z)$. This gives the expected result, namely

$$T(\mathbf{r}) = e^{-iP_x x} e^{-iP_y y} e^{-iP_z z} = e^{-i\mathbf{P}\cdot\mathbf{r}}.$$

Hence, the generator of 3-dimensional translations is the vector $\mathbf{P} = (P_x, P_y, P_z)$. The consequences of the separability of T_3 allows the results from Section 3.2.2 to be applied independently to each axis of translation. The intuitive generalization of T_1 to T_3 lets us immediately write down the explicit form of the generator for 3-dimensional translations. Since

$$P_j = -i\frac{\partial}{\partial x_j}, \tag{3.23}$$

we have

$$\mathbf{P} = -i\nabla. \tag{3.24}$$

Again, up to $\hbar$, Eqn. 3.24 is precisely the quantum mechanical linear momentum operator in 3 dimensions, often denoted $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$.

3.3 Symmetry, invariance, and conserved quantities

Physically, the generators $\mathbf{P}$ and J alter a (quantum) system by translation and rotation. These transformations correspond to the Hermitian operators $\hat{\mathbf{p}}$ and $\hat{L}_z$ that act on the state vectors belonging to the Hilbert space describing the system. Hence, the (real) eigenvalues of $\mathbf{P}$ and J (thus $\hat{\mathbf{p}}$ and $\hat{L}_z$) correspond to the physical observables (measurable quantities) of linear and angular momentum, respectively.

Armed with the explicit forms of these operators, the physical ramifications of the irreducible representations of T_3 and SO(2) can now be demonstrated.

According to Ehrenfest's theorem (see Appendix B.4), if a physical system represented by a Hamiltonian $\hat{H}$ is invariant under a time-independent transformation generated by an operator $\hat{A}$, then the physical observable corresponding to $\hat{A}$ is conserved. In other words, the expectation value of $\hat{A}$ is constant in time if the commutator $[\hat{H}, \hat{A}] = \hat{H}\hat{A} - \hat{A}\hat{H} = 0$.

The generators obtained in previous sections fit this framework. If a Hamiltonian $\hat{H}$ is invariant under translations or rotations, then $[\hat{H}, \mathbf{P}] = [\hat{H}, \hat{\mathbf{p}}] = 0$ or $[\hat{H}, J] = [\hat{H}, \hat{L}_z] = 0$, respectively. Therefore, the physical observables of linear and angular momentum are conserved in systems with translational and rotational symmetry. The following are examples of physical systems that exhibit these symmetries and the conserved quantities that result from them.

3.3.1 Conservation of linear momentum

Consider a free particle in three spatial dimensions. The Hamiltonian of this system is given by

$$\mathcal{H} = \frac{\mathbf{p}^2}{2m},$$

which gives the quantum operator

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m}.$$

Notice that

$$\left[\hat{\mathbf{p}}^2, \hat{\mathbf{p}}\right] = \left[(-i\hbar\nabla)^2, -i\hbar\nabla\right] = i\hbar^3\left[\nabla^2, \nabla\right] = i\hbar^3\left(\nabla^3 - \nabla^3\right) = 0,$$

where $\nabla^3 = \nabla \cdot \nabla \cdot \nabla$. It follows that

$$[\hat{H}, \hat{\mathbf{p}}] = \left[\frac{\hat{\mathbf{p}}^2}{2m}, \hat{\mathbf{p}}\right] = \frac{1}{2m}\left[\hat{\mathbf{p}}^2, \hat{\mathbf{p}}\right] = 0$$

Therefore, linear momentum is conserved in this system. This result is expected, as the Hamiltonian of a free particle is invariant under translations in space, which are generated by $\mathbf{P}$. The conservation of linear momentum is a direct consequence of the translational symmetry of the system.

3.3.2 Conservation of angular momentum

Now, consider the Hamiltonian describing a free particle in confined to a radially symmetric scalar potential $V(\mathbf{r})$. The quantum analog of the Hamiltonian is the operator

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}(\mathbf{r}),$$

where $\hat{V}(\mathbf{r})$ is the potential operator defined by $\hat{V}(\mathbf{r})\,|\mathbf{r}\rangle = V(\mathbf{r})\,|\mathbf{r}\rangle$.

Intuitively, a potential that depends solely on the radial coordinate should be invariant under rotations, as there is no angular dependence. According to Noether's theorem, the rotational symmetry of the system implies that angular momentum is conserved. This claim is equivalent to showing that $[\hat{H}, \hat{L}_z] = 0$, where $\hat{L}_z$ is the

operator corresponding to the generator of rotations in the xy-plane, derived as J in Section 3.1.

The angular momentum operator $\hat{L}_z$ is given by

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi},$$

where ϕ is the polar angle in the xy-plane. The following result is immediate:

$$[V(\mathbf{r}), \hat{L}_z] = 0,$$

since $V(\mathbf{r})$ does not have ϕ-dependence.

Recall that we can express $\hat{L}_z$ in Cartesian coordinates as

$$\hat{L}_z = -i\hbar \left(\hat{\mathbf{r}} \times \hat{\mathbf{p}}\right) \cdot \mathbf{e}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) = x\hat{p}_y - y\hat{p}_x, \tag{3.25}$$

where the *position operator* is defined as $\hat{\mathbf{r}} \left|\mathbf{r}\right\rangle = \mathbf{r} \left|\mathbf{r}\right\rangle$. To reduce clutter, the components $\hat{x}, \hat{y}, \hat{z}$ of $\hat{\mathbf{r}}$ are written without the hats.

First, we can reduce the commutator $[\hat{\mathbf{p}}^2, \hat{L}_z]$ to a simpler form:

$$\begin{aligned}
[\hat{\mathbf{p}}^2, \hat{L}_z] &= \left[\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2, x\hat{p}_y - y\hat{p}_x\right] \\
&= \left[\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2, x\hat{p}_y\right] + \left[\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2, -y\hat{p}_x\right] \\
&= \left[\hat{p}_x^2, x\hat{p}_y\right] + \left[\hat{p}_y^2, -y\hat{p}_x\right],
\end{aligned}$$

since the components of $\hat{\mathbf{p}}$ commute with each other. Further simplification is done using Eqns. B.1, B.2, B.4 and B.5:

$$\begin{aligned}
\left[\hat{p}_y^2, -y\hat{p}_x\right] &= \hat{p}_y[\hat{p}_y, -y\hat{p}_x] + [\hat{p}_y, -y\hat{p}_x]\hat{p}_y \\
&= \hat{p}_y\left(-y\cancelto{0}{[\hat{p}_y, \hat{p}_x]} - [\hat{p}_y, y]\hat{p}_x\right) + \left(-y\cancelto{0}{[\hat{p}_y, \hat{p}_x]} - [\hat{p}_y, y]\hat{p}_x\right)\hat{p}_y \\
&= \hat{p}_y\left(i\hbar - \cancel{y\hat{p}_y} + \cancel{y\hat{p}_y}\right)\hat{p}_x + \left(i\hbar - \cancel{y\hat{p}_y} + \cancel{y\hat{p}_y}\right)\hat{p}_x\hat{p}_y \\
&= 2i\hbar\hat{p}_y\hat{p}_x,
\end{aligned}$$

and by a relabeling of the variables, we also have

$$\left[\hat{p}_x^2, x\hat{p}_y\right] = -\left[\hat{p}_x^2, -x\hat{p}_y\right] = -2i\hbar\hat{p}_x\hat{p}_y = -\left[\hat{p}_y^2, -y\hat{p}_x\right].$$

Therefore, $[\hat{\mathbf{p}}^2, \hat{L}_z] = 0$. It follows from Eqn. B.3 that

$$[\hat{H}, \hat{L}_z] = \left[\frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}(\mathbf{r}), \hat{L}_z\right] = 0.$$

This result agrees with the previous claim that the emergence of conservation of angular momentum is due to the rotational symmetry of the system.

The profound connection between symmetries and conserved quantities is a fundamental principle in physics, and the results obtained in this section highlight the significance of representation theory in physics. The irreducible representations of T_3 and SO(2) provide the necessary mathematical framework to derive conservation laws without a preconceived notion of the physical universe.

3.4 3D rotations and the group SO(3)

As was done for translations in Section 3.2.3, we can generalize SO(2) to rotations in 3-dimensional space, albeit with less triviality. The group of rotations in Euclidean 3-space, which corresponds to 3-dimensional linear operators that fix the length of vectors, is the *special orthogonal group in 3D*, denoted by SO(3).

Consider a rotation in three dimensions about an axis (unit vector) $\mathbf{n}$ by an angle θ. The rotation $R_{\mathbf{n}}(\theta)$ is a linear transformation that maps a vector $\mathbf{v}$ to a new vector $\mathbf{v}'$ such that $|\mathbf{v}| = |\mathbf{v}'|$. The rotation angle $\theta \in [0, 2\pi)$ is a continuous parameter, and every one-parameter subgroup of SO(3) can be written as $\left\{ R_{\mathbf{n}}(\theta) \;\middle|\; \theta \in [0, 2\pi) \right\}$ for fixed $\mathbf{n}$.

The set of rotations in a plane perpendicular $\mathbf{n}$ is clearly isomorphic to SO(2). Thus, for a fixed axis of rotation $\mathbf{n}$, an infinitesimal rotation $R_{\mathbf{n}}(d\theta)$ can be used to obtain a generator of rotations about $\mathbf{n}$. The derivation is identical to that of J for SO(2) in Section 3.1.2. Hence, we can label the generator of rotations about $\mathbf{n}$ as $J_{\mathbf{n}}$, and the corresponding results from Section 3.1 can be applied to $J_{\mathbf{n}}$. Most notably, for arbitrary θ, we can write

$$R_{\mathbf{n}}(\theta) = e^{-i\theta J_{\mathbf{n}}}.$$

Consider the standard basis vectors $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ in 3-dimensional Euclidean space. The generators of rotations about the x, y, z axes are denoted by J_x, J_y, J_z, respectively. With some work [26], it can be shown that the generator $J_{\mathbf{n}}$ is decomposable into J_x, J_y, J_z for any $\mathbf{n}$ by projection onto the standard basis:

$$J_{\mathbf{n}} = n_x J_x + n_y J_y + n_z J_z, \tag{3.26}$$

where $n_\mu = \mathbf{n} \cdot \mathbf{e}_\mu$ for $\mu = x, y, z$. The general rotation operator about $\mathbf{n}$ becomes

$$R_{\mathbf{n}}(\theta) = e^{-i\theta(n_x J_x + n_y J_y + n_z J_z)}.$$

As in Section 3.1.3, the unitarity of the rotation operator requires that the generators J_x, J_y, J_z be Hermitian and therefore have real eigenvalues.

Therefore, the set $\{J_x, J_y, J_z\}$ forms a basis for the generators of the one-parameter abelian subgroups of SO(3). As a result, we can express SO(3) in terms of the Hermitian generator $\mathbf{J} = [J_x, J_y, J_z]^\top$. Namely, for an arbitrary rotation $R_{\mathbf{n}}(\theta)$, we can write

$$R_{\mathbf{n}}(\theta) = e^{-i\theta\, \mathbf{n}\cdot\mathbf{J}}.$$

3.4.1 Explicit form of J

Since the subspace generated by each component of $\mathbf{J}$ is isomorphic to SO(2), we can use the same arguments made in Section 3.1.5 to obtain the explicit forms of the generators J_x, J_y, J_z. For $\mu = x, y, z$, Eqn. 3.15 generalizes to rotations about $\mathbf{e}_\mu$ as follows:

$$J_\mu = -i\frac{\partial}{\partial \phi_\mu} = -i\,(\mathbf{r} \times \nabla) \cdot \mathbf{e}_\mu,$$

where ϕ_μ is the polar angle in the plane perpendicular to $\mathbf{e}_\mu$. This allows an explicit expression for $\mathbf{J}$:

$$\mathbf{J} = -i\,(\mathbf{r} \times \nabla)\,, \tag{3.27}$$

which, up to $\hbar$, is the quantum mechanical angular momentum operator in 3 dimensions [13], often written as $\hat{\mathbf{L}} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$.

3.4.2 Commutation relations of SO(3) generators

The algebraic structure of the generators of SO(3) is defined by the commutation relations of the basis generators J_x, J_y, J_z. By studying the underlying algebraic relationships between the generators, we can gain insight into the irreducible representations of SO(3) and the corresponding physical implications.

A consequence of the correspondence found in Section 3.4.1 is that the commutation relations of the generators J_x, J_y, J_z are identical to those of the angular momentum operators $\hat{L}_x, \hat{L}_y, \hat{L}_z$ up to $\hbar$. First, note that Eqn. 3.25 can be generalized to each component of $\mathbf{L}$:

$$\hat{L}_x = y\hat{p}_z - z\hat{p}_y, \tag{3.28}$$

$$\hat{L}_y = z\hat{p}_x - x\hat{p}_z, \tag{3.29}$$

$$\hat{L}_z = x\hat{p}_y - y\hat{p}_x. \tag{3.30}$$

Thus, we can write

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}. \tag{3.31}$$

The commutation relations of the angular momentum operators can then be found by direct computation:

$$\begin{aligned}
[\hat{L}_z, \hat{L}_x] &= [\hat{L}_z, y\hat{p}_z - z\hat{p}_y] \\
&= [\hat{L}_z, y\hat{p}_z] - [\hat{L}_z, z\hat{p}_y] \\
&= y[\hat{L}_z, \hat{p}_z] + [\hat{L}_z, y]\hat{p}_z - z[\hat{L}_z, \hat{p}_y] - [\hat{L}_z, z]\hat{p}_y \\
&= 0 - i\hbar x\hat{p}_z + i\hbar z\hat{p}_x + 0 \\
&= i\hbar(z\hat{p}_x - x\hat{p}_z) = i\hbar\hat{L}_y,
\end{aligned}$$

where the remaining details can be found in Appendix B.3. The appropriate permutation of the indices gives the other commutation relations. Altogether, the commutation relations of the angular momentum operators are

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \tag{3.32}$$

$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \tag{3.33}$$

$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \tag{3.34}$$

These commutation relations are identical to those of the generators J_x, J_y, J_z up to $\hbar$.

3.4.3 Irreducible representations of SO(3)

Due to the nontrivial interaction between J_x, J_y, J_z, the irreducible representations of SO(3) are not as straightforward to determine as those of SO(2). However, the commutation relations in Eqns. 3.32–3.34 provide the necessary foundation to proceed with the following analysis.

Let V be a finite dimensional vector space corresponding to a representation of SO(3). The generators J_x, J_y, J_z act on V as linear operators, and the J-analogue of Eqns. 3.32–3.34 must be satisfied. To obtain an irreducible representation of SO(3), we seek a subspace of V that is invariant under SO(3) rotations. Equivalently, a subspace of V that is invariant under the action of J_x, J_y, J_z will be invariant under SO(3).

The most straightforward procedure for constructing an invariant subspace of V is by choosing a "standard" vector that is an eigenvector of one of the generators, and then applying SO(3) operations to generate the rest of the basis [26]. As is customary in physics, we choose the z-axis as the standard axis of rotation.

Let $|m\rangle$ be a normalized eigenvector of J_z, in which $J_z |m\rangle = m |m\rangle$ for some real number m. For reasons presently unknown, define a new operator

$$J^2 = \mathbf{J} \cdot \mathbf{J} = J_x^2 + J_y^2 + J_z^2.$$

It follows that

$$\begin{aligned}
[J^2, J_z] &= [J_x^2 + J_y^2 + J_z^2, J_z] \\
&= [J_x^2, J_z] + [J_y^2, J_z] \\
&= J_x[J_x, J_z] + [J_x, J_z]J_x + J_y[J_y, J_z] + [J_y, J_z]J_y \\
&= -iJ_xJ_y - iJ_yJ_x + iJ_yJ_x + iJ_xJ_y = 0.
\end{aligned}$$

Appropriate permutations of variables yields identical results. Since J^2 commutes with J_x, J_y, J_z, it also commutes with their linear combinations. Consequently, J^2

commutes with all SO(3) rotations. Furthermore, we have

$$(J^2)^\dagger = \mathbf{J}^\dagger \cdot \mathbf{J}^\dagger = \mathbf{J} \cdot \mathbf{J} = J^2,$$

so J^2 is Hermitian.

Note that commuting Hermitian operators are simultaneously diagonalizable, so they have a common set of eigenvectors [28]. In this case, we choose the simultaneous eigenvectors of J^2 and J_z as the basis vectors of the invariant subspace of V. At this point, we have one eigenvector $|m\rangle$ of J_z, and it is not currently known to be an eigenvector of J^2. To obtain the other eigenvectors that span the invariant subspace, we first define two more operators. Let

$$J_\pm = J_x \pm iJ_y. \tag{3.35}$$

These operators have the following properties [12, 13, 26]:

$$\begin{aligned}
[J_z, J_\pm] &= [J_z, J_x] \pm i[J_z, J_y] = iJ_y \pm i(-iJ_x) = \pm(J_x \pm iJ_y) = \pm J_\pm, \\
[J_+, J_-] &= [J_x + iJ_y, J_x - iJ_y] = i[J_x, J_y] - i[J_y, J_x] = 2iJ_z, \\
J^2 &= J_x^2 + J_y^2 + J_z^2 \\
&= J_+J_- + i(J_xJ_y - J_yJ_x) + J_z^2 \\
&= J_+J_- + i(iJ_z) + J_z^2 \\
&= J_+J_- - J_z + J_z^2 \\
&= J_-J_+ + J_z + J_z^2 \\
J_\pm^\dagger &= J_\mp.
\end{aligned}$$

Notice that

$$\begin{aligned} J_z J_\pm |m\rangle &= [J_z, J_\pm] |m\rangle + J_\pm J_z |m\rangle \\ &= \pm J_\pm |m\rangle + m J_\pm |m\rangle \\ &= (m \pm 1) J_\pm |m\rangle . \end{aligned}$$

Therefore, $J_\pm |m\rangle$ are either eigenstates of J_z with eigenvalue $m \pm 1$ or zero. The name of $J_\pm$ as the *ladder operators* is justified by the fact that they raise or lower the eigenvalue of J_z by one unit as if climbing the rungs of a ladder. Assume that $J_+ |m\rangle \neq 0$. Then the eigenvector $J_+ |m\rangle$ can be normalized and written as $|m+1\rangle$. Similarly, $|m-1\rangle$ can be defined as $J_- |m\rangle$, assuming it is nonzero.

With the ladder operators, we can generate a set of eigenvectors of J_z by repeated application on the standard eigenvector $|m\rangle$. Since V is assumed to be finite, this process must terminate at some point. Let j denote the largest eigenvalue of J_z in the invariant subspace, and similarly let l denote the lowest. In other words, we have

$$J_+ |j\rangle = 0, \quad J_- |l\rangle = 0,$$

so that any further application of the corresponding ladder operator returns zero.

The span of eigenvectors $\{|l\rangle , |l+1\rangle , \ldots , |j-1\rangle , |j\rangle\}$ is indeed an invariant subspace of V under SO(3) rotations. Since $J_x = \frac{1}{2}(J_+ + J_-)$ and $J_y = \frac{1}{2i}(J_+ - J_-)$, it follows that their action on $\{|l\rangle , |l+1\rangle , \ldots , |j-1\rangle , |j\rangle\}$ is closed within the subspace.

Additionally, for the eigenvector $|j\rangle$, it follows that

$$\begin{aligned} J^2 |j\rangle &= (J_- J_+ + J_z + J_z^2) |j\rangle \\ &= (0 + j + j^2) |j\rangle \\ &= j(j+1) |j\rangle . \end{aligned}$$

Since J^2 commutes with all SO(3) rotations, for any $|m\rangle \in \{|l\rangle , |l+1\rangle, \ldots, |j-1\rangle , |j\rangle\}$ we have

$$J^2 |m\rangle = J^2 J_-^{(j-m)} |j\rangle = J_-^{(j-m)} J^2 |j\rangle = j(j+1) |m\rangle .$$

Therefore, every eigenvector of J_z is simultaneously an eigenvector of J^2 with degenerate eigenvalue $j(j+1)$. We gain further insight into this invariant subspace by noting that

$$\begin{aligned} 0 = \langle 0|0\rangle &= \langle l| J_-^\dagger J_- |l\rangle \\ &= \langle l| J_+ J_- |l\rangle \\ &= \langle l| J^2 - J_z^2 + J_z |l\rangle \\ &= (j(j+1) - l^2 + l) \cancelto{1}{\langle l|l\rangle} \\ &= j(j+1) - l(l-1), \end{aligned} \tag{3.36}$$

which implies $j(j+1) = l(l-1)$ or equivalently $j = -l$. Since j is the largest eigenvalue of J_z in the invariant subspace, we then have a $(2j+1)$-dimensional invariant subspace spanned by $\{|-j\rangle , |-j+1\rangle , \ldots , |j-1\rangle , |j\rangle\}$. Moreover, we can write

$$|-j\rangle = J_-^{(2j+1)} |j\rangle ,$$

which implies that $2j+1$ is a positive integer. Therefore,

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{3.37}$$

Therefore, each irreducible representation of SO(3) is characterized by the value of j, which is nonnegative and either an integer or a half-integer. The orthonormal basis of eigenvectors corresponding to the invariant space have two labels, one for the value of j which specifies the irreducible representation, and one for the value of $m \in \{-j, -j+1, \ldots, j-1, j\}$, which identifies the specific eigenvector. The results obtained above can be summarized by the following equations:

$$J^2 \left|j, m\right\rangle = j(j+1) \left|j, m\right\rangle, \tag{3.38}$$

$$J_z \left|j, m\right\rangle = m \left|j, m\right\rangle, \tag{3.39}$$

$$J_\pm \left|j, m\right\rangle = \sqrt{j(j+1) - m(m \pm 1)} \left|j, m \pm 1\right\rangle, \tag{3.40}$$

where the normalization constant of the ladder operators results from a calculation similar to Eqn. 3.36.

3.5 Physical implications of SO(3)

First and foremost, any physical system that is invariant under rotations in 3D space is subject to conservation of angular momentum, which follows the same arguments as in Section 3.3.2. Thus, the process of deriving the commutator of a Hamiltonian with the angular momentum operator generalizes to the 3D case and has the same physical consequences. The case of SO(2) versus SO(3) is similar enough that the results obtained in Section 3.3.2 can be extended to SO(3) with minimal effort. We therefore reserve the following discussion for new insights that arise from the 3D case.

The irreducible representations of SO(3) give rise to highly fundamental results in quantum mechanics. As discussed in Section 3.3, the eigenvalues of the components of $\mathbf{J}$ and thus $\hat{\mathbf{L}}$ correspond to physical observables. Recall that the components of $\hat{\mathbf{L}}$ act on the Hilbert space of a quantum system where each state vector describes a particular physical state of the system. With this in mind, the eigenvalues of $\hat{L}_x, \hat{L}_y, \hat{L}_z$ correspond to the physically measurable values of orbital angular momentum projected along the appropriate axis of rotation.

In quantum physics, there are multiple types of angular momentum. Up to this point, we have discussed the abstract generators of SO(3) ($\mathbf{J}$) and the corresponding orbital angular momentum operators ($\hat{\mathbf{L}}$). However, there is a second type of angular momentum that is intrinsic to particles, known as spin angular momentum, or plainly *spin*. Though not actually spinning, particles with intrinsic spin behave as if they are spinning about an axis resulting in nonzero angular momentum.

The standard generators of spin are denoted by $\hat{\mathbf{S}} = [\hat{S}_x, \hat{S}_y, \hat{S}_z]^\top$, which satisfy the same commutation relations as $\hat{\mathbf{L}}$ and $\mathbf{J}$ (Eqns. 3.32–3.34). The *total angular momentum* of a quantum system is then given by the sum of the orbital and spin angular momenta, $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$, where here $\hat{\mathbf{J}}$ is the total angular momentum operator.

A defining result of quantum theory is that continuous measurables in classical mechanics become discretized (or quantized) when applied to quantum systems. For example, in the classical regime, the angular momentum of an object is equal to the product of its moment of inertia and angular velocity. There are no physical restrictions of angular velocity, so the angular momentum can take on any real value. Of course, one can place an upper bound on angular momentum by means of relativity, but that is beyond the scope of this discussion and does not affect the comparisons made here.

In quantum physics, classical intuition breaks down. For example, the angular momentum of an electron, which is governed by quantum mechanics, can only be integer multiples of $\hbar/2$ rather than a continuum of values [12, 13]. The lack of physical intuition behind these properties is troubling for many. However, the representation theory of SO(3) provides an avenue to understand the emergence of discretization of observables in quantum mechanics.

3.5.1 Quantization of observables

Though not offering physical intuition, the irreducible representations of SO(3) provide a mathematical framework to describe the discretization of angular momentum in quantum mechanics.

Consider an arbitrary representation of SO(3). Recall that the eigenvectors $\left\{|j,m\rangle \;\middle|\; m = -j, -j+1, \ldots j+1, j\right\}$ of J_z and J^2 (derived in Section 3.4.3) form a basis for the invariant subspace of the representation space of SO(3). The corresponding irreducible representation of SO(3) is characterized by the value of j, which is nonnegative and either an integer or half-integer.

By construction, the eigenvalues of J_z are discrete. Consequently, the measurable values of orbital, spin, and total angular momentum are quantized in the same manner. The discretization of observables arises from the mathematics of the irreducible representations of SO(3) which does not require physical intuition to derive. This fact illustrates the power of representation theory in physics, as we have obtained one of the most fundamental and defining attributes of quantum mechanics without ever invoking physical principles.

The irreducible representations of SO(3) fall under two distinct classifications: integer values of j and half-integer values. If we restrict our attention to the spin states of

particles, the representations with integer values of j correspond to integer-spin particles, such as bosons or gravitons. Conversely, the representations with half-integer values of j correspond to half-integer-spin particles, such as fermions. The characteristics of integer-spin and half-integer-spin particles emerges from the underlying irreducible representation of SO(3).

The first case we investigate is spin-1 particles, which includes force carriers such as the photon for electromagnetism. The representation of SO(3) with $j = 1$ has eigenvectors $\{|1,1\rangle, |1,0\rangle, |1,-1\rangle\}$, which correspond to the possible spin states of a spin-1 particle. Thus, the measurable spin values of a spin-1 particle are $m = -1, 0, 1$ (normalized to $\hbar$), which comes directly from the eigenvalues of the generators of the irreducible representation of SO(3). One can use the ladder operators obtained in Section 3.4.3 to jump between the spin states of a spin-1 particle. As discussed in Section 3.1.3 for SO(2), the single-valued representations (integer-j representations) are faithful, and thus the global periodicity condition of SO(3) (rotations by 2π are equivalent to the identity) is satisfied. This periodicity condition is a defining property of integer-spin particles, and will be exploited in Chapter 5.

Identical to the conclusions of Section 3.1.4, the irreducible representations of SO(3) corresponding to half-integer values of j are double-valued. This fact has profound implications for the behavior of *spinors* under coordinate rotations. A spinor is a 2-component complex-valued vector that describes the spin state of a half-integer-spin particle, such as an electron [12]. A non-intuitive property of electrons, and thus spinors, is that a rotation by 2π results in a change of sign of the spinor. This antisymmetric behavior is a direct consequence of the double-valued representations of SO(3), as was seen in the $m = 1/2$ case of SO(2) in Section 3.1.4.

For the case of electrons, we can understand the possible spin states by investigating the eigenvectors corresponding to the $j = 1/2$ irreducible representation of SO(3).

According to Section 3.4.3, we have a 2-dimensional invariant space with eigenbasis $|\frac{1}{2}, \pm\frac{1}{2}\rangle$. In physics, these two states are often referred to as the spin-up ($m = +1/2$) and spin-down ($m = -1/2$) states. From these vectors, one can derive the matrix forms of the spin operators $\hat{S}_x, \hat{S}_y, \hat{S}_z$, which give the familiar Pauli-spin matrices [12]. These matrices are unitary, which agrees with the fact that physical transformations must preserve probabilities (Appendix B.1).

The main takeaway here is that one can derive extremely fundamental properties and results from quantum mechanics without any physical assumptions. The quantization of angular momentum in quantum mechanics is a direct consequence of the irreducible representations of SO(3), which is a powerful and elegant result that has far-reaching implications in physics.

3.5.2 Additional applications

We have merely scratched the surface of the physical ramifications of the irreducible representations of SO(3). With the foundation laid in this section, one can explore a variety of topics in quantum physics that are built upon the representation theory of SO(3). A brief list is provided below, and more detailed discussions can be found in references [12, 13, 26].

1. As with SO(2), if a system is invariant under 3D rotations (radially symmetric), then angular momentum is conserved. For each axis of rotation, one can directly calculate the commutator with the Hamiltonian as was done previously for SO(2) rotations. The same arguments in Section 3.3.2 can be applied to spin, orbital, and total angular momentum and will give the familiar results.

2. In a radially symmetric system, one has eigenvectors $|E, l, m\rangle$, which are simultaneous eigenvectors of $\hat{H}, J^2, J_z$. In position space, the eigenfunctions cor-

responding to these eigenvectors are separable into a radial component and a spherical harmonic. This is a classic result in undergraduate quantum physics.

3. The analysis of linear and angular momentum and the corresponding operators can be combined by finding simultaneous eigenvectors to construct a subspace invariant under both translations and rotations. The related group is known as the Euclidean group, which is the group of translations and rotations.

4. Multi-particle systems can be described by tensor products of the irreducible representations of SO(3). One such example is a 2-electron state, which results in the classification of singlet and triplet states. Taking this further, one arrives at the Clebsch-Gordan coefficients, which relate the individual angular momentum basis to the total angular momentum basis.

5. In the case of spin-1/2 particles, one can apply statistical mechanics to arrive at the renowned Pauli exclusion principle. Though not formally derived, the Pauli exclusion principle will emerge from the antisymmetric nature of fermions in Chapter 5.

Chapter 4

THE BRAID GROUP

The study of the braid group has far-reaching implications from knot theory to quantum computation [1, 4, 6, 7, 14, 15]. A basic understanding of topology is assumed in the following discussions. The necessary topological definitions are provided in Appendix A.

Definition 4.1. The *configuration space* of n ordered distinct points in the complex plane $\mathbb{C}$ is defined as $M_n = \{(z_1, \ldots, z_n) \in \mathbb{C}; z_i \neq z_j, \forall i \neq j\}$. Alternatively, consider $\mathcal{D}$ to be the collection of all hyperplanes $H_{i,j} = \{z_i = z_j\} \in \mathbb{C}^n$ for $1 \leq i < j \leq n$. Then we can define $M_n = \mathbb{C}^n \setminus \mathcal{D}$.

Note that $(z_1, z_2, z_3, \ldots, z_n)$ and $(z_2, z_1, z_3, \ldots, z_n)$ are different points in the configuration space M_n. Before studying the various interpretations of the braid group, we first define the braid group itself.

Definition 4.2. The *pure braid group* on n strands, denoted PB_n, is the fundamental group of M_n. One can write $PB_n = \pi_1(M_n)$.

4.1 Visualization of pure braids

We can think of a pure braid as a loop in M_n:

$$\begin{aligned} \beta : [0,1] &\to M_n \\ t &\mapsto \beta(t) = (\beta_1(t), \beta_2(t), \ldots, \beta_n(t)) \,, \end{aligned}$$

with some base point. Conventionally, we define the base point as the n-tuple of integers $(1, 2, 3, \ldots, n) \in \mathbb{C}^n$. Then a pure braid can be though of the motion of these points in the complex plane as t ranges from 0 to 1 in which $\beta_i(t)$ is defined and $\beta_i(t) \neq \beta_j(t)$ for every $t \in [0, 1]$ and $i \neq j \in \{1, 2, \ldots, n\}$. Because each β_i is a loop, it must start and end at the point i (e.g., $\beta_i(0) = \beta_i(1) = i$). Recall that the loops are actually equivalence classes of loops under homotopy. As a result, we can continuously deform the motion of the n points while maintaining the same pure braid (up to equivalence) so long as we preserve the pairwise distinction of the points for all time $t \in [0, 1]$.

A common visualization of pure braids is to plot the motion of the points in 3-dimensional space. For each $t \in [0, 1]$, we draw the points $(\beta_i(t), t)$ in $\mathbb{C} \times [0, 1]$ for every $i \in \{1, \ldots, n\}$. The space $\mathbb{C} \times [0, 1]$ can be thought of as a spacetime diagram, where the motion of the points is plotted in the complex plane at each time t, with the time being the vertical axis. The convention is to have $\mathbb{C} \times \{0\}$ placed above $\mathbb{C} \times \{1\}$, so that the motion of the points is plotted from top to bottom, as in Figure 4.1.

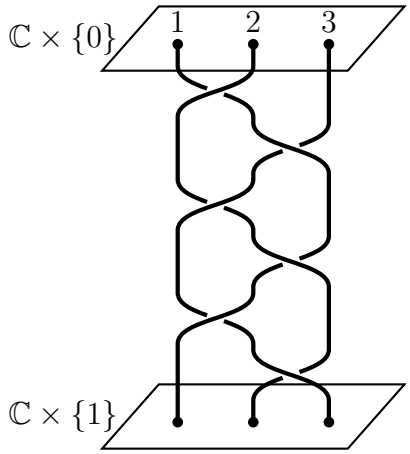

Figure 4.1: Pure braid. A pure braid on three strands is visualized as the trajectory of three particles as they move in $\mathbb{C}$, plotted over an arbitrary (normalized) time period. Each point ends at the same relative starting position in $\mathbb{C}$.

For every $i \in \{1, \ldots, n\}$, the motion of a single point starting at $(i, 0)$ and ending at $(i, 1)$ is known as the i-th *strand* of the pure braid. This can also be described by the i-th projection of the n-tuple $\beta(t)$. Thus, two braids are equivalent under homotopy if, for every moment of a continuous deformation of the n strands in $\mathbb{C} \times [0, 1]$, the (fixed) endpoints $((1, 0), (2, 0), \ldots, (n, 0))$ and $((1, 1), (2, 1), \ldots, (n, 1))$ are connected by strands that are pairwise disjoint where each strand intersects the plane $\mathbb{C} \times \{t\}$ exactly once for every $t \in [0, 1]$.

As pure braids are members of the pure braid group, multiplication is a well-defined operation. In the context of M_n, multiplication of pure braids involves the concatenation of loops. Visually, this is the process of stacking braids on top of each other, and then rescaling the time dimension so that t ranges from 0 to 1.

4.2 General braids

In the previous section, we defined pure braids in which the endpoints of each strand are identical at the beginning and end of the motion. This notion generalizes to define (non-pure) braids. First, we define a more general configuration space than M_n. The symmetric group S_n permutes the n distinct points in $\mathbb{C}$. Then the *configuration space of n unordered points in* $\mathbb{C}$ is the quotient space $N_n = M_n/S_n$.

Definition 4.3. The braid group on n strands is the fundamental group of N_n, denoted $B_n = \pi_1(N_n)$.

The visualization of a braid is the same as in the case of pure braids, only now the endpoints of each strand do not necessarily match the starting points, as illustrated Figure 4.2. For example, the i-th strand may start at the point $(i, 0)$ but end at the point $(j, 1)$ for $i, j \in \{1, \ldots, n\}$. The equivalence of strands is still defined as

before under the homotopy of loops. Loop concatenation defines the multiplication of braids, as before.

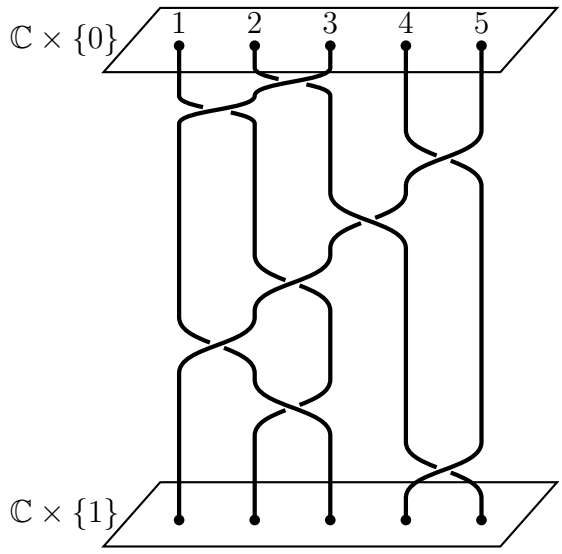

Figure 4.2: General braid. A general braid on five strands. The ending position of each particle is not necessarily in the same location of $\mathbb{C}$ as the starting position, which classifies this as a general braid.

Note that for $n \geq 1$ there is a natural inclusion $\iota : B_n \hookrightarrow B_{n+1}$ in which we add a new strand to any $\beta \in B_n$ that does not interact with the other strands.

4.3 Standard generators of the braid group

Originally proposed by Artin [1], each braid can be decomposed into a product of *standard generators* of the braid group. When visualizing braids in $\mathbb{R} \times [0, 1]$, a crossing of two strands is clearly indicated by one going over the other, as in Figure 4.3. Suppose each crossing occurs at a different time $t \in [0, 1]$. Then by rescaling the time component of an arbitrary braid, we can decompose it into a stack of simple braids with only one crossing between neighboring strands per braid. Each single crossing of strands can be obtained by performing a transposition between neighboring endpoints of the strands.

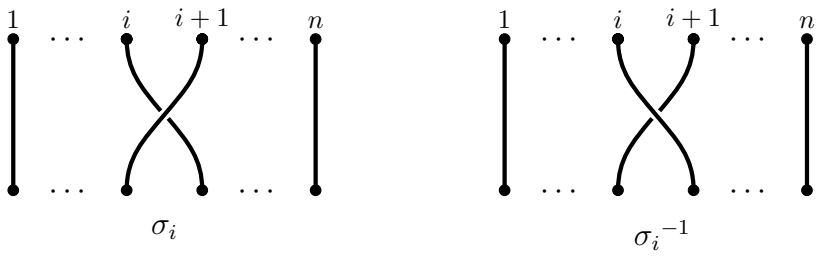

Figure 4.3: Artin generators. The action of σ_i on n strands. The i-th strand gets twisted behind the $(i+1)$-th strand when acted on by σ_i, and the opposite occurs for ${\sigma_i}^{-1}$. In that way, concatenating these two braids would result in no intertwining of strands, giving the identity braid.

For instance, swapping the endpoints of the i-th and $(i+1)$-th strands can be written as applying σ_i to the identity braid (i.e., the braid that starts without any crossings of strands). It must be noted that there are two distinct ways to swap the endpoints of two strands. From a top-down perspective looking at the plane $\mathbb{C} \times \{t\}$ for some time t, σ_i swaps (i, t) and $(i+1, t)$ in a clockwise rotation. The reverse of this operation (i.e., twisting the endpoints around in the counterclockwise direction) is denoted σ_i^{-1}. Both of these operations are illustrated in Figure 4.3. The standard generators of the braid group B_n are defined as the set $\{\sigma_1, \sigma_2, \ldots, \sigma_{n-1}\}$. An arbitrary braid can be constructed by concatenating (or stacking) the simple braids made from the standard generators before rescaling the time coordinate to $[0, 1]$.

A general braid can be written as a product of the standard generators and their inverses. One of the many metrics to characterize a braid involves the powers of the standard generators in its decomposition, as defined below.

Definition 4.4. The *degree* of a braid $\beta \in B_n$ is the sum of the powers of the standard generators in the decomposition of β.

4.4 Automorphisms of the free group

Consider the n-times punctured disk $\mathbb{D}_n$. The fundamental group of $\mathbb{D}_n$ involves loops that start and end at the same (fixed) base point in $\partial\mathbb{D}_n$. Up to homotopy, a clockwise-directional loop that encompasses the i-th hole in $\mathbb{D}_n$ corresponds to the i-th generator of the free group F_n of rank n, which is illustrated in Figure 4.4. In fact, $\pi_1(\mathbb{D}_n) = F_n$. This equality allows us to define a representation of the braid group on n strands as automorphisms of F_n.

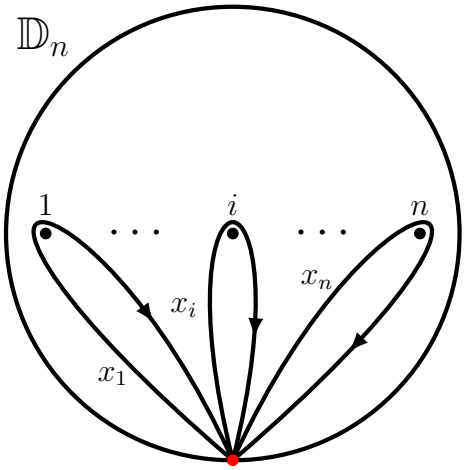

Figure 4.4: Fundamental group of the punctured disk. For each $i \in \{1, \ldots, n\}$, the clockwise-directional loop encircling the i-th hole in $\mathbb{D}_n$ corresponds to the i-th free generator of F_n (i.e., x_i). The red dot indicates the (arbitrary) base point for the loops in $\pi_1(\mathbb{D}_n)$.

Each braid $\beta \in B_n$ is realized as an automorphism of $\pi_1(\mathbb{D}_n) = F_n$ (up to isotopy) where each loop $\gamma \in \pi_1(\mathbb{D}_n)$ is sent to another loop $\beta(\gamma)$. In other words, we have a representation of the braid group defined by

$$\rho : B_n \to \mathrm{Aut}(F_n) \tag{4.1}$$

$$\beta \mapsto \rho_\beta. \tag{4.2}$$

The action of β on a loop γ is defined by the rearrangements of the n holes in $\mathbb{D}_n$, similar to the action of the standard generators of B_n on the base points in $\mathbb{C}\times\{1\}$. In terms of the standard generators of B_n, each σ_i corresponds to switching the places of hole i and hole $i+1$ by means of a clockwise rotation, as seen in Figure 4.5. This is identical to viewing the action of σ_i on the base points (Section 4.3) from above, looking down on the $\mathbb{C}\times\{t\}$-plane. As before, the inverse action ${\sigma_i}^{-1}$ is a counterclockwise rotation of the two adjacent holes i and $i+1$ in $\mathbb{D}_n$. These actions respect the group operation of loop concatenation.

Note that there is not a widely accepted convention for the standard generators of the braid group. For example, many sources will define the direction of the standard generators to be counterclockwise, where the punctures on $\mathbb{D}_n$ are swapped in the opposite manner as in this thesis. Moreover, when considering the braid group as the intertwining of strands in $\mathbb{C}\times[0,1]$, a generator σ_i may be defined as overlaying the i-th strand in front of the $(i+1)$-th strand rather than behind. This is a matter of convention, and the choice of the standard braid permutations is equivalent up to inverses.

The automorphism ρ_β is most simply defined in terms of the action of the standard generators of B_n on the generators $x_1,\ldots,x_n$ of F_n (visualized as loops in $\mathbb{D}_n$). For each i, it follows that

$$\rho_{\sigma_i}(x_i) = x_i x_{i+1} {x_i}^{-1}, \tag{4.3}$$

$$\rho_{\sigma_i}(x_{i+1}) = x_i, \tag{4.4}$$

$$\rho_{\sigma_i}(x_j) = x_j, \text{ for } j \neq i, i-1. \tag{4.5}$$

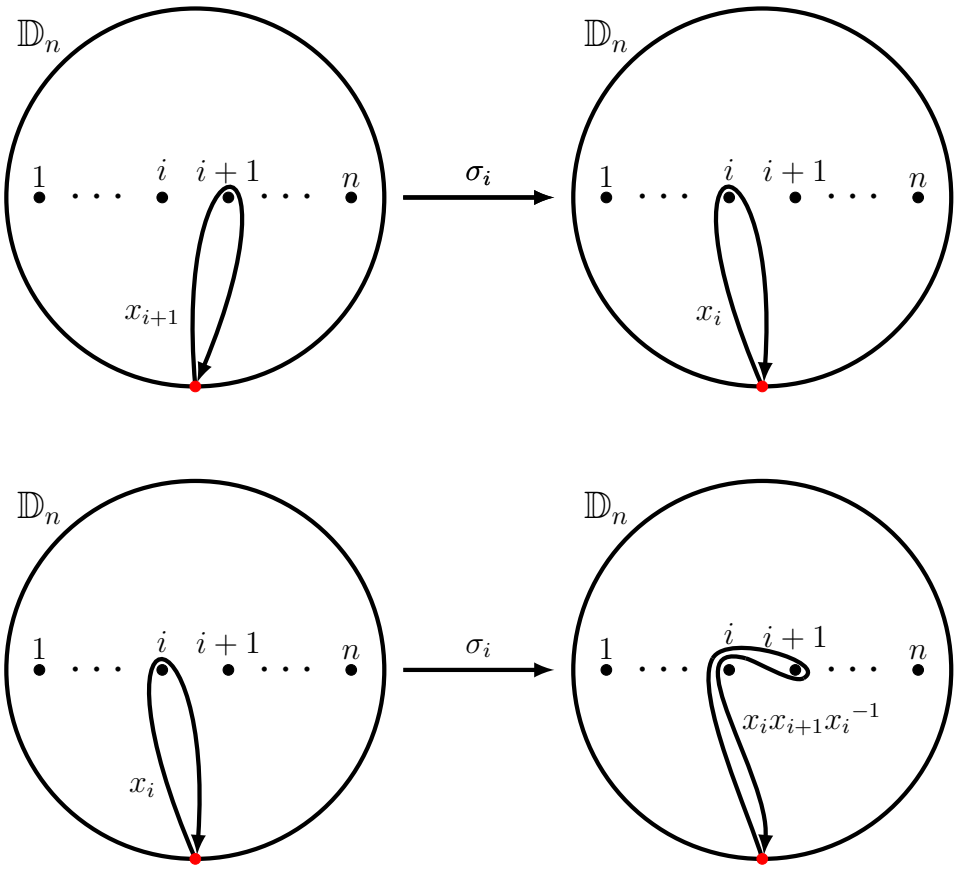

Figure 4.5: Artin generators realized on the punctured disk. The action of σ_i on the generators x_i and x_{i+1} as described by Eqns. 4.3–4.5. The image of x_i under σ_i is verified visually in Figure 4.6

Clearly, any two loops that are separated by at least one puncture will not interact while performing σ_i. The relations for adjacent loops can be verified graphically as illustrated in Figures 4.5 and 4.6.

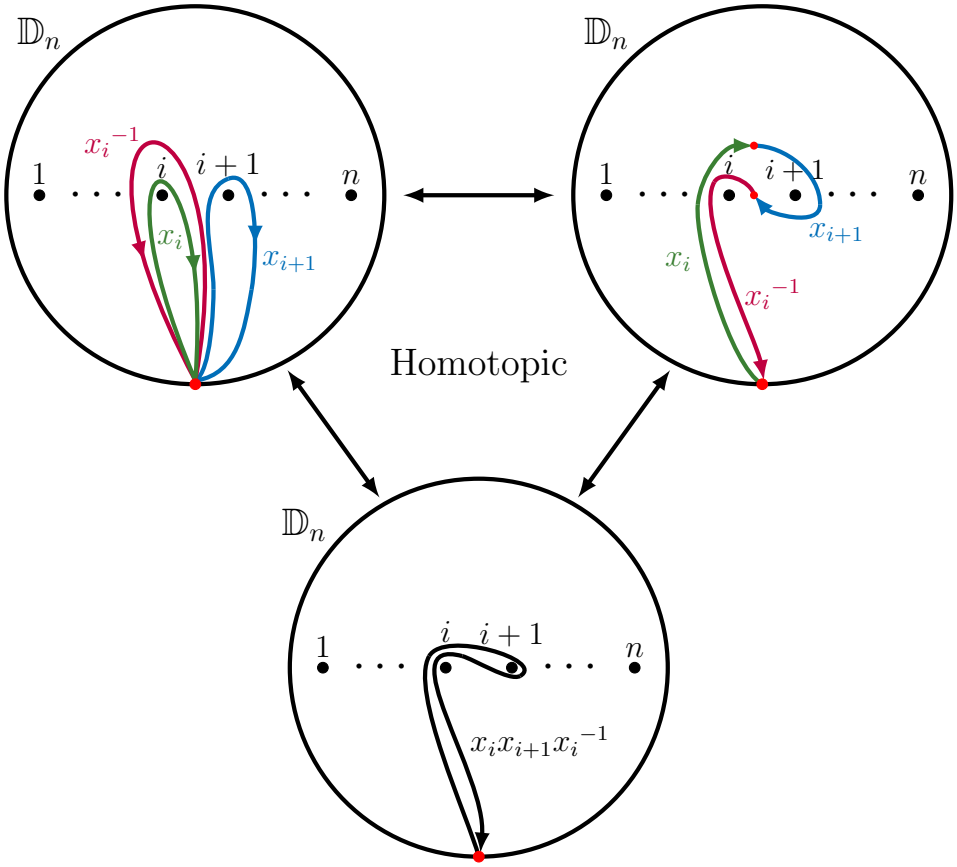

Figure 4.6: Graphical verification of Eqn. 4.3. Up to homotopy, the product $x_i x_{i+1} x_i^{-1}$ is visualized as the concatenation of the loops $x_i, x_{i+1}, x_i^{-1} \in \pi_1(\mathbb{D}_n)$. In the top right diagram, small red dots are used indicate the (homotopically deformed) points of concatenation.

For any σ_i, $\rho_{\sigma_i^{-1}}$ is well-defined. It follows that for any braid $\beta \in B_n$, we can decompose ρ_β the composition of the automorphisms of the standard generators $\sigma_1, \ldots, \sigma_{n-1}$ and their inverses that make up β.

Notice that for any σ_i, $\rho_{\sigma_i}(x_1 \cdots x_n) = x_1 \cdots x_n$. This is because the loop $x_1 \cdots x_n$ in $\mathbb{D}_n$, encompassing all holes, is parallel to the boundary $\partial\mathbb{D}_n$. Thus, the action of σ_i on $x_1 \cdots x_n$ is trivial does not affect the structure of the loop up to isotopy. More

generally, this implies that $\rho_\beta(x_1 \cdots x_n) = x_1 \cdots x_n$ for any $\beta \in B_n$. Paired with the observation that every generator is conjugate to another, Artin [1] showed that this is a necessary and sufficient condition for ρ_β to be an automorphism of F_n.

Theorem 4.1. *An automorphism $f \in Aut(F_n)$ is equal to ρ_β for some $\beta \in B_n$ if and only if*

1. *$f(x_i)$ is a conjugate of some x_j for every $i \in \{1, \ldots, n\}$, and*
2. *$f(x_1 \cdots x_n) = x_1 \cdots x_n$.*

In this interpretation of the braid group, we can express B_n in terms of the standard generators:

$$B_n = \left\langle \sigma_1, \ldots, \sigma_{n-1} \;\middle|\; \begin{array}{ll} \sigma_i\sigma_j = \sigma_j\sigma_i, & |i-j| > 1 \\ \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}, & |i-j| = 1 \end{array} \right\rangle. \tag{4.6}$$

The first condition that the standard generators commute if $|i - j| > 1$ is easily verified by thinking about the action of the generators on $\pi_1(\mathbb{D}_n)$ as in Figure 4.5 and as automorphisms in Eqn. 4.5. This follows from the fact that if punctures i and j are non-adjacent, then the actions of σ_i and σ_j are mutually exclusive and thus commutative. The second condition on the standard generators is most easily verified in Figure 4.7 by looking at the corresponding braids in $\mathbb{R} \times [0, 1]$ (here, the distinction between $\mathbb{R}$ and $\mathbb{C}$ is only a matter of choosing to show the starting and ending complex planes in traditional braid visualizations).

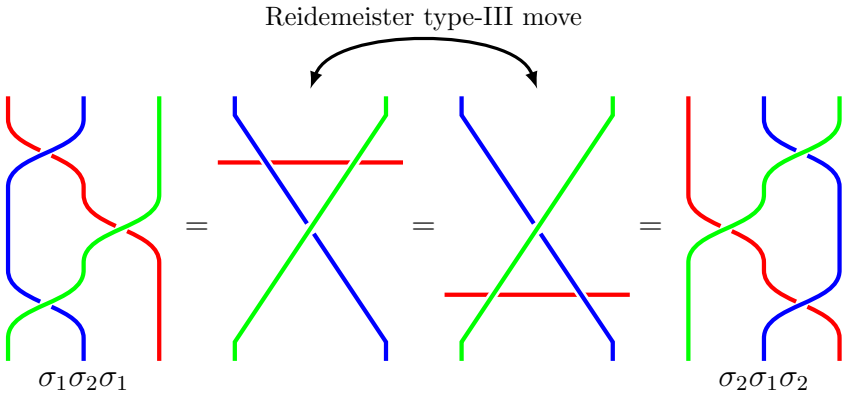

Figure 4.7: Graphical verification of Eqn. 4.6. Graphical verification of Eqn. 4.6. All of the above braids are equivalent up to homotopy. The transformation between the middle two braid diagrams is known as a Reidmeister type-III move, in which the red strand (or string in the context of knot theory) is moved completely under the crossing of the blue and green strands. In these middle two diagrams, the red strand is slightly translucent to indicate that it is behind all other strands in the diagram.

4.5 One-dimensional representations of $\mathbf{B_n}$

A straightforward nontrivial representation of the braid is defined on the standard generators of B_n and is given by:

$$p_\theta : B_n \to \mathbb{C}_{|z|=1} \tag{4.7}$$

$$\sigma_i \mapsto e^{i\theta}, \tag{4.8}$$

for $\theta \in \mathbb{R}$ [7]. Clearly, p_θ is a homomorphism, and it is unitary because

$$p_\theta(\sigma_i)^\dagger = \left(e^{i\theta}\right)^\dagger = e^{-i\theta} = \left(e^{i\theta}\right)^{-1} = p_\theta(\sigma_i)^{-1}, \tag{4.9}$$

where † denotes the conjugate transpose of a matrix.

Due to the abelian nature of the one-dimensional representations, we are free to simplify products of braids. For example,

$$\begin{aligned} p_\theta(\sigma_1\sigma_2{\sigma_i}^{-1}\sigma_2) &= p_\theta(\sigma_1)p_\theta(\sigma_2)p_\theta({\sigma_i}^{-1})p_\theta(\sigma_2) \\ &= e^{i\theta_1}e^{i\theta_2}e^{-i\theta_1}e^{i\theta_2} \\ &= e^{i(\theta_1-\theta_1+\theta_2+\theta_2)} \\ &= e^{i\cdot 2\theta_2} = p_\theta(\sigma_2^2), \end{aligned}$$

where the subscripts on θ are only used to keep track of the contributing generator to each multiple of θ. Hence, for a general braid $\beta \in B_n$, we can express $p_\theta(\beta)$ as a simplified product of the standard generators of B_n:

$$p_\theta(\beta) = p_\theta(\sigma_1^{m_1}\sigma_2^{m_2}\cdots\sigma_{n-1}^{m_{n-1}}) = e^{i\theta(m_1+m_2+\cdots+m_{n-1})} = e^{ik\theta},$$

where k can be thought of as the degree of the braid.

Notice that with a choice of $\theta = 2\pi n$ for $n \in \mathbb{Z}$, we recover the trivial representation of B_n. Similarly, $p_{\pi n}$ is a restriction of the sign representation of S_n, where the sign of a permutation is defined as the parity of the number of transpositions in its decomposition into a product of transpositions.

4.6 The Burau Representation

In Section 4.4, we defined a representation of the braid group B_n as automorphisms of the free group F_n. This representation is clearly nonabelian. Likewise, the Artin generators of B_n are nonabelian. Suppose we wish to abelianize the braid group. The details of the abelianization of B_n would require a quotient by the commutator $[a, b] = aba^{-1}b^{-1}$.

Sparing the details, let $B_{n,ab} = B_n / [B_n, B_n]$ be the abelianization of B_n, where $[B_n, B_n] = \left\{ [\beta_1, \beta_2] \;\middle|\; \beta_1, \beta_2 \in B_n \right\}$ is the commutator subgroup of B_n. Then, under the representation ρ from Section 4.4, the abelianization of Eqns. 4.3–4.5 become

$$x_i \overset{\sigma_i}{\longmapsto} \cancel{x_i} + x_{i+1} - \cancel{x_i} = x_{i+1} = \rho_{{\sigma_i}^{-1}}(x_i), \tag{4.10}$$

$$x_{i+1} \overset{\sigma_i}{\longmapsto} x_i, \tag{4.11}$$

$$x_j \overset{\sigma_i}{\longmapsto} x_j, \text{ for } j \neq i, i-1, \tag{4.12}$$

for each i. Thus, the generator $\sigma_i = {\sigma_i}^{-1}$, and corresponds to a transposition permutation in the symmetric group S_n. It follows that $B_{n,ab} \simeq S_n$. In this current construction, the abelianization of the braid group results in a loss of complexity. This raises the question whether there exists such a reframing of the braid group that allows an abelian operation on the free generators while preserving the inequivalence of the Artin generators with their inverses.

First, we define a topological space that will aid in the desired construction.

Definition 4.5. Let X be a topological space. A *covering* of X is a space $\widetilde{X}$ together with a continuous map $p : \widetilde{X} \to X$ such that, for every $x \in X$, there exists a path-connected open neighborhood U containing x such that $p^{-1}(U)$ is a disjoint union of open sets in $\widetilde{X}$ where each component of $p^{-1}(U)$ is mapped homeomorphically onto U by p. Each component of $p^{-1}(U)$ is called a *sheet* of the covering, where the i-th sheet is denoted by $\widetilde{X}_i$, and total the number of sheets in $p^{-1}(U)$ is called the *degree* of the covering.

Example 4.1. One of the simplest examples of a covering space is the covering of the circle S^1 by the real line by the parameterization map $p : \mathbb{R} \to S^1$ defined by $p(t) = (\cos t, \sin t)$. Clearly, there are infinitely many sheets in this covering.

Example 4.2. A similar example is the covering of the circle S^1 through $p : [0, 1] \to S^1$ defined by $p(t) = e^{2\pi it}$. This defines a one-degree covering of S^1. If we instead let our domain be $[0, 2]$, then we have a two-degree covering of S^1.

With this topological tool, we construct a countably infinite-degree covering of the punctured disk $\mathbb{D}_n$, denoted $\widetilde{\mathbb{D}}_n$, which can be visualized as an infinite stack of copies of $\mathbb{D}_n$, with a slight modification to be explained shortly. Let $\widetilde{\mathbb{D}}_{n,i}$ denote the i-th sheet of $\widetilde{\mathbb{D}}_n$, and consider the base sheet our covering to be $\widetilde{\mathbb{D}}_{n,0}$.

We start this construction with a countably infinite stack of copies of $\mathbb{D}_n$. Then, for every $i \in \mathbb{Z}$, for each of the n punctures on $\widetilde{\mathbb{D}}_{n,i}$, apply a cut from the hole to some point on the boundary of $\widetilde{\mathbb{D}}_{n,i}$, as illustrated in Figure 4.8 for the case when $n = 3$. Each cut results in two edges, which will be referred to as the left edge and the right edge. Through a homeomorphic deformation, connect the left edge of $\widetilde{\mathbb{D}}_{n,i}$ to the corresponding right edge of $\widetilde{\mathbb{D}}_{n,i+1}$, and the right edge of $\widetilde{\mathbb{D}}_{n,i}$ to the left edge of $\widetilde{\mathbb{D}}_{n,i-1}$, for every cut on every sheet.

Now, viewing a single sheet, say $\widetilde{\mathbb{D}}_{n,0}$, from above, when a loop with base point $\tilde{\xi}_0$ passes through a cut from the left, it traverses up to the next sheet, and ends at the base point $\tilde{\xi}_1$. Similarly, a loop passing through a cut from the right ends at the base point $\tilde{\xi}_{-1}$, on the sheet $\widetilde{\mathbb{D}}_{n,-1}$ below $\widetilde{\mathbb{D}}_{n,0}$. To keep track of the various loops, we use a free parameter t. For example, a loop γ that starts on $\widetilde{\mathbb{D}}_{n,j}$ would be written $t^j\gamma$, for $j \in \mathbb{Z}$. Notice that the substitution of a complex number for the free parameter t results in a possibly finite degree covering. As an example, if we set t to an n-th root of unity, then we obtain an n-th degree covering of $\mathbb{D}_n$. For the purposes of this construction, we will keep t as a free parameter for now. Figure 4.8 demonstrates how loops interact with the cuts on different sheets. The following example describes the action of the standard generators of B_3 on the covering space $\widetilde{\mathbb{D}}_3$.

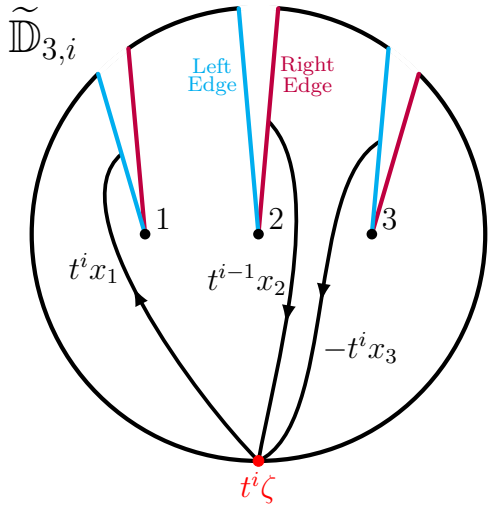

Figure 4.8: Punctured disk covering space. For the covering of $\mathbb{D}_3$, we observe the i-th sheet of $\widetilde{\mathbb{D}}_3$ with the cuts applied across each of the three punctures. The base point of the loop is indicated by the red dot, and labelled as $t^i\xi$. The power of t indicates that we are on the i-th sheet of the covering. The portions of three different loops are drawn to illustrate the behavior of loops as they pass through various edges on $\widetilde{\mathbb{D}}_{3,i}$. The loop that would traditionally be x_1 is labeled by $t^i x_1$ to indicate that it's starting on the i-th sheet. When it passes through the left edge corresponding to this puncture labeled with a 1, it traverses up to the sheet $\widetilde{\mathbb{D}}_{3,i+1}$ and ends at base point $t^{i+1}\xi$. Similarly, the loop that starts on $\widetilde{\mathbb{D}}_{3,i-1}$ and passes through the left edge of puncture 2 ends up coming out of the right edge of puncture 2 on $\widetilde{\mathbb{D}}_{3,i}$ and ends at the base point $t^i\xi$. This loop is labeled by the starting sheet, so it is $t^{i-1}x_2$. Finally, the loop that starts on $\widetilde{\mathbb{D}}_{3,i+1}$ and passes through the right edge of puncture 3 ends up coming out of the left edge of puncture 3 on $\widetilde{\mathbb{D}}_{3,i}$. This loop is labeled by $-t^i x_3$ since it is the inverse of $t^i x_3$, with the negative sign indicating that the loop direction is reversed.

Example 4.3. Consider the case when $n = 3$. Then we have the corresponding covering space $\widetilde{\mathbb{D}}_3$ of $\mathbb{D}_3$. See Figure 4.8 for the view of a single sheet with various loops interacting with the cuts on the sheet. The actions of the standard generators of B_3 in $\pi_1(\mathbb{D}_3)$ are known, and can be visually understood in Figure 4.5. In the context of the covering space $\widetilde{\mathbb{D}}_3$, the action of the generators σ_1 and σ_2 is observed by reducing the visualization to only the base points on each sheet and the loops themselves. This can be seen in Figure 4.9 for the case of σ_1.

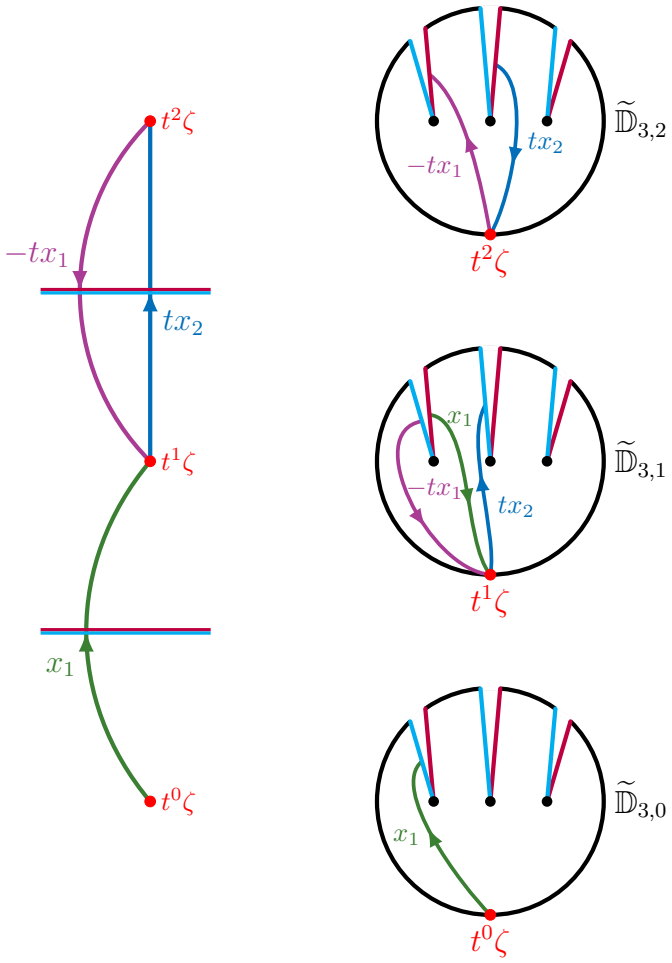

Figure 4.9: Covering space loop. The loop $\rho_{\sigma_1}(x_1) = {x_1}^{-1}x_2x_1$ cast onto the covering space $\widetilde{\mathbb{D}}_3$. On the right, the loops are depicted on the corresponding sheets of the covering, and on the left is a simplified version. The order of loops is $x_1, tx_2, -tx_1$. The horizontal lines on the left indicate the edges taking each loop up/down a sheet.

Now, we can express loop concatenation as an abelian operation, where Eqns. 4.3–4.5 become

$$x_1 \overset{\sigma_1}{\longmapsto} x_1 + tx_2 - tx_1 = (1-t)x_1 + tx_2, \tag{4.13}$$

$$x_2 \overset{\sigma_1}{\longmapsto} x_1, \tag{4.14}$$

$$x_3 \overset{\sigma_1}{\longmapsto} x_3. \tag{4.15}$$

Consider the vector $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Then the action of σ_1 on the loops x_1, x_2, x_3 is realized by the matrix $\begin{bmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ since

$$\begin{bmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (1-t)x_1 + tx_2 \\ x_1 \\ x_3 \end{bmatrix}. \tag{4.16}$$

The action of σ_2 is obtained similarly, where

$$\sigma_2 \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{bmatrix}. \tag{4.17}$$

Notice that these matrices have entries in the ring of Laurent polynomials, denoted $\Lambda = \mathbb{Z}[t, t^{-1}]$.

The result from Example 4.3 generalizes to the case of braids on n strands. Fix $n > 1$. Let I_k denote the $k \times k$ dimensional identity matrix, and let

$$U = \begin{bmatrix} 1-t & t \\ 1 & 0 \end{bmatrix}. \tag{4.18}$$

For $i \in \{1, \ldots, n-1\}$, the action of σ_i on $\pi_1\big(\widetilde{\mathbb{D}}_n\big)$ is realized as an $n \times n$ matrix with entries in $\Lambda = \mathbb{Z}[t, t^{-1}]$.

The Burau representation of B_n is then defined by:

$$\psi_n : B_n \to \mathrm{GL}_n(\Lambda) \tag{4.19}$$

$$\sigma_i \mapsto \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & I_{n-i-1} \end{bmatrix}. \tag{4.20}$$

The Burau representation need only be defined on the standard generators, since any braid $\beta \in B_n$ decomposes into a product of $\sigma_1, \ldots, \sigma_{n-1}$ and their inverses. Notice that if we set $t = 1$, we recover the defining representation of S_n, as expected when we use a degree 1 covering space of $\mathbb{D}_n$ and force the action of the generators to be abelian.

By the Cayley-Hamilton, it follows that [14]

$$U^2 - \mathrm{tr}\,(U)\,U - \det(U) I_2 = U^2 - (1-t)U - tI_2 = 0. \tag{4.21}$$

Rewriting gives

$$t^{-1}U(U - (1-t)I_2) = I_2, \tag{4.22}$$

or simply

$$U^{-1} = t^{-1}(U - (1-t)I_2) = \begin{bmatrix} 0 & 1 \\ t^{-1} & 1-t^{-1} \end{bmatrix}. \tag{4.23}$$

The block structure of the matrix representation of σ_i results in

$$\psi_n(\sigma_i)^{-1} = \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & U^{-1} & 0 \\ 0 & 0 & I_{n-i-1} \end{bmatrix}.$$

Furthermore, by direct computation [14], it follows that

$$\psi_n(\sigma_i)\psi_n(\sigma_j) = \psi_n(\sigma_j)\psi_n(\sigma_i) \text{ for } |i-j| > 1, \tag{4.24}$$

$$\psi_n(\sigma_i)\psi_n(\sigma_{i+1})\psi_n(\sigma_i) = \psi_n(\sigma_{i+1})\psi_n(\sigma_i)\psi_n(\sigma_{i+1}) \text{ for } |i-j| = 1. \tag{4.25}$$

Moreover, the Burau representation is compatible with the natural inclusion map $\iota : B_n \hookrightarrow B_{n+1}$ for $n \geq 1$ and $\beta \in B_n$:

$$\psi_{n+1}(\iota(\beta)) = \begin{bmatrix} \psi_n(\beta) & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.26}$$

4.7 The Reduced Burau Representation

Recall that in Section 4.4, every braid in B_n is described by an automorphism of the free group F_n. Equivalently, we thought of each generator σ_i as a clockwise rearrangement of adjacent punctures of $\mathbb{D}_n$ and recorded the resulting transformation of the n free generators of $F_n = \pi_1(\mathbb{D}_n)$, which were the loops around the punctures. It is intuitive to understand why the loop $\ell = x_1 \cdots x_n$ is invariant under any braid-

realized automorphism, as the loop encompasses all n punctures in $\mathbb{D}_n$. This was discussed in Section 4.4, and was formalized in Theorem 4.1 as a necessary and sufficient condition for the realization of braids as automorphisms on the free group.

The invariant loop ℓ gives insight into the Burau representation. As was done with the generators $x_1, \ldots, x_n$ when constructing the Burau representation from $\pi_1(\mathbb{D}_n)$, the loop ℓ is written as an element of the free Λ-module of rank n, denoted Λ^n, when cast onto the covering space $\widetilde{\mathbb{D}}_n$:

$$x_1 \cdots x_n \in F_n \mapsto x_1 + tx_2 + t^2x_3 + \cdots + t^{n-1}x_n = \sum_{i=1}^{n} t^{i-1}x_i \in \Lambda^n. \tag{4.27}$$

The right-hand side of Eqn. 4.27 manifests as a vector $\nu = [1, t, t^2, \ldots, t^{n-1}]^\top$, where the vector components correspond to the coefficients of the free generators $x_1, \ldots, x_n$. Note that this vector comes from the coordinate representation of Λ^n rather than directly from a vector space since Λ is not a field, but this serves as useful way to describe the elements of Λ^n. The coordinate description of Λ^n is useful in constructing a non-trivial invariant vector space with respect to the Burau representation.

In particular, notice that for any $\sigma_i \in B_n$, the matrix representation $\psi_n(\sigma_i)$ leaves ν invariant. This can be verified directly by computing the matrix multiplication:

$$\psi_n(\sigma_i)\nu = \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & I_{n-i-1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ t^{i-1} \\ t^i \\ \vdots \\ t^{n-1} \end{bmatrix}$$

$$= \begin{bmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ t^{i-1} \\ t^i \\ \vdots \\ t^{n-1} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ \vdots \\ t^{i-1} - t^i + t^i \\ t^i \\ \vdots \\ t^{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ t^{i-1} \\ t^i \\ \vdots \\ t^{n-1} \end{bmatrix} = \nu,$$

for $i = 1, 2, \ldots, n-1$. Since ν is invariant under the generators of B_n, it is also invariant under any arbitrary braid in B_n. Therefore, we have a non-trivial ψ_n-invariant subspace defined by span $\{\nu\}$, which implies that ψ_n is reducible (provided $n > 1$).

The goal of the following procedure is to obtain an irreducible representation of B_n from the Burau representation. First, given the invariance of ν under ψ_n, we find a basis for a ψ_n-invariant vector space by looking for eigenvalues of $\psi_n(\sigma_i)$ for $i = 1, 2, \ldots, n-1$ [6].

Let $\mathbf{v}_i$ be defined as the vector with i-th component $-t$, $(i+1)$-th component equal to 1, and all other components equal to zero, for $i = 1, 2, \ldots, n-1$. Again, by direct computation, we can verify that $\mathbf{v}_i$ is an eigenvector of $\psi_n(\sigma_i)$ with eigenvalue $-t$:

$$\psi_n(\sigma_i)\mathbf{v}_i = \begin{bmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-i-1} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -t \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ t^2 \\ -t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = -t\mathbf{v}_i.$$

Proposition 4.2. *Define* $\mathbf{v}_i$ *as above. Then span* $\{\mathbf{v}_i\}$ = *span* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ *form a non-trivial* $(n-1)$*-dimensional* ψ_n*-invariant vector space.*

Proof. The vectors $\mathbf{v}_i$ are eigenvectors of $\psi_n(\sigma_i)$ with eigenvalue $-t$ for each $i = 1, 2, \ldots, n-1$. These eigenvectors are clearly linearly independent, and so span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ is a non-trivial $(n-1)$-dimensional vector space. Clearly $\psi_n(\sigma_i)\mathbf{v}_j = \mathbf{v}_j$ for $|i-j| > 1$, which is easily verified due to the block diagonal structure of ψ_n. Leveraging this block structure of $\psi_n(\sigma_i)$ also allows the following

calculations

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -t \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -t \\ 1-t \\ 1 \end{bmatrix} = \begin{bmatrix} -t \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -t \\ 1 \end{bmatrix},$$
$$\begin{bmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -t \\ 1 \end{bmatrix} = \begin{bmatrix} -t^2 \\ 0 \\ 1 \end{bmatrix} = t \begin{bmatrix} -t \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -t \\ 1 \end{bmatrix},$$

which implies that

$$\psi_n(\sigma_i)\mathbf{v}_{i-1} = \mathbf{v}_{i-1} + \mathbf{v}_i,$$
$$\psi_n(\sigma_i)\mathbf{v}_{i+1} = -t\mathbf{v}_i + \mathbf{v}_{i+1}.$$

Therefore, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ are $\psi_n(\sigma_i)$ invariant for all $i = 1, 2, \ldots, n-1$. Hence, $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ is an $(n-1)$-dimensional ψ_n-invariant vector space. □

We can now define the reduced Burau representation in terms of the (unreduced) Burau representation:

$$\psi_n^{\mathbf{r}} : B_n \to \mathrm{GL}_{n-1}(\Lambda) \tag{4.28}$$
$$\sigma_i \mapsto \psi_n(\sigma_i)\big|_{\operatorname{span}\{\mathbf{v}_i\}}. \tag{4.29}$$

In practice, the matrices for $\psi_n^{\mathbf{r}}(\sigma_i)$ are found by means of a slightly different approach. The exact arguments are beyond the scope of this thesis, but can be found in [19].

Let $V \in \mathrm{GL}_3(\Lambda)$ be defined by

$$V = \begin{bmatrix} 1 & t & 0 \\ 0 & -t & 0 \\ 0 & 1 & 1 \end{bmatrix}. \tag{4.30}$$

Then the reduced Burau representation matrices are given by

$$\psi_n^{\mathbf{r}}(\sigma_1) = \begin{bmatrix} -t & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & I_{n-3} \end{bmatrix}, \tag{4.31}$$

$$\psi_n^{\mathbf{r}}(\sigma_i) = \begin{bmatrix} I_{n-2} & 0 & 0 \\ 0 & V & 0 \\ 0 & 0 & I_{n-i-2} \end{bmatrix}, \tag{4.32}$$

$$\psi_n^{\mathbf{r}}(\sigma_{n-1}) = \begin{bmatrix} I_{n-3} & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & -t \end{bmatrix}, \tag{4.33}$$

for $i = 2, 3, \ldots, n-2$ and $n \geq 3$ [14]. The exact form of the above matrices may differ by a transpose in some sources, but the results are equivalent.

The Burau representation has been shown to be faithful for $n \leq 3$ and degenerate for $n \geq 5$ [14]. Whether the Burau representation is faithful for $n = 4$ remains an open problem, but has been shown to be faithful almost everywhere [5].

4.8 Unitary Representation Matrices

As explained in [6], we can use U from the construction of the (unreduced) Burau representation to show that there are no choices of the parameter t such that the

representation matrices are directly unitary. However, we can use the reduced Burau representation matrices to construct a new representation of B_n that is in fact unitary [6, 21].

First, let $t = s^2$ for some $s \in \mathbb{C}^*$. In other words, we are restricting t to be a square of a nonzero complex number. Define the $(n-1) \times (n-1)$ matrices

$$P_{n-1} = \mathrm{diag}(s, s^2, \ldots, s^{n-1}), \tag{4.34}$$

$$J_{n-1} = \begin{bmatrix} s+s^{-1} & -1 & \cdots & 0 \\ -1 & s+s^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & -1 & s+s^{-1} \end{bmatrix}, \tag{4.35}$$

where diag denotes the diagonal matrix with the given diagonal entries. Furthermore, define $\widetilde{\psi}_n(\beta) = P_{n-1}(\psi_n^{\mathrm{r}}(\beta))^\dagger (P_{n-1})^{-1}$ for $\beta \in B_n$. Note that

$$(P_{n-1})^{-1} = \mathrm{diag}(s^{-1}, s^{-2}, \ldots, s^{-(n-1)}).$$

Suppose that for specific choices of $s \in \mathbb{C}^*$, we can decompose $J_{n-1}(s) = X^\dagger X$ for some $(n-1) \times (n-1)$ matrix X. Then

$$(J_{n-1}(s))^\dagger = \left(X^\dagger X\right)^\dagger = X^\dagger \left(X^\dagger\right)^\dagger = X^\dagger X = J_{n-1}(s),$$

which implies that $J_{n-1}(s)$ is Hermitian. In general, $J_{n-1}(s)$ is Hermitian if $s+s^{-1} \in \mathbb{R}$. In other words, either $s \in \mathbb{R}^*$, or if $s \in \mathbb{C}^*$, then

$$\begin{aligned} s + s^{-1} = \overline{s} + \overline{s^{-1}} &\iff s + \frac{\overline{s}}{|s|^2} = \overline{s} + \frac{s}{|s|^2} \\ &\iff s - \overline{s} = \frac{s - \overline{s}}{|s|^2} \\ &\iff |s| = 1, \end{aligned}$$

where $s\overline{s} = |s|^2$.

Thus, $s \in \mathbb{R}^* \cup \mathbb{C}_{|z|=1}$, which gives a more specific description of the diagonals of $J_{n-1}(s)$:

$$s + s^{-1} = s + \overline{s} = 2\mathrm{Re}(s), \tag{4.36}$$

where $\mathrm{Re}(z)$ denotes the real part of $z \in \mathbb{C}$. However, if J_{n-1} is to be decomposed into $X^\dagger X$, then we can further restrict the allowable values of s. Let $x_{i,j}$ denote the (i,j)-th entry of X. Then it must be that

$$2\mathrm{Re}(s) = \sum_{i=1}^{n-1} \overline{x}_{i,j} x_{i,j} = \sum_{i=1}^{n-1} |x_{i,j}|^2 \geq 0,$$

for all $j = 1, 2, \ldots, n-1$. Since X is not the zero matrix, $\mathrm{Re}(s) > 0$. Furthermore, we place a restriction on X by observing that

$$\left(X^\dagger X\right)_{i,j} = \sum_{k=1}^{n-1} \overline{x}_{k,i} x_{k,j} = -\delta_{i,j\pm 1},$$

where δ_{ij} is the Kronecker delta.

For such a choice of s (hence X and t) and any braid $\beta \in B_n$, the corresponding matrix

$$X\widetilde{\psi}_n(\beta)X^{-1} = XP_{n-1}(\psi_n^{\mathbf{r}}(\beta))^{\dagger}(P_{n-1})^{-1}X^{-1} = (XP_{n-1})\,(\psi_n^{\mathbf{r}}(\beta))^{\dagger}(XP_{n-1})^{-1}$$

is unitary.

Example 4.4. If $s = e^{i\frac{\pi}{6}} = \frac{\sqrt{3}}{2} + \frac{i}{2}$, then $2\mathrm{Re}(s) = \sqrt{3}$. In the 2×2 case ($n = 3$), trial and error calculations in `MATLAB` resulted in one of many possibilities for X as

$$X = \begin{bmatrix} -\frac{1}{\sqrt{2\left(\sqrt{2}+\sqrt{3}\right)}} & \sqrt{\frac{1}{2}\left(\sqrt{2}+\sqrt{3}\right)} \\ \sqrt{\frac{1}{2}\left(\sqrt{2}+\sqrt{3}\right)} & -\frac{1}{\sqrt{2\left(\sqrt{2}+\sqrt{3}\right)}} \end{bmatrix}$$

Direct computation can verify that

$$X^{\dagger}X = J_2(e^{i\frac{\pi}{6}}) = \begin{bmatrix} \sqrt{3} & -1 \\ -1 & \sqrt{3} \end{bmatrix}$$

Recall that $t = s^2 = e^{i\frac{\pi}{3}}$. The corresponding reduced Burau representation matrices are:

$$\psi_3^{\mathbf{r}}(\sigma_1) = \begin{bmatrix} -e^{i\frac{\pi}{3}} & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \psi_3^{\mathbf{r}}(\sigma_2) = \begin{bmatrix} 1 & e^{i\frac{\pi}{3}} \\ 0 & -e^{i\frac{\pi}{3}} \end{bmatrix}$$

Notice that the involution $\eta_n : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ defined by $A \mapsto A^{\dagger}$ is an automorphism of $\mathrm{GL}_n(\mathbb{C})$. Then $\eta_2 \circ \psi_3^{\mathbf{r}}$ is a homomorphism and thus a representation of B_3. In other words, applying the hermitian adjoint to the reduced Burau representation matrices results in a new representation.

Finally, traditional conjugation of $\eta_2 \circ \psi_3^{\mathbf{r}}$ by the matrix XP_{n-1} gives a representation $\mathcal{U} : B_3 \to U(2)$ defined by

$$\begin{aligned}\mathcal{U}(\sigma_1) &= (XP_{n-1})\,(\psi_n^{\mathbf{r}}(\sigma_1))^{\dagger}(XP_{n-1})^{-1} \\ &= \frac{1}{2}e^{-i\frac{\pi}{6}}\begin{bmatrix} \sqrt{3}\,e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)} & 1 \\ 1 & -\sqrt{3}\,e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)} \end{bmatrix}\end{aligned}$$

$$\begin{aligned}\mathcal{U}(\sigma_2) &= (XP_{n-1})\,(\psi_n^{\mathbf{r}}(\sigma_2))^{\dagger}(XP_{n-1})^{-1} \\ &= \frac{1}{2}e^{-i\frac{\pi}{6}}\begin{bmatrix} -\sqrt{3}\,e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)} & 1 \\ 1 & \sqrt{3}\,e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)} \end{bmatrix},\end{aligned}$$

where $U(2)$ is the group of 2×2 unitary matrices. The unitarity of $\mathcal{U}(\sigma_1)$ and $\mathcal{U}(\sigma_2)$ can be verified by direct computation. Additionally, $\mathcal{U}$ satisfies the braid relation:

$$\mathcal{U}(\sigma_1)\mathcal{U}(\sigma_2)\mathcal{U}(\sigma_1) = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \mathcal{U}(\sigma_2)\mathcal{U}(\sigma_1)\mathcal{U}(\sigma_2).$$

For the permutation group, transpositions are their own inverse, and so the representation matrices for transpositions are involutory. The additional complexity of the braid group is highlighted by the fact that the representation matrices corresponding to ${\sigma_1}^{-1}$ and ${\sigma_2}^{-1}$ are not the same as those corresponding to σ_1 and σ_2. This is easily verified by obtaining the inverses of $\mathcal{U}(\sigma_1)$ and $\mathcal{U}(\sigma_2)$ through application of

the Hermitian adjoint due to the unitarity of $\mathcal{U}$:

$$\mathcal{U}({\sigma_1}^{-1}) = \mathcal{U}(\sigma_1)^\dagger = \frac{1}{2}e^{i\frac{\pi}{6}} \begin{bmatrix} \sqrt{3}\, e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)} & 1 \\ 1 & -\sqrt{3}\, e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)} \end{bmatrix}$$

$$\mathcal{U}({\sigma_2}^{-1}) = \mathcal{U}(\sigma_2)^\dagger = \frac{1}{2}e^{i\frac{\pi}{6}} \begin{bmatrix} -\sqrt{3}\, e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)} & 1 \\ 1 & \sqrt{3}\, e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)} \end{bmatrix}.$$

Thus, σ_1^2 and σ_2^2 are not the identity braid, which is in contrast to the permutation group. This aspect of the braid group results in non-trivial consequences for permutations of objects that are governed by the braid group rather than the permutation group. This is the focus of the following chapter.

Chapter 5

ANYONS: A CONSEQUENCE OF BRAIDING PARTICLES

5.1 Braiding action on a quantum system

As discussed in [7], each of the standard generators of the braid group can be realized as unitary operators on a quantum system. For the one-dimensional representations of the braid group (Section 4.5), the action on a quantum system is merely a phase shift by $e^{ik\theta}$ for some $\theta \in \mathbb{R}$ and $k \in \mathbb{Z}$ the degree of the braid. We can see this explicitly by considering the action of $\beta \in B_n$ on a wavefunction $\psi(r_1, \ldots, r_n)$ by permuting the identical particles fixed at positions $r_1, r_2, \ldots, r_n$:

$$\psi(r_{1'}, r_{2'}, \ldots, r_{n'}) = p_\theta(\beta)\psi(r_1, r_2, \ldots, r_n) = e^{ik\theta}\psi(r_1, r_2, \ldots, r_n),$$

where $r_{i'}$ denotes particle i's position after the braid β has been applied.

Physically, we can think of each σ_i as a clockwise exchange of identical particles i and $i+1$ that live in 2 spatial dimensions. Using a third time dimension, the spacetime trajectories of the particles can be realized as a braid. In general, particles that obey the braid group permutation rules are known as *anyons* [27], and their statistics are determined by the braid group representation that describes the system.

Higher degree representations of the braid group are generally nonabelian. Thus, the action of the braid group on a quantum system will not always be as straightforward as a phase shift. This is a direct result of the "omniscience" the braid group, in the sense that a braid is only well-defined if all particle trajectories are known. As will be investigated later, a permutation of just two anyons requires the knowledge of the

positions of all other anyons in the system. These so-called *nontrivial braiding effects* of the braid group are a key feature of anyons and distinguish them from bosons and fermions.

Before studying the consequences of nontrivial braiding effects, it is instructive to first consider the action of the braid group on a quantum system with only two anyons. In Example 4.4, a 2-dimensional, nonabelian, unitary representation of B_3 was constructed. Consider a degenerate set of two quantum states with orthonormal basis $\psi_1(r_1, r_2, r_3)$ and $\psi_2(r_1, r_2, r_3)$. Each basis state can be thought of a column vector, written as $|1\rangle$ and $|2\rangle$, for ψ_1 and ψ_2 respectively.

For the unitary representation $\mathcal{U}$ constructed in Example 4.4, the action of σ_1 on the basis states is given by

$$\begin{aligned}
\mathcal{U}(\sigma_1)\,|1\rangle &= \frac{1}{2}e^{-i\frac{\pi}{6}}\left(\sqrt{3}\,e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)}\,|1\rangle + |2\rangle\right), \\
\mathcal{U}(\sigma_1)\,|2\rangle &= \frac{1}{2}e^{-i\frac{\pi}{6}}\left(|1\rangle - \sqrt{3}\,e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)}\,|2\rangle\right).
\end{aligned}$$

Similarly, the action of σ_2 on the basis states yields

$$\begin{aligned}
\mathcal{U}(\sigma_2)\,|1\rangle &= \frac{1}{2}e^{-i\frac{\pi}{6}}\left(-\sqrt{3}\,e^{-i\arctan\left(\frac{1}{\sqrt{2}}\right)}\,|1\rangle + |2\rangle\right), \\
\mathcal{U}(\sigma_2)\,|2\rangle &= \frac{1}{2}e^{-i\frac{\pi}{6}}\left(|1\rangle + \sqrt{3}\,e^{i\arctan\left(\frac{1}{\sqrt{2}}\right)}\,|2\rangle\right).
\end{aligned}$$

As a result of this nonabelian representation, we can see that the action of the braid group on this system corresponds to nontrivial rotations in the many-particle Hilbert space that describes the quantum system [7,18]. This differs from the one-dimensional representations of the braid group, where the action on the Hilbert space is merely a phase shift on every basis state.

More generally, if we have a set of $m \geq 2$ degenerate states in terms of $r_1, \ldots, r_n$ with orthonormal basis $\psi_1, \psi_2, \ldots, \psi_m$, then the basis states transform under the action of $\sigma_k \in B_n$ as

$$\psi_i' = \sum_j \left[\Xi(\sigma_k)\right]_{ij} \psi_j,$$

where $\left[\Xi(\sigma_k)\right]_{ij}$ is the (i,j)-th entry of the unitary matrix $\Xi(\sigma_k)$ for some representation $\Xi : B_n \to U(m)$.

Just as with fermions, there are specific rules to follow when writing down the quantum state of multi-anyon systems. As established earlier in this section, the specific combination or *fusion* rules of anyons depend on the specific braid group representation acting on the system. For one-dimensional representations, the corresponding anyons are known as *abelian anyons*, in which the resulting action of the braid group is a phase shift on the quantum state. For higher-dimensional (nonabelian) representations, the anyons are known as *nonabelian anyons*, and the action of the braid group is a nontrivial rotation in the Hilbert space of the quantum system as illustrated above.

The simplest cases of nonabelian anyon fusion rules are the *Fibonacci anyons* and the *Ising anyons* [7, 16, 25]. These fusion rules are useful in realizing fault-tolerant topological quantum computers due to the nontrivial braiding effects that arise from the nonabelian representations of the braid group [6, 8, 16]. More information on the fusion rules of anyons can be found at the aforementioned references.

Now that the braid group (representations) have been established in a physical context, we can investigate the consequences of nontrivial braiding effects on a system of anyons, starting with the simplest nontrivial example of two non-interacting anyons.

5.2 Two Non-Interacting Anyons

In classical mechanics, the physics of a system is encapsulated in the Lagrangian and/or Hamiltonian of the system, which describe the system in terms of generalized coordinates and momenta [23]. For this reason, the most efficient method to study the dynamics of anyons is to construct the corresponding Lagrangian and Hamiltonian. This approach will lead to a clear understanding of the complexity of anyons and their braiding statistics as compared to traditional particles such as bosons and fermions.

The interaction term in the Lagrangian for two anyons due to the braiding of the anyons is given by

$$\mathcal{L}_{\text{int}} = \hbar\alpha\dot{\phi}, \tag{5.1}$$

where a dot indicates a total time derivative $\frac{d}{dt}$ and $\phi = \arctan\left(\frac{y_2-y_1}{x_2-x_1}\right)$ is the relative angle between the two anyons with positions $\mathbf{r}_1 = (x_1, y_1)$ and $\mathbf{r}_2 = (x_2, y_2)$. As in the previous section, $\alpha \in [0, 1]$ is the *winding angle* or braiding statistic of the anyons. The parameter α can also be thought of as an angle modulo π. Though the relative angle ϕ is ambiguous for identical particles, the derivative $\frac{d\phi}{dt} = \dot{\phi}$ is well-defined.

Notice that if we take $\alpha \to 0$, the interaction term vanishes as expected for bosons. Similarly, for $\alpha > 0$, ϕ becomes singular if $\mathbf{r}_1 = \mathbf{r_2}$, which motivates the Pauli exclusion principle for fermions. In fact, this means that for any $\alpha > 0$, the corresponding anyons exhibit some form of the Pauli exclusion principle.

The classical kinetic energy of this system is

$$T = \frac{1}{2}m\left(\dot{\mathbf{r}}_1^2 + \dot{\mathbf{r}}_2^2\right), \tag{5.2}$$

as expected. Then the Lagrangian for this system is

$$\mathcal{L}\left(\dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, \dot{\phi}\right) = T + \mathcal{L}_{\text{int}} = \frac{1}{2}m\left(\dot{\mathbf{r}}_1^2 + \dot{\mathbf{r}}_2^2\right) + \hbar\alpha\dot{\phi}, \tag{5.3}$$

which can also be viewed as the Lagrangian for 2 interacting bosons/fermions.

We can redefine the Lagrangian in terms of the relative and center-of-mass coordinates

$$\mathbf{R} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}, \tag{5.4}$$

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2, \tag{5.5}$$

where $\mathbf{r}$ is the relative position vector and $\mathbf{R}$ is the center-of-mass position vector of the two anyons. Note that we are assuming that the mass of the two particles are equal ($m_1 = m_2$). Classically, the momentum of a particle is given by the product of its mass and velocity. Then the corresponding center-of-mass and relative momenta:

$$\mathbf{P} = 2m\dot{\mathbf{R}} = 2m\frac{\mathbf{r}_1 + \mathbf{r}_2}{2} = m\dot{\mathbf{r}}_1 + m\dot{\mathbf{r}}_2 = \mathbf{p}_1 + \mathbf{p}_2, \tag{5.6}$$

$$\mathbf{p} = \mu\dot{\mathbf{r}} = \frac{m}{2}\left(\dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2\right) = \frac{\mathbf{p}_1 - \mathbf{p}_2}{2}, \tag{5.7}$$

where m is the mass of each anyon and μ is the reduced mass of the system.

With this in mind, we derive the following identity:

$$\dot{\mathbf{R}} + \frac{1}{4}\dot{\mathbf{r}} = \frac{\left(\dot{\mathbf{r}}_1 + \dot{\mathbf{r}}_2\right)^2}{4} + \frac{\left(\dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2\right)^2}{4} = \frac{\dot{\mathbf{r}}_1^2 + \dot{\mathbf{r}}_2^2}{2}. \tag{5.8}$$

Thus, the Lagrangian decomposes into relative and center-of-mass components upon making the substitution from the above identity:

$$\mathcal{L} = \underbrace{m\dot{\mathbf{R}}^2}_{\mathcal{L}_R} + \underbrace{\frac{m\dot{\mathbf{r}}^{\,2}}{4} + \hbar\alpha\dot{\phi}}_{\mathcal{L}_r}, \tag{5.9}$$

where the squared velocities indicate magnitude squared. Observe that the center-of-mass component of the Lagrangian, $\mathcal{L}_R$, is independent of the braiding parameter α. We can further simplify the relative component of the Lagrangian, $\mathcal{L}_r$, by noting that we can briefly write the coordinate $\mathbf{r}$ in polar form by representing it as a complex number $\mathbf{r} = re^{i\phi}$. It follows that

$$|\dot{\mathbf{r}}(r,\phi)|^2 = \left|\frac{d}{dt}\mathbf{r}(r,\phi)\right|^2 = \left|\left(\dot{r} + ir\dot{\phi}\right)e^{i\phi}\right|^2 = \dot{r}^2 + r^2\dot{\phi}^2. \tag{5.10}$$

Hence, we rewrite the relative component of the Lagrangian as

$$\mathcal{L}_r = \frac{m\left(\dot{r}^2 + r^2\dot{\phi}^2\right)}{4} + \hbar\alpha\dot{\phi}. \tag{5.11}$$

Recall that the classical relative angular momentum can be described by:

$$p_\phi = \frac{d\mathcal{L}}{d\dot{\phi}} = \frac{mr^2}{2}\dot{\phi} + \hbar\alpha. \tag{5.12}$$

Now, the Hamiltonian for this system can be constructed:

$$\begin{aligned}\mathcal{H} &= P\dot{R} + p_r\dot{r} + p_\phi\dot{\phi} - \mathcal{L} \\ &= \frac{P^2}{4m} + \frac{p_r^2}{m} + \frac{mr^2}{4}{p_\phi}^2 \\ &= \frac{P^2}{4m} + \frac{p_r^2}{m} + \frac{(p_\phi - \hbar\alpha)^2}{mr^2}.\end{aligned} \tag{5.13}$$

Once again, the center-of-mass component of the Hamiltonian is independent of α, and so we can focus on the relative component of the Hamiltonian, which is

$$\mathcal{H}_r = \frac{p_r^2}{m} + \frac{(p_\phi - \hbar\alpha)^2}{mr^2}. \tag{5.14}$$

For the purposes of this work, we need not carry out to find the energy eigenstates corresponding to the quantum operator of this relative Hamiltonian. More about this is found in [15].

5.3 Anyons in Harmonic Potential

Placing anyons in a harmonic potential alters the Hamiltonian obtained in Section 5.2. The potential energy of a 2-anyon system is given by

$$V(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}m\omega^2 \left(\mathbf{r}_1{}^2 + \mathbf{r}_2{}^2\right) = m\omega^2 \left(\mathbf{R}^2 + \frac{1}{4}\mathbf{r}^2\right), \tag{5.15}$$

where ω is the angular frequency of the harmonic potential. We can make the same substitution as in Section 5.2 to write the potential in terms of the relative and center-of-mass coordinates. As is a recurring theme, the center-of-mass component of the potential has no dependence on the braiding parameter α, and corresponds to a 2-dimensional quantum harmonic oscillator problem for a particle of mass $2m$.

Note that we can generalize Eqn. 5.11 (now omitting the subscript r) to an N-anyon system in a harmonic potential by writing

$$\mathcal{L} = \sum_{i=1}^{N} \frac{m}{2}\dot{\mathbf{r}}_i{}^2 + \hbar\alpha \sum_{\substack{i=1 \\ j\neq i}}^{N} \dot{\phi}_{ij} - \frac{m\omega^2}{2}\sum_{i=1}^{N} \mathbf{r}_i{}^2, \tag{5.16}$$

where $\phi_{ij} = \arctan\left(\frac{y_i - y_j}{x_i - x_j}\right)$ is the relative angle between anyons i and j. For brevity, we write $x_{ij} = x_i - x_j$ and $y_{ij} = y_i - y_j$ to denote the relative coordinates between the anyons.

More generally, let $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ be the relative coordinate between anyons i and j, and define $r_{ij}^2 = |\mathbf{r}_{ij}|^2$. Then we can solve directly for $\dot{\phi}_{ij}$ as follows:

$$\begin{aligned}
\dot{\phi}_{ij} = \frac{d\phi_{ij}}{dt} = \frac{d}{dt}\arctan\left(\frac{y_{ij}}{x_{ij}}\right) &= \frac{\frac{d}{dt}\left(\frac{y_{ij}}{x_{ij}}\right)}{1 + \left(\frac{y_{ij}}{x_{ij}}\right)^2} \\
&= \frac{x_{ij}\dot{y}_{ij} - \dot{x}_{ij}y_{ij}}{x_{ij}^2\left[1 + \left(\frac{y_{ij}}{x_{ij}}\right)^2\right]} \\
&= \frac{x_{ij}\dot{y}_{ij} - \dot{x}_{ij}y_{ij}}{x_{ij}^2 + y_{ij}^2} \\
&= \frac{\mathbf{r}_{ij} \times \dot{\mathbf{r}}_{ij}}{r_{ij}^2}.
\end{aligned}$$

Setting $\hbar = 1$, we can rewrite the Lagrangian as [4]

$$\mathcal{L} = \frac{m}{2}\sum_{i=1}^{N}\left[\dot{\mathbf{r}}^{\,2} - \omega^2 \mathbf{r}_i^{\,2}\right] + \alpha\sum_{i<j}^{N}\frac{\mathbf{r}_{ij} \times \dot{\mathbf{r}}_{ij}}{r_{ij}^2}, \tag{5.17}$$

which can be expanded as

$$\mathcal{L} = \frac{m}{2}\sum_{i=1}^{N}\left[\dot{\mathbf{r}}^{\,2} - \omega^2 \mathbf{r}_i^{\,2}\right] + \alpha\sum_{i<j}^{N}\dot{\mathbf{r}}_{ij} \cdot \frac{(-y_{ij}\hat{x} + x_{ij}\hat{y})}{r_{ij}^2}. \tag{5.18}$$

The last term in Eqn. 5.18 is of similar form to the vector (gauge) potential associated with the i-th anyon [4, 15, 17]:

$$\mathbf{A}_i(\mathbf{r}_i) = \alpha\sum_{j\neq i}\frac{\hat{z} \times \mathbf{r}_{ij}}{r_{ij}^2} = \alpha\sum_{j\neq i}\frac{-y_{ij}\hat{x} + x_{ij}\hat{y}}{r_{ij}^2}, \tag{5.19}$$

where $\hat{z}$ is the unit vector perpendicular to the r_{ij}-plane. This is analogous to the Coulomb potential for charged particles. Here, α serves as the coupling constant, which dictates the strength of the interaction between anyons in the system.

For the i-th anyon, the contribution to the Hamiltonian can be written as

$$\mathcal{H}_i = \frac{1}{2m}\big(\mathbf{p}_i - \mathbf{A}_i(\mathbf{r}_i)\big)^2 + \frac{m\omega^2}{2} r_i^2, \tag{5.20}$$

where $p_i - A_i(\mathbf{r}_i)$ is known as the *canonical momentum* of the system. This is a required modification since we must account for the motion of the anyons in the presence of the gauge potential in addition to their mechanical momentum.

More generally, a gauge potential is added to the Lagrangian and thus Hamiltonian such that the equations of motion are invariant under the relevant gauge transformation. In the case of electromagnetism, the prescription of invariance under local changes in electric and magnetic potentials leads to the introduction of the gauge potential in the Lagrangian [11, 12, 17].

Hence, the Hamiltonian for the N-anyon system in a harmonic potential is given by

$$\mathcal{H} = \frac{1}{2m}\sum_{i=1}^{N}\big(\mathbf{p}_i - \mathbf{A}_i(\mathbf{r}_i)\big)^2 + \frac{m\omega^2}{2}\sum_{i=1}^{N} r_i^2. \tag{5.21}$$

Substituting Eqn. 5.19 into Eqn. 5.21, we have

$$\mathcal{H} = \frac{1}{2m}\sum_{i=1}^{N} p_i^2 + \frac{m\omega^2}{2}\sum_{i=1}^{N} r_i^2 - \frac{\alpha}{2m}\sum_{\substack{i=1\\ j\neq i}}^{N}\frac{\ell_{ij}}{r_{ij}^2} + \frac{\alpha^2}{2m}\sum_{\substack{i=1\\ j,k\neq i}}^{N}\frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{ik}}{r_{ij}^2 r_{ik}^2}, \tag{5.22}$$

where $\ell_{ij} = (\mathbf{r}_i - \mathbf{r}_j)\times(\mathbf{p}_i - \mathbf{p}_j)$ is the relative angular momentum of anyon i and j. For details on the derivation of Eqn. 5.22, see Appendix C.

5.4 Nontrivial braiding effects

The last term in Eqn. 5.22 can be thought of as a long-range 2- and 3-particle interaction term. As previously established, anyons obeying braiding statistics exhibit nontrivial braiding effects. To better understand these nontrivial braiding effects, first consider two different trajectories of anyons, as depicted in Figure 5.1 [15].

From a top-down perspective, trajectory A involves a clockwise exchange of anyons 1 and 2 while anyon 3 does not take part in the exchange. In trajectory B, anyon 1 is again swapped with anyon 2 in a clockwise fashion, but anyon 3 is now involved in the exchange by being in between anyons 1 and 2. The differences between the symmetric group and the braid group now become apparent. In the symmetric group, both trajectories would be equivalent. However, the two trajectories are distinct in the braid group, which is clear by comparing the braid of each trajectory in the bottom row of Figure 5.1.

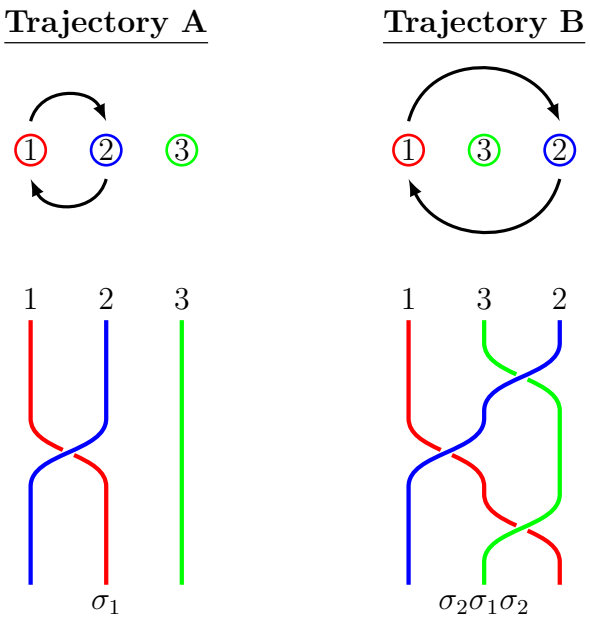

Figure 5.1: Anyon trajectories. Two possible trajectories of three anyons viewed from above (top) and represented as braids (bottom).

Even in the one-dimensional (abelian) representation of the braid group, the nontrivial braiding effects emerge for the trajectories depicted in Figure 5.1. The braid representations for trajectory A and trajectory B are $\sigma_1 \mapsto e^{i\theta}$ and $\sigma_2\sigma_1\sigma_2 \mapsto e^{3i\theta}$, respectively. As long as the choice of θ is not an integer multiple of π, it follows that $e^{i\theta} \neq e^{3i\theta}$. This is precisely the nontrivial braiding effect.

If there were some wavefunction $\psi(r_1, r_2, r_3)$ that described the system, then the action of the braid group on ψ would yield different phase changes for the two trajectories. Despite the fact that both trajectories involve the same exchange of anyons 1 and 2, the relative position of anyon 3 significantly impacts the outcome of the exchange. It is then evident that for a system of N anyons, every particle exchange must also simultaneously consider the relative positions of all other anyons in the system in order to properly encode the braiding action.

Clearly, nontrivial braiding affects must be accounted for when studying a system that obeys braiding statistics. In fact, nontrivial braiding is already encoded into the Hamiltonian found Section 5.3. For a 2-anyon system, the long-range interaction term manifests in a familiar form.

Isolating the last term in Eqn. 5.22 for $N = 2$ anyons, we see that

$$\frac{\alpha^2}{2m}\left(\frac{\mathbf{r}_{12}\cdot\mathbf{r}_{12}}{r_{12}^2 r_{12}^2} + \frac{\mathbf{r}_{21}\cdot\mathbf{r}_{21}}{r_{21}^2 r_{21}^2}\right) = \frac{\alpha^2}{m r_{12}^2},$$

which is analogous to the Coulombic interaction between two charged particles. Only when we have $N \geq 3$ anyons will the nontrivial braiding effects truly present themselves in the long-range interaction term of the Hamiltonian. For example, if we take

$N = 3$, the last term in Eqn. 5.22 becomes

$$\frac{\alpha^2}{m}\Bigg(\underbrace{\frac{1}{r_{12}^2} + \frac{1}{r_{13}^2} + \frac{1}{r_{23}^2}}_{\text{Coulomb-like interaction}} + \underbrace{\frac{\mathbf{r}_{12}\cdot\mathbf{r}_{13}}{r_{12}^2 r_{13}^2} + \frac{\mathbf{r}_{21}\cdot\mathbf{r}_{23}}{r_{21}^2 r_{23}^2} + \frac{\mathbf{r}_{31}\cdot\mathbf{r}_{32}}{r_{31}^2 r_{32}^2}}_{\text{Nontrivial braiding}} \Bigg).$$

With three anyons, the nontriviality of the long-range interaction becomes clear, as there are now cross terms that involve the relative positions of all three anyons. This result is similar to the observations made in Figure 5.1. The consequence of these three-body interaction terms is that the system is no longer separable into independent 2-anyon systems, which greatly complicates the quantum mechanics and statistical analysis of the system [15]. The nontrivial braiding effects of anyonic systems highlights the rich complexity of the braid group and its physical implications.

5.5 Conclusion

In this chapter, we have seen how the permutation of particles according to the braid group adds additional structure to the physical system as compared to the symmetric group. The study of anyons has emerged as an area of interest in condensed matter physics, such as in the fractional quantum Hall effect [3]. Arguably the largest area of study involving anyons is in the field of topological quantum computation, where the nontrivial braiding effects of anyons are the key ingredient towards realizing a fault-tolerant quantum computer [6, 8].

The braid group is but one example of the rich mathematical structure that underlies the physical world. As demonstrated in this thesis, representation theory serves as a bridge between abstract mathematics and the physical world, and offers unique insight into the fundamental nature of our universe.

BIBLIOGRAPHY

[1] E. Artin. Theory of braids. *The Annals of Mathematics*, 48(1):101, Jan. 1947.

[2] S. Axler. *Linear Algebra Done Right*. Springer International Publishing, 2024.

[3] K. I. Bolotin, F. Ghahari, M. D. Shulman, H. L. Stormer, and P. Kim. Observation of the fractional quantum hall effect in graphene. *Nature*, 462(7270):196–199, Nov. 2009.

[4] G. Date, M. V. N. Murthy, and R. Vathsan. Classical and quantum mechanics of anyons, 2003.

[5] A. Datta. A strong characterization of the entries of the burau matrices of 4-braids: The burau representation of the braid group b_4 is faithful almost everywhere, 2022.

[6] C. Delaney, E. C. Rowell, and Z. Wang. Local unitary representations of the braid group and their applications to quantum computing, 2016.

[7] A. Deshmukh. An introduction to anyons.

[8] B. Field and T. Simula. Introduction to topological quantum computation with non-abelian anyons. 2018.

[9] W. Fulton. *Algebraic Topology: A First Course*. Graduate Texts in Mathematics. Springer New York, 1997.

[10] J. Gonzalez-Meneses. Basic results on braid groups, 2010.

[11] D. J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, June 2017.

[12] D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, Aug. 2018.

[13] B. C. Hall. *Quantum Theory for Mathematicians*. Springer New York, 2013.

[14] C. Kassel and V. Turaev. *Homological Representations of the Braid Groups*, page 93–150. Springer New York, 2008.

[15] A. Khare. *Fractional Statistics and Quantum Theory*. WORLD SCIENTIFIC, Feb. 2005.

[16] V. Lahtinen and J. K. Pachos. A Short Introduction to Topological Quantum Computation. *SciPost Phys.*, 3:021, 2017.

[17] K. Moriyasu. *An Elementary Primer for Gauge Theory*. WORLD SCIENTIFIC, Oct. 1983.

[18] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. Non-abelian anyons and topological quantum computation. *Reviews of Modern Physics*, 80(3):1083–1159, Sept. 2008.

[19] M. Palmer and A. Soulié. The burau representations of loop braid groups. 2021.

[20] D. Rolfsen. Tutorial on the braid groups, 2010.

[21] C. C. Squier. The burau representation is unitary. *Proceedings of the American Mathematical Society*, 90(2):199–202, 1984.

[22] G. Strang. *Linear Algebra and Its Applications*. Thomson, Brooks/Cole, 2006.

[23] J. R. Taylor. *Classical Mechanics*. University Science Books, board book edition, 1 2005.

[24] J.-L. Thiffeault. The burau representation of the braid group and its application to dynamics. Presentation given at Topological Methods in Mathematical Physics 2022, Seminar GEOTOP-A, September 2022.

[25] S. Trebst, M. Troyer, Z. Wang, and A. W. W. Ludwig. A short introduction to fibonacci anyon models. *Progress of Theoretical Physics Supplement*, 176:384–407, 2008.

[26] W.-K. Tung. *Group theory in physics: An introduction to symmetry principles, group representations, and special functions in classical and quantum physics*. World Scientific Publishing, Singapore, Singapore, Jan. 1985.

[27] F. Wilczek. Quantum mechanics of fractional-spin particles. *Physical Review Letters*, 49(14):957–959, Oct. 1982.

[28] B. Zwiebach. *Mastering Quantum Mechanics: Essentials, Theory, and Applications*. MIT Press, 2022.

APPENDICES

Appendix A

RELEVANT TOPOLOGICAL DEFINITIONS

The braid group is formally defined in terms of topology. In order to understand the braid group, we must first understand the underlying topological properties that are used to define the braid group. The following is a brief introduction to the relevant topological concepts [9, 14].

Similar to an isomorphism in algebra, the notion of topological equivalence is given by the following definition.

Definition A.1. Consider X and Y to be two topological spaces. A *homotopy* between two continuous functions $f, g : X \to Y$ is a continuous function $H : X \times [0,1] \to Y$ such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$. If such a homotopy exists, we say that f and g are *homotopic*.

The homotopy H can be thought of as a continuous deformation of f into g. The interval $[0,1]$ represents the "time" parameter of the deformation. At time equal to 0, the function H is equal to f, and at time equal to 1, the function H is equal to g. If two functions are homotopic, then they belong to the same homotopy class, which is an equivalence class of functions under the relation of homotopy.

Definition A.2. A *loop* on a topological space X is a continuous function $\ell : [0,1] \to X$ such that $\ell(0) = \ell(1)$. In other words, the path of ℓ starts and ends at the same point in X. Often, this point is called the *base point* of the loop.

Equipped with the above definitions, the equivalence of loops on a topological space is defined as follows.

Definition A.3. A *homotopy class of loops* on a topological space X is an equivalence class of loops under the relation of homotopy. Simply put, a homotopy of loops is a continuous transformation of one loop into another. If two loops $\ell_1, \ell_2 : [0,1] \to X$ with base point $\xi \in X$ are homotopic, then there exists a continuous map $H : [0,1] \times [0,1] \to X$ such that:

1. $H(0,t) = \xi = H(1,t)$ for all $t \in [0,1]$, and
2. $H(s,0) = \ell_1(s)$ and $H(s,1) = \ell_2(s)$ for all $s \in [0,1]$.

Property 1 ensures that the starting/ending point of the loop remains fixed throughout the deformation from ℓ_1 to ℓ_2, and property 2 follows from the definition of a homotopy.

Definition A.4. The *fundamental group* of a topological space X with base point ξ is defined as the collection of loops on X with base point ξ modulo homotopy. In other words, the fundamental group is the collection of equivalence classes of loops under homotopy. This is written as

$$\pi(X, \xi) := \{\text{loops } \ell \text{ on } X \text{ with base point } \xi\} / \text{homotopy}.$$

Often times, the base point of a loop is arbitrary, so we can write $\pi(X)$ instead of $\pi(X, \xi)$ to denote the fundamental group of X.

The group structure of the fundamental group is defined as operations on the loops themselves. Consider two loops $\ell_1, \ell_2 : [0,1] \to X$ with base point ξ. Then the product $\ell_1 \cdot \ell_2$ is defined in terms of *concatenation* of the two loops. Specifically, this

defines a new loop $(\ell_1 \cdot \ell_2)(t) = \mathcal{L}(t) : [0,1] \to X$ where $\mathcal{L}(t) = \ell_1(2t)$ on $\left[0, \frac{1}{2}\right]$ and $\mathcal{L}(t) = \ell_2(2t-1)$ on $\left[\frac{1}{2}, 1\right]$. Loop concatenation can be thought of as stitching the loops together at the shared base point. As t ranges from 0 to 1, we can think of the first half of the deformation as traversing the first loop at twice the original speed, and then traveling along the second loop at twice the original speed in the second half of the deformation.

In the above description, recall that each loop ℓ is actually an equivalence class $[\ell]$ under the relation of homotopy. So the concatenation of two loops ℓ_1 and ℓ_2 is actually the concatenation of any two loops belonging to the equivalence classes $[\ell_1]$ and $[\ell_2]$, which becomes the equivalence class $[\ell_1 \cdot \ell_2]$.

In the fundamental group, the inverse of an element is the identical topological path traversed in the opposite direction. If $\gamma : [0,1] \to X$ is a loop on X, then $\gamma^{-1}(t) := \gamma(1-t)$.

Just as how homotopy describes a continuous transformation from one continuous path on a topological space to another, topological equivalence is established in a broader sense in the following definition.

Definition A.5. A continuous, bijective function $f : X \to Y$ between two topological spaces X and Y such that the inverse $f^{-1} : Y \to X$ is also continuous and bijective is called a *homeomorphism*. If there exists such a homeomorphism, then we say X is *homeomorphic* to Y. Moreover, an *embedding* of topological spaces is a continuous function that is a homeomorphism when restricted to its image. If we have a continuous family of homeomorphisms $f_t : X \to Y$ for $t \in [0,1]$, then we say that X and Y are *isotopic*. The isotopy of X and Y can be written as a function $H : X \times [0,1] \to Y$ such that

1. $H(x,0) = x$ for all $x \in X$,

2. $H(x, 1) = f(x)$ for all $x \in X$, and

3. $H(-, t)$ is an embedding of X onto Y for all $t \in [0, 1]$.

Evidently, isotopy defines a stronger and more broad notion of topological equivalence, which will be important in defining representations of the braid group.

Appendix B

PHYSICS BACKGROUND

B.1 Physics conventions and Dirac notation

More information on physics notation and conventions can be found in [12, 13, 26].

For quantum mechanics, the physical state of the system is often represented by an abstract vector belonging to some Hilbert space. The convention is to assume that the Hilbert space is complex and separable. In this context, the inner product is typically defined to be linear in the second argument. For instance, for two vectors ϕ, ψ and a scalars λ, α, we have

$$\langle \phi, \lambda\psi \rangle = \lambda \langle \phi, \psi \rangle,$$
$$\langle \alpha\phi, \psi \rangle = \overline{\alpha} \langle \phi, \psi \rangle,$$

where the overline denotes complex conjugation.

A linear operator on the Hilbert space is said to be *Hermitian* if it is self-adjoint under the inner product. The Hermitian conjugate or adjoint (conjugate transpose) of an operator A is denoted by $A^\dagger$.

Let $\mathbf{H}$ denote a quantum Hilbert space. The term "quantum" here is used to identify the vectors of $\mathbf{H}$ as representing quantum states of a system. A *ket* is a vector belonging to $\mathbf{H}$. For some vector ψ in $\mathbf{H}$, we write $|\psi\rangle$. A *bra* is the dual of a ket,

and is denoted by $\langle\psi|$. For any $\phi \in \mathbf{H}$, the bra is defined by

$$\langle\phi|\,(\psi) = \langle\phi, \psi\rangle.$$

However, this notation is not standard in the literature. Instead, a *bra-ket* is the inner product of a bra and a ket, and is denoted by $\langle\phi|\psi\rangle$.

For a linear operator $\hat{A}$ on $\mathbf{H}$, the action of $\hat{A}$ on a ket $|\psi\rangle$ is denoted by $\hat{A}\,|\psi\rangle$. Moreover, for a bra $\langle\phi|$, we have a corresponding linear functional $\langle\phi|\,\hat{A}$ defined by

$$\langle\phi|\,\hat{A}(\psi) = \langle\phi|\,\hat{A}\,|\psi\rangle\,.$$

This inner product can either be thought of as the application of $\langle\phi|\,\hat{A}$ to $|\psi\rangle$ or as the application of $\langle\phi|$ to the vector $\hat{A}\,|\psi\rangle$. The notion of an adjoint is given by

$$\langle\phi|\,\hat{A} = \langle\hat{A}^{\dagger}\phi|\,.$$

One defines an outer product of two vectors $|\psi\rangle$ and $|\phi\rangle$ as the linear operator $|\psi\rangle\,\langle\phi|$, which acts on a vector $|\chi\rangle$ as

$$(|\psi\rangle\,\langle\phi|)(\chi) = |\psi\rangle\,\langle\phi|\chi\rangle = \langle\phi|\chi\rangle\,|\psi\rangle\,.$$

For a set of orthonormal basis vectors $\{|n\rangle\}$, one can expand an arbitrary vector $|\psi\rangle$ as

$$|\psi\rangle = \left(\sum_n |n\rangle\,\langle n|\right)|\psi\rangle = \sum_n |n\rangle\,\langle n|\psi\rangle\,,$$

where $\langle n|\psi\rangle$ are the components of $|\psi\rangle$ in the basis $\{|n\rangle\}$. This is simply a change of basis, and $\sum_n |n\rangle \langle n|$ is equivalent to the identity operator.

For a continuous basis labelled by $|x\rangle$ where x is a continuous parameter, the *wavefunction* $\psi(x)$ is used to express $|\psi\rangle$ in terms of the orthonormal states $|x\rangle$. The wavefunction is the projection of $|\psi\rangle$ onto $|x\rangle$:

$$\langle x|\psi\rangle = \psi(x).$$

Suppose we have a quantum mechanical object that exists in the super position of orthonormal states $|1\rangle, |2\rangle$. The state of the object is given by the wavefunction $\Psi(x,t)$ whose square magnitude gives the probability density of the object being at position x at time t. The wavefunction $\Psi(x,t)$ must be normalized and thus square integrable.

For some physically measurable quantity A, often called an *observable*, the *expectation value* of A with the associated operator $\hat{A}$ is given by

$$\langle A\rangle = \int dx\ \overline{\Psi}(x,t)\hat{A}\Psi(x,t),$$

which can be rewritten in terms of the operator $\hat{A}$ as

$$\langle A\rangle = \langle\Psi|\,\hat{A}\,|\Psi\rangle\,,$$

and so A must be Hermitian. In general, we care about Hermitian operators because they correspond to physical observables.

Since the expectation value corresponds to a physical measurement, it must be real. Therefore,

$$\langle A \rangle = \overline{\langle A \rangle} \iff \langle \Psi | \hat{A} \Psi \rangle = \langle \hat{A} \Psi | \Psi \rangle \,.$$

We can decompose the state of the object into a superposition of the orthonormal states $|1\rangle$ and $|2\rangle$:

$$|\Psi\rangle = \alpha \, |1\rangle + \beta \, |2\rangle \,,$$

where $\alpha, \beta \in \mathbb{C}$ and $\alpha^2 + \beta^2 = 1$. The *probability* of measuring the object in state $|1\rangle$ is given by

$$|\langle 1 | \psi \rangle|^2 = |\alpha|^2 \,,$$

and similarly for state $|2\rangle$.

With this in mind, consider some operator $\hat{U}$ that does not change the probabilities of measuring the object in states $|1\rangle$ and $|2\rangle$. Then $\hat{U}$ must preserve the inner product on the relevant Hilbert space. In particular, we have

$$\langle \Psi | \Psi \rangle = \langle \hat{U} \Psi | \hat{U} \Psi \rangle = \langle \Psi | \hat{U}^\dagger \hat{U} | \Psi \rangle \,,$$

which is only true if $\hat{U}^\dagger \hat{U} = \hat{I}$, where $\hat{I}$ is the identity operator. In other words, we must have $U^\dagger = U^{-1}$. Such operators are called *unitary*. Thus, one can describe unitary operators as probability-preserving transformations.

B.2 Commutator Identities

For linear operators A and B, the commutator is defined as

$$[A, B] = AB - BA.$$

The commutator satisfies the following properties:

$$[A, B] = -[B, A] \tag{B.1}$$

$$[A, -B] = -AB + BA = -[A, B]. \tag{B.2}$$

$$\begin{aligned}[A, B + C] &= A(B + C) - (B + C)A \\ &= AB + AC - BA - CA \\ &= AB - BA + AC - CA \\ &= [A, B] + [A, C].\end{aligned} \tag{B.3}$$

$$\begin{aligned}[A^2, B] &= [AA, B] \\ &= AAB - BAA \\ &= AAB - ABA + ABA - BAA \\ &= A(AB - BA) + (AB - BA)A \\ &= A[A, B] + [A, B]A.\end{aligned} \tag{B.4}$$

$$\begin{aligned}[A, BC] &= ABC - BCA \\ &= ABC - BAC + BAC - BCA \\ &= (AB - BA)C + B(AC - CA).\end{aligned} \tag{B.5}$$

B.3 Commutation relations for SO(3)

This section includes various commutation relations that are used in Chapter 3. The definitions of the operators are given in the relevant section of said chapter.

$$[y, \hat{p}_y] = y\hat{p}_y - \hat{p}_y y = \cancel{y\hat{p}_y} - \overbrace{\left(-i\hbar + \cancel{y\hat{p}_y}\right)}^{\text{product rule}} = i\hbar,$$

$$[\hat{L}_z, \hat{p}_z] = [x\hat{p}_y - y\hat{p}_x, \hat{p}_z] = [x\hat{p}_y, \hat{p}_z] - [y\hat{p}_x, \hat{p}_z] = 0.$$

$$[\hat{L}_z, z] = [x\hat{p}_y - y\hat{p}_x, z] = [x\hat{p}_y, z] - [y\hat{p}_x, z] = 0.$$

$$[\hat{L}_z, \hat{p}_y] = [x\hat{p}_y - y\hat{p}_x, \hat{p}_y] = \cancelto{0}{[x\hat{p}_y, \hat{p}_y]} - [y\hat{p}_x, \hat{p}_y] = -y\cancelto{0}{[\hat{p}_x, \hat{p}_y]} - [y, \hat{p}_y]\hat{p}_x = -i\hbar\hat{p}_x.$$

$$[\hat{L}_z, y] = [x\hat{p}_y - y\hat{p}_x, y] = [x\hat{p}_y, y] - \cancelto{0}{[y\hat{p}_x, y]} = x[\hat{p}_y, y] + \cancelto{0}{[x, y]}\hat{p}_y = -i\hbar x.$$

B.4 Conserved quantities in quantum mechanics

Suppose $\hat{G}$ is an operator on a quantum Hilbert space of states. The quantity $\langle G \rangle$ is conserved if

$$\frac{d\langle G \rangle}{dt} = 0.$$

The time-dependent Schrödinger equation is given by

$$\hat{H}\psi = i\hbar\frac{d\psi}{dt},$$

which implies

$$\frac{d\psi}{dt} = \frac{1}{i\hbar}\hat{H}\psi,$$

where $\hat{H}$ is the Hamiltonian operator. Since the eigenvalues of $\hat{H}$ correspond to the energy of the system, $\hat{H}$ must be Hermitian.

Then if $\hat{G}$ is time-independent we have

$$\begin{aligned}
\frac{d\langle G\rangle}{dt} &= \frac{d}{dt}\langle\psi|\hat{G}|\psi\rangle \\
&= \left\langle \frac{d\psi}{dt}\middle|\hat{G}\middle|\psi\right\rangle + \left\langle\psi\middle|\hat{G}\middle|\frac{d\psi}{dt}\right\rangle + \cancelto{0}{\left\langle\psi\middle|\frac{\partial\hat{G}}{\partial t}\middle|\psi\right\rangle} \\
&= \left\langle \frac{1}{i\hbar}\hat{H}\psi\middle|\hat{G}\middle|\psi\right\rangle + \left\langle\psi\middle|\hat{G}\middle|\frac{1}{i\hbar}\hat{H}\psi\right\rangle \\
&= \frac{i}{\hbar}\left(\langle\hat{H}\psi|\hat{G}|\psi\rangle - \langle\psi|\hat{G}|\hat{H}\psi\rangle\right) \\
&= \frac{i}{\hbar}\left(\langle\psi|\hat{H}^\dagger\hat{G}|\psi\rangle - \langle\psi|\hat{G}\hat{H}|\psi\rangle\right) \\
&= \frac{i}{\hbar}\left(\langle\psi|\hat{H}\hat{G}|\psi\rangle - \langle\psi|\hat{G}\hat{H}|\psi\rangle\right) \text{ because } \hat{H} \text{ is Hermitian} \\
&= \frac{i}{\hbar}\langle\psi|\left(\hat{H}\hat{G} - \hat{G}\hat{H}\right)|\psi\rangle \\
&= \frac{i}{\hbar}\langle\psi|\left[\hat{H},\hat{G}\right]|\psi\rangle = 0 \iff \left[\hat{H},\hat{G}\right] = 0.
\end{aligned} \tag{B.6}$$

This is often referred to as Ehrenfest's theorem [12]. Thus, if $[\hat{H},\hat{G}] = 0$, it follows that

$$\begin{aligned}
\hat{H}\hat{G} - \hat{G}\hat{H} = 0 &\iff \hat{H}\hat{G} = \hat{G}\hat{H} \\
&\iff \hat{G}^{-1}\hat{H}\hat{G} = \hat{H}.
\end{aligned}$$

Therefore, $\hat{G}^{-1}\hat{H}\hat{G}$ and $\hat{H}$ share the same eigenvalues (observables), which is only true if $\hat{H}$ is invariant under $\hat{G}$. If G (corresponding to the operator $\hat{G}$) generates a group of transformations, then $\hat{H}$ is invariant under the group of transformations generated by G. If $\hat{G}$ is unitary, this invariance is often expressed as

$$\hat{G}^\dagger\hat{H}\hat{G} = \hat{H}.$$

Running the argument in reverse, if $\hat{H}$ is invariant under the transformations generated by G, then $[\hat{H}, \hat{G}] = 0$, which by Eqn. B.6 implies that $\langle G \rangle$ is conserved (for time-independent $\hat{G}$).

Appendix C

MULTI-ANYON SYSTEM WITH HARMONIC POTENTIAL

C.1 Deriving the additional Hamiltonian terms

The last term in Eqn. 5.22 is the result of squaring the canonical momentum in Eqn. 5.21. To see this, let's isolate one of the terms. Fix i. Then,

$$\big(\mathbf{p}_i - \mathbf{A}_i(\mathbf{r}_i)\big)^2 = p_i^2 - 2\mathbf{p}_i \cdot \mathbf{A}_i(\mathbf{r}_i) + A_i^2(\mathbf{r}_i).$$

By Eqn. 5.19, we have

$$\mathbf{A}_i{}^2(\mathbf{r}_i) = \left(\alpha \sum_{j\neq i} \frac{-y_{ij}\hat{x} + x_{ij}\hat{y}}{r_{ij}^2}\right)^2 = \alpha^2 \sum_{j,k\neq i} \frac{y_{ij}y_{ik} + x_{ij}x_{ik}}{r_{ij}^2 r_{ik}^2} = \alpha^2 \sum_{j,k\neq i} \frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{ik}}{r_{ij}^2 r_{ik}^2},$$

which is the last term in Eqn. 5.22.

Moreover, the cross term in the expansion of $\big(\mathbf{p}_i - \mathbf{A}_i(\mathbf{r}_i)\big)^2$ is

$$\begin{aligned}
-2\mathbf{p}_i \cdot \mathbf{A}_i(\mathbf{r}_i) &= -2\mathbf{p}_i \cdot \left(\alpha \sum_{j\neq i} \frac{-y_{ij}\hat{x} + x_{ij}\hat{y}}{r_{ij}^2}\right) \\
&= -2\alpha \sum_{j\neq i} \frac{\mathbf{p}_i \cdot (-y_{ij}\hat{x} + x_{ij}\hat{y})}{r_{ij}^2} \\
&= -2\alpha \sum_{j\neq i} \frac{-p_{ix}y_{ij} + p_{iy}x_{ij}}{r_{ij}^2} \\
&= -2\alpha \sum_{j\neq i} \frac{(\mathbf{r}_{ij}\times\mathbf{p}_i)\cdot\hat{z}}{r_{ij}^2}.
\end{aligned}$$

For each j, there is a corresponding term in Eqn. 5.22 with

$$-2\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ji}^2}=-\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ij}^2}+\alpha\frac{(\mathbf{r}_{ij}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ij}^2},$$

where we rewrote one of the two terms to have $\mathbf{r}_{ij}$ instead of $\mathbf{r}_{ji}$. Then, for fixed i and j, the ij- and ji-term can be combined in the following manner:

$$\begin{aligned}
-2\alpha\frac{(\mathbf{r}_{ij}\times\mathbf{p}_i)\cdot\hat{z}}{r_{ji}^2}-2\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ji}^2} &= -\alpha\frac{(\mathbf{r}_{ij}\times\mathbf{p}_i)\cdot\hat{z}}{r_{ij}^2}+\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_i)\cdot\hat{z}}{r_{ji}^2}\\
&\quad+\alpha\frac{(\mathbf{r}_{ij}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ij}^2}-\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_j)\cdot\hat{z}}{r_{ji}^2}\\
&= -\alpha\frac{(\mathbf{r}_{ij}\times(\mathbf{p}_i-\mathbf{p}_j))\cdot\hat{z}}{r_{ij}^2}\\
&\quad-\alpha\frac{(\mathbf{r}_{ji}\times(\mathbf{p}_j-\mathbf{p}_i))\cdot\hat{z}}{r_{ji}^2}\\
&= -\alpha\frac{(\mathbf{r}_{ij}\times\mathbf{p}_{ij})\cdot\hat{z}}{r_{ij}^2}+\alpha\frac{(\mathbf{r}_{ji}\times\mathbf{p}_{ji})\cdot\hat{z}}{r_{ji}^2}\\
&= -\alpha\frac{\ell_{ij}}{r_{ij}^2}-\alpha\frac{\ell_{ji}}{r_{ji}^2}.
\end{aligned}$$

Then, summing over all $i\neq j$ yields the second-to-last term in Eqn. 5.22.