Abstract

Berry Phase of the H_3 Molecular System

Jialiang Wu

In this paper, we lay down the mathematical foundations necessary to describe the H_3 molecular system. We survey the general theories and then apply them to the H_3 case.

- We develop in parallel the mathematical formalism for classical system and quantum system, and point out their distinguishing features. We introduce the concepts of space, logic, observable, and state; we describe the laws of state evolution and updating the state after observation. The formalisms for several quantum systems are constructed: the n-spin system, the spinless H atom and H_2^+ ion, and the real H_2 and H_3 molecular systems.
- The representations of molecular symmetry group is examined. We develop the concept of the symmetries of a molecule and represent them by a group of matrices. We also discuss the representations classification of that group, which is summarized as its characteristic table.
- The shape space and three internal coordinate systems of the H_3 molecule are built up.
- Differential geometric language is introduced to describe the H_3 quantum system: fiber bundles, the special case of Hermitian vector bundles, (on which the connections, parallel translations, and the covariant derivative are defined), and holonomy group. The Berry phase of the H_3 system is shown to be the holonomy in the Hermitian vector bundles of H_3 quantum system.

This mathematical framework provides a foundation to address many unsolved issues of the H_3 system.

Dissertation Director: Dr. Daniel Dix

Berry Phase of the H_3 Molecular System

by

Jialiang Wu

Bachelor of Science in Mathematics

Zhongshan University, 1996

Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in the

Department of Mathematics

University of South Carolina

2003

____
Department of Mathematics
Director of Thesis

____
Department of Mathematics
Second Reader

____
Dean of The Graduate School

Dedication

To Christ Jesus who replaces the emptiness and devastation in my life by faith, hopes, and love lasting forever. In the sense of eternity, our works finds its worth.

# Acknowledgements

My sincere thanks goes to my advisor, Professor Daniel Dix, who faithfully committed his wise guidance, undying patience, and loving discipline to me. I highly appreciate that he helped me to grow independence, persistence, and comprehensive views in scientific research. His love of math and his attitudes toward people have lifelong impacts on me, and his first edition of Ground Rules for Master's Thesis Drafts designed for me will be always on my desk.

I am grateful to Professor George McNulty for many helpful discussions and technical assistance in Latex.

My appreciations go to Professor Ralph Howard. Thank Ralph very much for reading my thesis and offering comments for the thesis improvements.

My wonderful parents, Xinfa Wu and Zhonghao Li own my great gratitude. They unconditionally love me and support me, without whom my happiness and accomplishments would have never happened.

I am indebted to all my good friends for their love and supports. Especially I wish to thank Binglong Chen who teaches me "living" math and shares his beautiful visions, Yuan Ren, my "dear comrade" staying and walking with me through many young-age days in darkness, Jianhua Cao who shared with me every joyful and difficult moments for my three years in Columbia.

Upon God's glories we will celebrate that our encounters in this world are not incidental, but in His good will.

Abstract

In this paper, we lay down the mathematical foundations necessary to describe the H_3 molecular system. We survey the general theories and then apply them to the H_3 case.

- We develop in parallel the mathematical formalism for classical system and quantum system, and point out their distinguishing features. We introduce the concepts of space, logic, observable, and state; we describe the laws of state evolution and updating the state after observation. The formalisms for several quantum systems are constructed: the n-spin system, the spinless H atom and H_2^+ ion, and the real H_2 and H_3 molecular systems.
- The representations of molecular symmetry group is examined. We develop the concept of the symmetries of a molecule and represent them by a group of matrices. We also discuss the representations classification of that group, which is summarized as its characteristic table.
- The shape space and three internal coordinate systems of the H_3 molecule are built up.
- Differential geometric language is introduced to describe the H_3 quantum system: fiber bundles, the special case of Hermitian vector bundles, (on which the connections, parallel translations, and the covariant derivative are defined), and holonomy group. The Berry phase of the H_3 system is shown to be the holonomy in the Hermitian vector bundles of H_3 quantum system.

This mathematical framework provides a foundation to address many unsolved issues of the H_3 system.

Contents

List of Figures

Chapter 1

Introduction

In this thesis we lay down a mathematical framework for the study of the H_3 molecular system. H_3 denotes a system of 3 protons and 3 electrons. The protons are treated as classical particles but the electrons must be treated quantum mechanically. Several topics essential to the understanding of the H_3 system will be addressed in the following chapters.

In chapter 2, we develop in parallel the mathematical formalism for classical systems and quantum systems. The former is a good preparation to understand the latter. Quantum theory is the proper physical language to describe the phenomena in the microscopic world including the H_3 molecular system. We consider the following questions.

- How do we represent the *state* of the quantum system, which contains all the information the observer knows about the system at a given level of theory and approximation?

- How does the state evolve with time?

- How do we represent the physical quantities (called *observables*) associated to the system, such as position, momentum, angular momentum, energy etc. in H_3?

- How do we predict the measurement outcome of some observable for a given quantum state?

- The quantum system will change after observables are measured, therefore how do we take the measured value into account to update the state?

In chapter 3, we apply the quantum formalism to several microscopic systems from simple to complicated: the spin system, the artificially spinless H atom and H_2^+ ion, and the real H_2 and H_3 molecule systems.

In chapter 4, the symmetry information of a molecular system is examined. This information is important because it allows us to simplify the electron distribution when the nuclear configuration is symmetric. The following questions are investigated.

- What are the symmetries of a molecule? How can they be mathematically represented? We will show that the symmetries of a molecule form a group called the *symmetry group*.
- How can the symmetry group be represented by matrices?
- When are two matrix representations of the same symmetry group equivalent?
- What type of representations of a symmetry group are the simplest?
- How can more complex representations be understood in terms of the simple ones?

Moreover, we show that the information of a symmetry group can be organized by a table called its *character table*. We explain the theory through three examples: the isosceles triangle, the equilateral triangle and the tetrahedron.

In chapter 5, we discuss the conformation (i.e. the shape) and the configuration (i.e. the coordinates) of the H_3 system. We first discuss the shapes of the H_3 system in general Euclidean coordinates: how to mathematically represent a shape and how to classify the shapes into the three different categories: the non-collinear, the collinear, and the one-point-coincident. And then we discuss the shape space of the H_3 system using three particular internal coordinate systems which treat the three nuclei in a symmetrical way.

In chapter 6 we introduce the differential geometric concepts: fiber bundles, and the special case of Hermitian vector bundles, on which we define connections, parallel translation, and the covariant derivative. We illustrate these concepts and results in two concrete cases: a surface in $\mathbb{R}^3$ and the H_3 system.

In chapter 7 holonomy groups and Berry phase are considered. Suppose a quantum system undergoes an evolution so that after some time it comes back to its original state. Such an evolution traces out a cycle in quantum mechanical state space. The result of the evolution will be reflected in the phase of the wave function in the form of a geometric phase factor, usually called *Berry phase*. This phase factor can be measured by interfering the initial and the final states. In this chapter, we explain the Berry phase of the H_3 system using the geometric language i.e. in terms of holonomy in a Hermitian vector bundle.

We will see in chapter 6 and chapter 7 that there are several unsolved problems in the H_3 system mentioned.

- Given a nuclear conformation, what is the dimension of the vector space of all ground state electronic wave functions of the H_3 system?
- The ground state electronic energy is a function of the conformation, so where on the shape space is it a smooth function? Where are its singularities, if any?
- For each conformation, the symmetry group of that conformation has a representation in the vector space of all electronic wave functions. How can the representation be understood in terms of irreducible representations?
- Does there exist an atlas of smooth local trivializations for the Hermitian vector bundle of the H_3 system?
- Each fiber E_b should be a direct sum of eigenspaces of $\mathbf{S}^2$ (since $\tilde{H}, S_3, \mathbf{S}^2$ commute). How does this direct sum decomposition depend on the base point b?

Our work in this thesis provides a mathematical foundation for this future study.

Chapter 2

Survey on Mathematical Formalism of Classical Mechanics and Quantum Mechanics

2.1. Preliminaries: Tensor Products

2.1.1. Tensor Product of Two Vector Spaces.

Definition. The tensor product of two $\mathbb{F}$-vector spaces V and W, denoted $V \otimes W$, is a vector space spanned by elements of the form $v \otimes w$, where $v \in V, w \in W$, and such that the following rules are satisfied, for any scalar $\alpha \in \mathbb{F}$,

(1) $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$,

(2) $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$,

(3) $\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w)$

The definition is the same no matter which scalar field $\mathbb{F}$ is used.

Here is another equivalent definition of tensor product using the language of the Universal Property:

Definition. (Tensor Product by Universal Property) If V_1, V_2 are $\mathbb{C}$-vector spaces, then $V_1 \otimes V_2$ is another $\mathbb{C}$-vector space, and $\otimes : V_1 \times V_2 \to V_1 \otimes V_2 : (v_1, v_2) \mapsto v_1 \otimes v_2$ is a $\mathbb{C}$-bilinear mapping with property that if W is an arbitrary $\mathbb{C}$-vector space and $b : V_1 \times V_2 \to W$ is an arbitrary $\mathbb{C}$-bilinear mapping, then there exists one and only one $\mathbb{C}$-linear mapping $\tilde{b} : V_1 \otimes V_2 \to W$ s.t. $b(v_1, v_2) = \tilde{b}(v_1 \otimes v_2)$ for all $v_1 \in V_1$ and $v_2 \in V_2$.

EXAMPLE. Let $V_1 = V_2 = \mathbb{C}^2$. We can define the tensor product in two ways:
definition 1: $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}$, and $V_1 \otimes V_2 = \mathbb{C}^4$.
definition 2: $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \tilde{\otimes} \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) = \begin{pmatrix} ac \\ bc \\ ad \\ bd \end{pmatrix}$, and $V_1 \tilde{\otimes} V_2 = \mathbb{C}^4$.

Both definitions have the desired properties. Also $V_1 \otimes V_2$ is obviously isomorphic to $V_1 \tilde{\otimes} V_2$ by the rule $\alpha : \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$.

DEFINITION. Let V be a vector space. The *vector space symmetric tensor product* $V \otimes_S V$ is defined to be a vector space together with a symmetric bilinear mapping $\otimes_S : V \times V \to V \otimes_S V$ such that if $b : V \times V \to W$ is any bilinear and symmetric mapping, then there is a unique linear mapping $\tilde{b} : V \otimes_S V \to W$ such that $b(v_1, v_2) = \tilde{b}(v_1 \otimes_S v_2)$, for all $v_1, v_2 \in V$.

EXAMPLE. For any $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right), \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \in \mathbb{C}^2$, we define

$$\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes_S \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) := \frac{1}{2}\left[\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) + \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)\right] = \frac{1}{2}\left[\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) + \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)\right] = \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \otimes_S \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right).$$

DEFINITION. Let V be a vector space. The *vector space wedge product* (or *vector space antisymmetric tensor product*) $V \wedge V$ is defined to be a vector space together with a antisymmetric bilinear mapping $\wedge : V \times V \to V \wedge V$ such that if $b : V \times V \to W$ is any bilinear and alternating mapping, then there is a unique linear mapping $\tilde{b} : V \wedge V \to W$ such that $b(v_1, v_2) = \tilde{b}(v_1 \wedge v_2)$ for all $v_1, v_2 \in V$.

EXAMPLE. For any $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right), \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \in \mathbb{C}^2$, we define
$\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \wedge \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) := \frac{1}{2}\left[\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) - \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)\right] = -\frac{1}{2}\left[\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) - \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)\right] = -\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \wedge \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right).$
If we define $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}$, then $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \wedge \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) = \frac{1}{2}\left[\begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} - \begin{pmatrix} ca \\ cb \\ da \\ db \end{pmatrix}\right] = \frac{1}{2}\begin{pmatrix} 0 \\ ad-bc \\ bc-ad \\ 0 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 0 \\ bc-ad \\ ad-bc \\ 0 \end{pmatrix}$. Then $\mathbb{C}^2 \wedge \mathbb{C}^2 := \left\{\begin{pmatrix} 0 \\ b \\ c \\ 0 \end{pmatrix} : b + c = 0\right\} \subset \mathbb{C}^2 \otimes \mathbb{C}^2$.

We can see that vector space symmetric tensor product and wedge products are vector space tensor products having addition structures: symmetry and antisymmetry respectively.

2.1.2. Tensor Product of Two Linear Maps.

Definition. Let $\hat{A}_1 : V_1 \to W_1, \hat{A}_2 : V_2 \to W_2$ be two linear maps between vector spaces, then the tensor product of $\hat{A}_1$ and $\hat{A}_2$ is :

$$\hat{A}_1 \otimes \hat{A}_2 : V_1 \otimes V_2 \to W_1 \otimes W_2 : v_1 \otimes v_2 \mapsto (\hat{A}_1 v_1) \otimes (\hat{A}_2 v_2)$$

extended uniquely by the universal property.

Example. Define $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) := \begin{pmatrix} a\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \\ b\left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}$. For $i = 1, 2$, define linear maps $\hat{M}_i : \mathbb{R}^2 \to \mathbb{R}^2$, where $\hat{M}_i := \left(\begin{smallmatrix} a_i & b_i \\ c_i & d_i \end{smallmatrix}\right)$. The tensor product of these two linear maps is defined to be

$$\begin{aligned} \hat{M}_1 \otimes \hat{M}_2 &= \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \otimes \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \\ &:= \begin{pmatrix} a_1\left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right) & b_1\left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right) \\ c_1\left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right) & d_1\left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right) \end{pmatrix} \\ &= \begin{pmatrix} a_1a_2 & a_1b_2 & b_1a_2 & b_1b_2 \\ a_1c_2 & a_1d_2 & b_1c_2 & b_1d_2 \\ c_1a_2 & c_1b_2 & d_1a_2 & d_1b_2 \\ c_1c_2 & c_1d_2 & d_1c_2 & d_1d_2 \end{pmatrix}. \end{aligned}$$

We need to check that it is well-defined. Let $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right), \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right) \in \mathbb{R}^2$. On one hand, we have

$$\begin{aligned} &\left[\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\right] \otimes \left[\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix}\right] \\ &= \begin{pmatrix} a_1a+b_1b \\ c_1a+d_1b \end{pmatrix} \otimes \begin{pmatrix} a_2c+b_2d \\ c_2c+d_2d \end{pmatrix} \\ &= \begin{pmatrix} (a_1a+b_1b)(a_2c+b_2d) \\ (a_1a+b_1b)(c_2c+d_2d) \\ (c_1a+d_1b)(a_2c+b_2d) \\ (c_1a+d_1b)(c_2c+d_2d) \end{pmatrix}. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} &\left[\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \otimes \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\right] \left[\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix}\right] \\ &= \begin{pmatrix} a_1a_2 & a_1b_2 & b_1a_2 & b_1b_2 \\ a_1c_2 & a_1d_2 & b_1c_2 & b_1d_2 \\ c_1a_2 & c_1b_2 & d_1a_2 & d_1b_2 \\ c_1c_2 & c_1d_2 & d_1c_2 & d_1d_2 \end{pmatrix} \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} \\ &= \begin{pmatrix} (a_1a+b_1b)(a_2c+b_2d) \\ (a_1a+b_1b)(c_2c+d_2d) \\ (c_1a+d_1b)(a_2c+b_2d) \\ (c_1a+d_1b)(c_2c+d_2d) \end{pmatrix}. \end{aligned}$$

Done.

2.1.3. Tensor Product of Two Hilbert Spaces. For the purpose of this thesis a Hilbert space is a complete inner product space over the complex field, and is always *separable*, i.e. there is a countable maximal orthogonal set (see [**19**]).

DEFINITION. (Tensor Product of Two Hilbert Spaces by Universal Property) Assume that $\mathcal{H}_1, \mathcal{H}_2$ are $\mathbb{C}$-Hilbert spaces. Let $\mathcal{H}_1 \otimes \mathcal{H}_2$ be a $\mathbb{C}$-Hilbert space, and $\otimes : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 : (\phi_1, \phi_2) \mapsto \phi_1 \otimes \phi_2$ be a $\mathbb{C}$-bilinear continuous mapping with property that if W is an arbitrary $\mathbb{C}$-Hilbert space and $b : \mathcal{H}_1 \times \mathcal{H}_2 \to W$ is an arbitrary bilinear continuous mapping, then there exist one and only one $\mathbb{C}$-linear continuous mapping $\tilde{b} : \mathcal{H}_1 \otimes \mathcal{H}_2 \to W$ s.t. $b(\phi_1, \phi_2) = \tilde{b}(\phi_1 \otimes \phi_2)$ for all $\phi_1 \in \mathcal{H}_1$ and $\phi_2 \in \mathcal{H}_2$.

Here is one construction of $\mathcal{H}_1 \otimes \mathcal{H}_2$ followed by Reed and Simon [**34**]. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. For each $\phi_1 \in \mathcal{H}_1, \phi_2 \in \mathcal{H}_2$, let $\phi_1 \otimes \phi_2$ denote the conjugate bilinear form which acts on $\mathcal{H}_1 \times \mathcal{H}_2$ by

$$(\phi_1 \otimes \phi_2)(\psi_1, \psi_2) = (\psi_1, \phi_1)_{\mathcal{H}_1} (\psi_2, \phi_2)_{\mathcal{H}_2},$$

where $(\cdot, \cdot)_{\mathcal{H}_1}$ and $(\cdot, \cdot)_{\mathcal{H}_2}$ are the inner products defined on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively.

Let $\mathcal{H}$ be the set of all finite linear combinations of such bilinear forms; we define an inner product $(\cdot, \cdot)_{\mathcal{H}}$ on $\mathcal{H}$ by defining

$$(\phi \otimes \psi, \eta \otimes \mu)_{\mathcal{H}} = (\phi, \eta)_{\mathcal{H}_1} (\psi, \mu)_{\mathcal{H}_2}$$

and extending by linearity to $\mathcal{H}$.

FACT. $(\cdot, \cdot)_{\mathcal{H}}$ is well-defined and positive definite.[**34**]

DEFINITION. We define $\mathcal{H}_1 \otimes \mathcal{H}_2$ to be the completion of $\mathcal{H}$ under the inner product $(\cdot, \cdot)_{\mathcal{H}}$ defined above, which is a subspace of the space of all conjugate bilinear continuous functionals on $\mathcal{H}_1 \times \mathcal{H}_2$. $\mathcal{H}_1 \otimes \mathcal{H}_2$ is called the *tensor product* of $\mathcal{H}_1$ and $\mathcal{H}_2$.

Here we give an alternative construction of the Hilbert space tensor product (see §2.2.2 and §2.2.4 for definitions of σ-algebra and measures)

$$\mathcal{L}^2(X_1, \Sigma_1, \mu_1, \mathbb{C}) \otimes \mathcal{L}^2(X_2, \Sigma_2, \mu_2, \mathbb{C}),$$

where for $i = 1, 2$, X_i are sets, Σ_i are the sigma algebras defined on X_i, μ_i are measures defined on Σ_i, $\mathcal{L}^2(X_i, \Sigma_i, \mu_i, \mathbb{C})$ are the sets of equivalence classes of square integrable complex valued functions defined on X_i.

We define the bilinear mapping

$$\otimes : \mathcal{L}^2(X_1, \Sigma_1, \mu_1, \mathbb{C}) \times \mathcal{L}^2(X_2, \Sigma_2, \mu_2, \mathbb{C}) \to \mathcal{L}^2(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \times \mu_2, \mathbb{C})$$

by the rule $\otimes(\psi_1, \psi_2) = \psi_1 \otimes \psi_2$, where $(\psi_1 \otimes \psi_2)(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$; $\Sigma_1 \otimes \Sigma_2$ is the smallest sigma algebra generated by measurable rectangles [**36**]; $\mu_1 \times \mu_2$ is the product measure of μ_1 and μ_2 (see [**36**] p. 304).

Then for any given continuous bilinear mapping $b : \mathcal{L}^2(X_1, \Sigma_1, \mu_1, \mathbb{C})$ $\times \mathcal{L}^2(X_2, \Sigma_2, \mu_2, \mathbb{C}) \to W$, and for any Hilbert space W, we define the mapping

$$\tilde{b} : \mathcal{L}^2(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \times \mu_2, \mathbb{C}) \to W$$

by the rule $\tilde{b}(\psi_1 \otimes \psi_2) = b(\psi_1, \psi_2)$.

We must check that $\tilde{b}$ is well-defined and linear. To see $\tilde{b}$ is linear in the range of $\otimes$,

$$\begin{aligned}\tilde{b}(\psi_1 \otimes \psi_2 + \alpha\psi_1' \otimes \psi_2) &= \tilde{b}((\psi_1 + \alpha\psi_1') \otimes \psi_2)\\ &= b(\psi_1 + \alpha\psi_1', \psi_2)\\ &= b(\psi_1, \psi_2) + \alpha b(\psi_1', \psi_2)\\ &= \tilde{b}(\psi_1 \otimes \psi_2) + \alpha\tilde{b}(\psi_1' \otimes \psi_2)\end{aligned}$$

□

To see that $\tilde{b}$ is well-defined on the range of $\otimes$, we need to show that $\psi_1 \otimes \psi_2 = \tilde{\psi}_1 \otimes \tilde{\psi}_2 \Rightarrow b(\psi_1, \psi_2) = b(\tilde{\psi}_1, \tilde{\psi}_2)$.

We will show that one of the following three conditions holds:

(1) $\psi_1(x_1) = 0$ and $\tilde{\psi}_1(x_1) = 0$ for μ_1-a.e. $x_1 \in X_1$;

(2) $\psi_2(x_2) = 0$ and $\tilde{\psi}_2(x_2) = 0$ for μ_2-a.e. $x_2 \in X_2$;

(3) There is a constant c such that $\psi_2 = c\tilde{\psi}_2$ μ_2-a.e. on X_2, and $\tilde{\psi}_1 = c\psi_1$ μ_1-a.e. on X_1.

Each of these three conditions imply that $b(\psi_1, \psi_2) = b(\tilde{\psi}_1, \tilde{\psi}_2)$.

Define $S := \{(x_1, x_2) \in X_1 \times X_2 | \psi_1(x_1)\psi_2(x_2) = \tilde{\psi}_1(x_1)\tilde{\psi}_2(x_2)\}$. Then $(\mu_1 \times \mu_2)(S^c) = 0$, where $S^c = (X_1 \times X_2) \setminus S$. $\int_{X_1 \times X_2} \chi_{S^c} \, d\mu_1 d\mu_2 = \int_{X_1} (\int_{X_2} \chi_{S^c} \, d\mu_2) d\mu_1 = 0$.

Define $S_1 = \{x_1 \in X_1 | \int_{X_2} \chi_{S^c}(x_1, x_2) \, d\mu_2(x_2) = 0\}$. Then $\mu_1(X_1 \setminus S_1) = 0$. For all $x_1 \in S_1$, $\chi_{S^c}(x_1, x_2) = 0$ for μ_2-a.e. $x_2 \in X_2$; $\chi_S(x_1, x_2) = 1$ iff $\psi_1(x_1)\psi_2(x_2) = \tilde{\psi}_1(x_1)\tilde{\psi}_2(x_2)$ for μ_2-a.e. $x_2 \in X_2$. $S_2(x_1) = \{x_2 \in X_2 | \psi_1(x_1)\psi_2(x_2) = \tilde{\psi}_1(x_1)\tilde{\psi}_2(x_2)\}$ is meaningful for all $x_1 \in S_1$, and $\mu_2(X_2 \setminus S_2(x_1)) = 0$.

If $\forall x_1 \in S_1$ both $\psi_1(x_1) = 0$ and $\tilde{\psi}_1(x_1) = 0$ then $b(\psi_1, \psi_2) = 0 = b(\tilde{\psi}_1, \psi_2)$ regardless of ψ_2 and $\tilde{\psi}_2$; if not then $\exists x_1^* \in S_1$ such that either $\psi_1(x_1^*) \neq 0$ or $\tilde{\psi}_1(x_1^*) \neq 0$. Interchanging ψ_1 and $\tilde{\psi}_1$ as necessary we have $\psi_1(x_1^*) \neq 0$. We know that $\mu_2(X_2 \setminus S_2(x_1)) = 0$ for all $x_1 \in S_1$. If $\exists x_1^* \in S_1$ such that $\psi_1(x_1^*) \neq 0$ then define $c_1 := \frac{\tilde{\psi}_1(x_1^*)}{\psi_1(x_1^*)}$.

Define $S_2 := \{x_2 \in X_2 | \int_{X_1} \chi_{S^c}(x_1, x_2) \, d\mu_1(x_1) = 0\}$. Then $\mu_2(X_2 \setminus S_2) = 0$ and $S_2(x_1^*) := \{x_2 \in X_2 | \psi_1(x_1^*)\psi_2(x_2) = \tilde{\psi}_1(x_1^*)\tilde{\psi}_2(x_2)\} = \{x_2 \in X_2 | \psi_2(x_2) = c_1\tilde{\psi}_2(x_2)\}$ and $\mu_2(X_2 \setminus S_2(x_1^*)) = 0$. If $x_2 \in S_2 \cap S_2(x_1^*)$ then $\int_{X_1} \chi_{S^c}(x_1, x_2) \, d\mu_1(x_1) = 0$ i.e. $\chi_{S^c}(x_1, x_2) = 0$ μ_1-a.e. $x_1 \in X_1$, $\chi_S(x_1, x_2) = 1$ μ_1-a.e. $x_1 \in X_1$, i.e. $\psi_1(x_1)\psi_2(x_2) = \tilde{\psi}_1(x_1)\tilde{\psi}_2(x_2)$ for μ_1-a.e. $x_1 \in X_1 \Rightarrow c_1\psi_1(x_1)\tilde{\psi}_2(x_2) = \tilde{\psi}_1(x_1)\tilde{\psi}_2(x_2)$ for μ_1-a.e. $x_1 \in X_1 \Rightarrow [c_1\psi_1(x_1) - \tilde{\psi}_1(x_1)]\tilde{\psi}_2(x_2) = 0$ for μ_1-a.e. $x_1 \in X_1$.

If $\forall x_2 \in S_2 \cap S_2(x_1^*)$ $\tilde{\psi}_2(x_2) = 0$ then $\psi_2(x_2) = 0$ for all such x_2 as well and hence $b(\psi_1, \psi_2) = 0 = b(\tilde{\psi}_1, \tilde{\psi}_2)$ regardless of ψ_1 and $\tilde{\psi}_1$. So otherwise $\exists x_2^* \in S_2 \cap S_2(x_1^*)$ such that $\tilde{\psi}_2(x_2^*) \neq 0$, then $S_1(x_2^*) := \{x_1 \in X_1 | c_1\psi_1(x_1) = \tilde{\psi}_1(x_1)\}$ satisfies $\mu_1(X_1 \setminus S_1(x_2^*)) = 0$. Hence $\psi_2(x_2) = c_1\tilde{\psi}_2(x_2)$ for all $x_2 \in S_2(x_1^*)$, $\mu_2(S_2(x_1^*)^c) = 0$ and $\tilde{\psi}_1(x_1) = c_1\psi_2(x_2)$ for all $x_1 \in S_1(x_2^*)$, $\mu_1(S_1(x_2^*)^c) = 0$ as desired.

$b(\psi_1, \psi_2) = b(\psi_1, c\tilde{\psi_2}) = cb(\psi_1, \tilde{\psi_2})$ and $b(\tilde{\psi_1}, \tilde{\psi_2}) = b(c\psi_1, \tilde{\psi_2}) = cb(\psi_1, \tilde{\psi_2})$. So $\tilde{b}$ is well-defined in the range of $\otimes$. $\square$

The fact below shows that $\tilde{b}$ is densely defined.

FACT.

$$\begin{aligned} \psi =&\{\sum_{i=1}^{n} \psi_1^i \otimes \psi_2^i \in \mathcal{L}^2(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \times \mu_2, \mathbb{C}) \\ &|\psi_1^i \in \mathcal{L}^2(X_1, \Sigma_1, \mu_1, \mathbb{C}), \\ &\psi_2^i \in \mathcal{L}^2(X_2, \Sigma_2, \mu_2, \mathbb{C}), i = 1, \cdots, n\} \end{aligned}$$

is dense in $\mathcal{L}^2(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2, \mathbb{C})$.(See [**39**] for the proof.)

Remark: the universal property ensures that

$$\mathcal{L}^2(X_1 \times X_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2, \mathbb{C}) \cong \mathcal{L}^2(X_1, \Sigma_1, \mu_1, \mathbb{C}) \otimes \mathcal{L}^2(X_2, \Sigma_2, \mu_2, \mathbb{C}).$$

since if we take $W = \mathcal{L}^2(X_1) \otimes \mathcal{L}^2(X_2)$, then the induced mapping $\tilde{b}$ is an isomorphism.

EXAMPLE. As we will discuss in §3.2.1, $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ is the Hilbert space of one spinless electron system. $\mathbb{C}^2$ is the Hilbert space of one spin. The tensor product allows us to combine the two attributes. The Hilbert space of one real electron system (i.e. 1 electron system with spin) is represented by:

$$\begin{aligned} \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \otimes \mathbb{C}^2 &= \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \otimes \mathcal{L}^2(\{0,1\}, \mathcal{P}(\{0,1\}), \#, \mathbb{C}) \\ &\cong \mathcal{L}^2(\mathbb{R}^3 \times \{0,1\}, \mathbb{C}), \end{aligned}$$

where $\mathcal{P}(\{0,1\})$, i.e. the power set of $\{0,1\}$, is the sigma algebra of $\{0,1\}$, and $\#$ is the measure of counting.

For $\psi \in \mathcal{L}^2(\mathbb{R}^3 \times \{0,1\}, \mathbb{C})$, we have $\psi(x_1, x_2, x_3, \sigma) \in \mathbb{C}$, for all $x_1, x_2, x_3 \in \mathbb{R}, \sigma \in \{0,1\}$.

Notation: $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})^2 = \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \oplus \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) = \mathcal{L}^2(\mathbb{R}^3 \times \{0,1\}, \mathbb{C})$.

Given a Hilbert space $\mathcal{H}$, define $\hat{\mathcal{I}} : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ as $\psi_1 \otimes \psi_2 \mapsto \psi_2 \otimes \psi_1$ uniquely extended by the universal property. Therefore

(1) $\hat{\mathcal{S}} := \frac{1}{2}(\hat{1}+\hat{\mathcal{I}})$, called the *symmetrization operator*, is the projection of $\mathcal{H}$ into $\mathcal{H}_{+1}$, the eigenspace with eigenvalue $+1$, because $\hat{\mathcal{I}}[\frac{1}{2}(\hat{1}+\hat{\mathcal{I}})\psi] = \frac{1}{2}(\hat{\mathcal{I}}+\hat{\mathcal{I}}^2)\psi = \frac{1}{2}(\hat{\mathcal{I}}+\hat{1})\psi$, i.e. $\frac{1}{2}(\hat{1}+\hat{\mathcal{I}})\psi$ is an eigenfunction of $\hat{\mathcal{I}}$ with eigenvalue $+1$. We define the *Hilbert space symmetric tensor product* $\mathcal{H} \otimes_S \mathcal{H} := \mathcal{H}_{+1}$.

(2) $\hat{\mathcal{A}} := \frac{1}{2}(\hat{1} - \hat{\mathcal{I}})$, called the *anti-symmetrization operator*, is the projection of $\mathcal{H}$ into $\mathcal{H}_{-1}$, the eigenspace with eigenvalue -1, because $\hat{\mathcal{I}}[\frac{1}{2}(\hat{1} - \hat{\mathcal{I}})\psi] = \frac{1}{2}(\hat{\mathcal{I}} - \hat{\mathcal{I}}^2)\psi = -\frac{1}{2}(\hat{1} - \hat{\mathcal{I}})\psi$, i.e. $\frac{1}{2}(\hat{1} - \hat{\mathcal{I}})\psi$ is an eigenfunction of $\hat{\mathcal{I}}$ with eigenvalue -1. We define the *Hilbert space wedge product* (or Hilbert space antisymmetric tensor product) $\mathcal{H} \wedge \mathcal{H} := \mathcal{H}_{-1}$.

Therefore, $\mathcal{H} \otimes_S \mathcal{H}$ and $\mathcal{H} \wedge \mathcal{H}$ are subspaces of $\mathcal{H} \otimes \mathcal{H}$.

2.2. Mathematical Formalism of Classical Statistical Mechanics

To understand the mathematical formalism of quantum mechanics, we will first discuss the corresponding concepts in classical mechanics.

2.2.1. Phase Space.

Definition. The *phase space* determines all the constituents of the system, and all their possible configurations and dynamical behaviors. The *phase space* X of a classical mechanical system with n degrees of freedom is a $2n$ dimensional manifold, with local coordinates $q_1, \cdots, q_n, p_1, \cdots, p_n$, where $(q_1, \cdots, q_n) = \mathbf{q}$ represent the generalized position coordinates and $(p_1, \cdots, p_n) = \mathbf{p}$ represent the generalized momentum coordinates.

Example. For a particle in $\mathbb{R}^3$, the phase space is $\mathbb{R}^6 = \{(q_1, q_2, q_3, p_1, p_2, p_3) | q_i, p_i \in \mathbb{R}, i = 1, 2, 3\}$, where $\mathbf{q} = (q_1, q_2, q_3)$ is the position of the particle, $\mathbf{p} = (p_1, p_2, p_3)$ is the momentum of the particle.

2.2.2. Logic.

Definition. Let X be a set. Then a *σ-algebra* Σ is a collection of subsets of X such that the following hold:

(1) The empty set is in Σ;

(2) If A is in Σ, then so is $X \setminus A$;

(3) If A_n is a sequence of elements of Σ, then the union of the A_ns is in Σ.

If Σ is a σ-algebra and A is a subset of X, then A is called *measurable* if A is a member of Σ. Measurable sets are also called *events*. If S is any collection of subsets of X, then we can always find a σ-algebra containing S, namely the power set of X. By taking the intersection of all σ-algebras containing S, we obtain the smallest such σ-algebra, which is called *the σ-algebra generated by S*.

Definition. The *Borel σ-algebra* $\mathcal{B}(X)$ is defined to be the σ-algebra generated by the collection of open sets (or equivalently, by the closed sets) of the topological space X. A *Borel set* is an element of the Borel σ-algebra.

Definition. The *logic* of a classical mechanical system is $\Sigma = \mathcal{B}(X)$, the Borel σ-algebra on the phase space X.

The events of the logic are in an idealized sense definable and testable in terms of experimentally relevant quantities such as the coordinates $q_1, \cdots, q_n, p_1, \cdots, p_n$. $A \in \Sigma$ if there is an idealized decision algorithm for whether $s \in A$ involving only answering questions about the values $f(s)$ of various continuous functions $f : X \to \mathbb{R}$.

For $A, B \in \Sigma$, we define $A \leq B$ iff $A \subset B$. We define $A^\perp := X - A$. We define $A \perp B$ iff $A \cap B = \emptyset$.

2.2.3. Observables. A physical quantity relative to this system is called an *observable*, defined as a function $f : X \to \mathbb{R}$ measurable with respect to the σ-algebra Σ, i.e. $f^{-1}(R) \in \Sigma$, $\forall R \in \mathcal{B}(\mathbb{R})$.

For instance, if the system is that of a single particle of mass m which moves in $\mathbb{R}^3$ under the influence of a potential force, then $n = 3, X = \mathbb{R}^6, \Sigma = \mathcal{B}(\mathbb{R}^6)$, and an important observable is the Hamiltonian $\hat{H}$, which is given by

$$\hat{H}(q_1, q_2, q_3, p_1, p_2, p_3) = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V(q_1, q_2, q_3).$$

If $(\mathbf{q}, \mathbf{p}) = (q_1, q_2, q_3, p_1, p_2, p_3) \in \mathbb{R}^6$, the function $(\mathbf{q}, \mathbf{p}) \to \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2)$ is the kinetic energy observable of the particle. The function $(\mathbf{q}, \mathbf{p}) \to V(q_1, q_2, q_3)$ is the potential energy observable of the particle. These are also important observables.

Definition. Let be u, v be any two smooth functions of the variables $(\mathbf{q}, \mathbf{p})$. Then the expression

$$\{u, v\} = \sum_{i=1}^{n} (\frac{\partial u}{\partial p_i}\frac{\partial v}{\partial q_i} - \frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i})$$

is called the *Poisson bracket* of u and v.

Fact. The collection $\mathcal{P}$ of all smooth observables have the following properties, for any $A, B, C \in \mathcal{P}$

(1) distributive law (Leibnitz rule): $\{A, BC\} = \{A, B\}C + B\{A, C\}$;
(2) antisymmetric law: $\{A, B\} = -\{B, A\}$;
(3) linearity: $\{A, \beta_1 B_1 + \beta_2 B_2\} = \beta_1\{A, B_1\} + \beta_2\{A, B_2\}$;
(4) Jacobi identity: $\{A, \{B, C\}\} = \{\{A, B\}, C\} + \{B, \{A, C\}\}$.

2.2.4. State.

Definition. A *measure* is a map $m : \Sigma \to \mathbb{R}^+ \cup \{0\}$ such that $m(\emptyset) = 0$ and, if A_n is a countable sequence in Σ and the A_n are pairwise disjoint, then

$$m(\cup_n A_n) = \sum_n m(A_n).$$

If, in addition, $m(X) = 1$ for X a measure space, then m is said to be a *probability measure.*

Definition. A *state* in classical statistical mechanics is defined as a probability measure on the phase space:

$$\mu : \Sigma \longrightarrow [0,1].$$

The *pure states* are the extreme points of the convex set of all states, which are represented by the Dirac delta measures $\delta_{(\mathbf{q},\mathbf{p})}(S) = \begin{cases} 1, & \text{if } (\mathbf{q},\mathbf{p}) \in S \\ 0, & \text{otherwise.} \end{cases}$.

The state represents the observer's partial knowledge of the phase point $(\mathbf{q},\mathbf{p})$ and/or describes an ensemble of identically prepared systems.

Definition. A measure ν is absolutely continuous with respect to another measure μ, denoted as $\nu << \mu$, if for every set E with $\mu(E) = 0$, we have $\nu(E) = 0$.

This makes sense as long as μ is a positive measure, such as Lebesgue measure, but ν can be any measure, possibly a complex measure.

By the Radon-Nikodym theorem [**36**], $\nu << \mu$ implies that

$$\nu(E) = \int_E f \, d\mu, \; \forall E \in \Sigma$$

where the integral is the Lebesgue integral, for some integrable function f. The function f, uniquely determined μ a.e. on X by μ and ν, is like a derivative, and is called the *Radon-Nikodym derivative* $\frac{d\nu}{d\mu}$.

For an example of an absolutely continuous state, i.e. a state which is absolutely continuous with respect to Lebesgue measure $d^n\mathbf{q}d^n\mathbf{p} = dq_1 \cdots dq_n dp_1, \cdots dp_n$, we define the density function $\rho(\mathbf{q},\mathbf{p}) = \frac{e^{-\beta \hat{H}(\mathbf{q},\mathbf{p})}}{Z(\beta)}$, where $\beta > 0$ is a parameter inversely proportional to the temperature, and $Z(\beta) = \int_{\mathbb{R}^{2n}} e^{-\beta \hat{H}(\mathbf{q},\mathbf{p})} \, d^n\mathbf{q}d^n\mathbf{p}$ is called the partition function. ρ is a nonnegative real valued function, and $\int_{\mathbb{R}^{2n}} \rho \, d^n\mathbf{q}d^n\mathbf{p} = 1$. The state $d\mu = \rho(\mathbf{q},\mathbf{p}) \, d^n\mathbf{q}d^n\mathbf{p}$ is usually called the *canonical ensemble.*

2.2.5. State Evolution. The law of evolution in time of the state of the system is specified by a smooth function $\hat{H} : X \to \mathbb{R}$, called the *Hamiltonian*. Let $(\mathbf{q}(t), \mathbf{p}(t)) = (q_1(t), \cdots, q_n(t), p_1(t), \cdots, p_n(t))$ be a dynamical trajectory of the system; such a trajectory satisfies the Hamilton's equations of motion :

$$\begin{cases} \frac{d\mathbf{q}}{dt} = \nabla_{\mathbf{p}} \hat{H} \\ \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{q}} \hat{H} \end{cases}$$

where $\nabla_{\mathbf{q}} := \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \vdots \\ \frac{\partial}{\partial q_n} \end{pmatrix}$, $\nabla_{\mathbf{p}} := \begin{pmatrix} \frac{\partial}{\partial p_1} \\ \vdots \\ \frac{\partial}{\partial p_n} \end{pmatrix}$.

Let $\phi_t(\mathbf{q}_0, \mathbf{p}_0) := (\mathbf{q}(t), \mathbf{p}(t))$ be the solution of these classical equations of motion, where $(\mathbf{q}(0), \mathbf{p}(0)) = (\mathbf{q}_0, \mathbf{p}_0)$. Denote the initial state as μ_0, and the state after time t as μ_t. Let $S \in \mathcal{B}(X)$, then define

$$\mu_t(S) = \mu(\phi_{-t}(S)).$$

For a pure initial state, i.e. if $\mu_0 = \delta_{(\mathbf{q},\mathbf{p})}$, then

$$\mu_t(S) = \delta_{\phi_t(\mathbf{q},\mathbf{p})}(S).$$

For a given absolutely continuous initial state $d\mu_0 = \rho_0 d^n\mathbf{q} d^n\mathbf{p}$, the density function ρ_t of the state $d\mu_t = \rho_t\, d^n\mathbf{q} d^n\mathbf{p}$ at time t satisfies the Liouville's equation [**9**] :

$$\frac{d\rho_t}{dt} = -\{\hat{H}, \rho\} := -(\nabla_{\mathbf{p}} \hat{H} \cdot \nabla_{\mathbf{q}} \rho - \nabla_{\mathbf{p}} \rho \cdot \nabla_{\mathbf{q}} \hat{H}).$$

Example. For a 1-dimensional harmonic oscillator, the Hamiltonian is given by $\hat{H}(q,p) = \frac{p^2}{2m} + \frac{1}{2}kq^2 = \frac{p^2}{(\sqrt{2m})^2} + \frac{q^2}{(\sqrt{\frac{2}{k}})^2} = E$, where m is the mass of the system, k is a constant, q is the displacement, p is the momentum, $\frac{p^2}{2m}$ represents the kinetic energy, $\frac{1}{2}kq^2$ represents the potential energy, E is the total energy of the system. For different E values, the level curves of the Hamiltonian correspond to different ellipses having the same major axis and minor axis. Then $\nabla_q \hat{H} = \hat{H}_q = kq, \nabla_p \hat{H} = \hat{H}_p = p/m$ and hence the Hamilton's equation is

$$\begin{cases} \frac{dq}{dt} = \hat{H}_p = p/m, \\ \frac{dp}{dt} = -\hat{H}_q = -kq. \end{cases}$$

Given the initial state (q_0, p_0), the state $(\tilde{q}(t; q_0, p_0), \tilde{p}(t; q_0, p_0))$ at time t can be solved from the above equations as $q = C\cos(\omega t + \phi), p = -mC\omega\sin(\omega t + \phi)$, where $\omega = \sqrt{\frac{k}{m}}$; C, ϕ are constants calculated as follows.

$$\begin{aligned}
&q_0 = C\cos\phi,\ p_0 = -Cm\omega\sin\phi \Rightarrow \cos\phi = q_0/C,\ \sin\phi = -\frac{p_0}{Cm\omega};\\
&\hat{H}(q_0, p_0) = \frac{p_0^2}{2m} + \frac{1}{2}kq_0^2 = \frac{C^2m^2\omega^2}{2m}\sin^2\phi + \frac{1}{2}kC^2\cos^2\phi = kC^2/2.\\
\Rightarrow &C = \sqrt{\frac{2\hat{H}(q_0, p_0)}{k}},\\
&\phi = \arg(q_0/C - ip_0/(Cm\omega)) = \arg(q_0 - ip_0/(m\omega)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&q = C\cos(\omega t + \phi) = \sqrt{\frac{2\hat{H}(q_0, p_0)}{k}}\cos(\sqrt{\frac{k}{m}}t + \arg(q_0 - ip_0/(m\omega))) = \tilde{q}(t; q_0, p_0),\\
&p = -Cm\omega\sin(\omega t + \phi) = -\sqrt{2\hat{H}(q_0, p_0)m}\sin(\sqrt{\frac{k}{m}}t + \arg(q_0 - ip_0/(m\omega)))\\
&= \tilde{p}(t; q_0, p_0).
\end{aligned}$$

Suppose the initial density function $\rho_0(q, p) = \rho(q, p, 0)$ is given. The Liouville's equation is $\frac{\partial\rho}{\partial t} = -\frac{p}{m}\frac{\partial\rho}{\partial q} + kq\frac{\partial\rho}{\partial p}$. We claim that the solution for the Liouville's equation is $\rho(q, p, t) = \rho(\tilde{q}(-t; q, p), \tilde{p}(-t; q, p), 0) = \rho_0(\tilde{q}(-t; q, p), \tilde{p}(-t; q, p))$.

To see,

$$\begin{aligned}
\rho(\tilde{q}(t; q, p), \tilde{p}(t; q, p), t) &= \rho_0(q, p)\\
0 = \frac{\partial}{\partial t}\rho_0(q, p) &= \frac{\partial}{\partial t}\rho(\tilde{q}(t; q, p), \tilde{p}(t; q, p), t)\\
&= \frac{\partial\rho}{\partial q}\cdot\frac{\partial\tilde{q}}{\partial t} + \frac{\partial\rho}{\partial p}\frac{\partial\tilde{p}}{\partial t} + \frac{\partial\rho}{\partial t}
\end{aligned}$$

$$= \frac{\partial \rho}{\partial q} \cdot \frac{\tilde{p}}{m} + \frac{\partial \rho}{\partial p} \cdot (-k)\tilde{q} + \frac{\partial \rho}{\partial t}.$$

□

2.2.6. Predicting Measurement Outcomes. Suppose the observable $f : X \to \mathbb{R}$ is to be measured and we wish to predict the distribution of the measured values of this observable when the system is in the state $\mu\colon \Sigma \to [0,1]$. If we suppose the measurement is *noiseless* then the probability distribution of the values of f is the so-called *marginal distribution* $\nu_f\colon \mathcal{B}(\mathbb{R}) \to [0,1]$, defined by the rule: $\nu_f(R) = \mu(f^{-1}(R))$ for all $R \in \mathcal{B}(\mathbb{R})$.

Example. X is a product space $\mathbb{R}^2$, with coordinates (q,p), $f(q,p) = q$, $d\mu = \rho(q,p)\, dpdq$ then

$$d\nu_f(q) = \left(\int_{-\infty}^{\infty} \rho(q,p)\, dp \right) dq,$$

where $\int_{-\infty}^{\infty} \rho(q,p)\, dp$ is called the *marginal density function*. We will use the notation $d\nu_f(q)$ and $\nu_f(dq)$ interchangeably.

Realistic measurements all have some noise, which means that instead of measuring $f(x)$ the measurement apparatus introduces small perturbations, which we model using a normally distributed random variable with mean zero and small variance ϵ^2. Define

$$\alpha_\epsilon(y) = \frac{1}{(2\pi\epsilon^2)^{1/4}} \exp[-\frac{y^2}{4\epsilon^2}].$$

Then $\alpha_\epsilon(y)^2$ is the probability density function for the noise perturbations. The predicted probability distribution of the noisy measured values is $\nu_f * \alpha_\epsilon^2$. Thus the probability that the noisy measured values will be in the set $R \in \mathcal{B}(\mathbb{R})$ is given by

$$\begin{aligned}(\nu_f * \alpha_\epsilon^2)(R) &= \int_{-\infty}^{\infty} \chi_R(z) \int_{-\infty}^{\infty} \alpha_\epsilon(z-y)^2\, \nu_f(\, dy)\, dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi_R(z) \alpha_\epsilon(z-y)^2\, dz\, \nu_f(\, dy) \\ &= \int_{-\infty}^{\infty} (\chi_R * \alpha_\epsilon^2)(y)\, \nu_f(\, dy).\end{aligned}$$

In the limit as $\epsilon \to 0^+$ this expression tends to $\nu_f(R)$ as expected.

2.2.7. Updating the state after the measurement. First let us suppose a noiseless measurement of the observable f has been conducted while the system is in the state μ, and the outcome of the measurement is an observed probability distribution $\tilde{\nu}$ of the measured values of f. This observed distribution $\tilde{\nu}$ may be different from the predicted distribution ν_f, and our problem is to know how to take this new information into account. If there is some $R \in \mathcal{B}(\mathbb{R})$ such that $\nu_f(R) = 0$ and $\tilde{\nu}(R) > 0$ then there is a genuine conflict between theoretical prediction and observation. Let us assume the contrary, namely that $\tilde{\nu} << \nu_f$.

The experimentally obtained information $\tilde{\nu}$ concerns the possible values $y = f(x)$ of the observable, if the phase point x is in X. However, the experiment yields no information about which $x \in f^{-1}(\{y\})$ gives rise to a particular y value. The state μ does however provide such information by means of the *fibre measures* μ_y, which will altogether satisfy

$$(*) \qquad \mu(S) = \int_{-\infty}^{\infty} \mu_y(S)\, d\nu_f(y)$$

for all $S \in \Sigma$. This can be understood as follows. Suppose $S \in \Sigma$ and define $\nu_{f,S} \colon \mathcal{B}(\mathbb{R}) \to [0, \mu(S)]$ by the rule $\nu_{f,S}(R) = \mu(S \cap f^{-1}(R))$ for all $R \in \mathcal{B}(\mathbb{R})$. Clearly $\nu_{f,S}$ is a positive measure and $\nu_{f,S} << \nu_f$. Therefore by the Radon-Nikodym theorem [**36**]

$$\mu(S) = \nu_{f,S}(\mathbb{R}) = \int_{-\infty}^{\infty} d\nu_{f,S}(y) = \int_{-\infty}^{\infty} \frac{d\nu_{f,S}}{d\nu_f}(y)\, d\nu_f(y).$$

This suggests that $\mu_y(S) = \frac{d\nu_{f,S}}{d\nu_f}(y)$ except for y in some ν_f-null set (depending on S). The fact that there exists a universal ν_f-null set $N \in \mathcal{B}(\mathbb{R})$ such that for all $y \in \mathbb{R} \setminus N$ the fibre probability measure μ_y exists and satisfies $\mu_y(S) = \frac{d\nu_{f,S}}{d\nu_f}(y)$ *for all* $S \in \Sigma$ is a nontrivial theorem in measure theory, proved for example in [**31**] (Parthasarathy). μ_y is also called a *conditional probability*. It is concentrated on on the set $f^{-1}(\{y\})$ in the sense that $\mu_y(X \setminus f^{-1}(\{y\})) = 0$.

No experimental information gives us any reason to modify the fibre measures μ_y, so we define the *updated state* $\tilde{\mu}$ by the rule:

$$\tilde{\mu}(S) = \int_{-\infty}^{\infty} \mu_y(S)\, d\tilde{\nu}(y)$$

for all $S \in \Sigma$. This makes sense since $\mu_y(S)$ is defined for all $y \in \mathbb{R} \setminus N$, and $\tilde{\nu}(N) = 0$. Hence we essentially replace ν_f in (*) by $\tilde{\nu}$. It is not difficult to check that $\tilde{\mu}(f^{-1}(R)) = \tilde{\nu}(R)$ for all $R \in \mathcal{B}(\mathbb{R})$, and $\tilde{\mu}_y = \mu_y$ for all $y \in \mathbb{R} \setminus N$. The updated state thus encodes all the information the observer possesses about the system immediately after the measurement.

If the measurement is noisy, this does not affect the formula for the updated state, since in classical mechanics we do not suppose that the noise in the measurement apparatus actually perturbs the system, so as to modify the fibre measures μ_y. However in a noisy measurement one would expect the observed distribution $\tilde{\nu}$ to be "blurred" when compared to the distribution of noiselessly measured values. Thus the updated state represents what the observer knows about the system, not what "really is". Because of this it is not surprising that the state might need to be updated after a measurement.

Example. Let $X = \{1, 2, 3, 4\} \times \{1, 2, 3, 4, 5, 6\} = \{(1, 1), (1, 2), \cdots, (1, 6), (2, 1), (2, 2), \cdots, (2, 6), \cdots, (4, 1), (4, 2), \cdots, (4, 6)\}$. Let us see a formalistic construction on this classical system:

- define the logic to be $\Sigma = \mathcal{P}(X)$, the power set of X;
- define an observable to be $f : X \to \{1, \cdots, 6\} : (i, j) \mapsto j$;
- assume that the atomic probability at $(i, j) \in X$ is $P_{i,j}$, where $0 \leq P_{i,j} \leq 1$, $\sum_{i=1}^{4} \sum_{j=1}^{6} P_{i,j} = 1$;
- the state before measurement is $\mu(S) = \sum_{(i,j) \in S} P_{i,j}\ \forall S \in \Sigma$;

- for $R \subset \{1,2,3,4,5,6\}$, the predicted distribution of the measured values is

$$\nu_f(R) = \mu(f^{-1}(R)) = \sum_{(i,j)\in f^{-1}(R)} P_{i,j} = \sum_{f(i,j)\in R} P_{i,j} = \sum_{i=1}^{4} \sum_{j=1, j\in R}^{6} P_{i,j}$$

$$= \sum_{j=1, j\in R}^{6} (\sum_{i=1}^{4} P_{i,j}) = \sum_{j=1,j\in R}^{6} P_j^f,$$

where $P_j^f := \sum_{i=1}^{4} P_{i,j}$ is the marginal probability;

$$\nu_{f,S}(R) = \mu(f^{-1}(R) \cap S) = \sum_{(i,j)\in f^{-1}(R)\cap S} P_{i,j} = \sum_{j=1,j\in R}^{6} (\sum_{i=1;(i,j)\in S}^{4} P_{i,j})$$

$$= \sum_{j=1,j\in R}^{6} \frac{\sum_{i=1;(i,j)\in S}^{4} P_{i,j}}{\sum_{i=1}^{4} P_{i,j}} \sum_{i=1}^{4} P_{i,j} = \sum_{j=1,j\in R}^{6} \frac{\sum_{i=1;(i,j)\in S}^{4} P_{i,j}}{P_j^f} P_j^f$$

$$= \sum_{j=1,j\in R}^{6} \mu_j(S) P_j^f;$$

define $\mu_j(S) = \dfrac{\sum_{i=1;(i,j)\in S}^{4} P_{i,j}}{P_j^f}$;

- assume that the observed distribution of measured value at $j \in \{1, \cdots, 6\}$ is $\tilde{\nu}(\{j\}) = \tilde{P}_j$, where $0 \le \tilde{P}_j \le 1, \sum_{j=1}^{6} \tilde{P}_j = 1$ and $\tilde{P}_j = 0$ whenever $\sum_{i=1}^{4} P_{ij} = 0$ (the absolute continuity condition). Then the observed distribution of measured value at $R \subset \{1, \cdots, 6\}$ is $\tilde{\nu}(R) = \sum_{j=1,j\in R}^{6} \tilde{P}_j$. Then the updated state defined for any $S \in \mathcal{P}(X)$ is $\tilde{\mu}(S) = \sum_{j=1}^{6} \frac{\nu_{f,S}(\{j\})}{\nu_f(\{j\})} \tilde{\nu}(\{j\}) = \sum_{j=1}^{6} \frac{\sum_{i \text{ s.t. } (i,j)\in S} P_{i,j}}{P_j^f} \tilde{P}_j = \sum_{(i,j)\in S} \frac{P_{ij}}{P_j^f} \tilde{P}_j = \sum_{(i,j)\in S} \tilde{P}_{ij}$, where $\tilde{P}_{i,j} := \tilde{\mu}(\{(i,j)\})$ $= \frac{P_{i,j}}{P_j^f} \tilde{P}_j$ is updated probability distribution. For instance, if $S = \{(1,2), (2,2), (3,2), (3,3), (3,4)\}$, then $\tilde{\mu}(S) = \frac{P_{1,2}+P_{2,2}+P_{3,2}}{\sum_{i=1}^{4} P_{i,2}} \tilde{P}_2 + \frac{P_{3,3}}{\sum_{i=1}^{4} P_{i,3}} \tilde{P}_3 + \frac{P_{3,4}}{\sum_{i=1}^{4} P_{i,4}} \tilde{P}_4 = \frac{P_{1,2}+P_{2,2}+P_{3,2}}{P_2^f} \tilde{P}_2 + \frac{P_{3,3}}{P_3^f} \tilde{P}_3 + \frac{P_{3,4}}{P_4^f} \tilde{P}_4$.

Example. Let $X = \mathbb{R}^2$, with coordinates (x_1, x_2), and we are suppose to measure $f(x_1, x_2) = x_2$. Let $\rho(x_2)$ be the predicted density function of x_2, and $\tilde{\rho}(x_2)$ the observed density function of x_2. Let $\eta(x_1, x_2)$ be the density function of the state μ

prior to the measurement, and $\tilde{\eta}$ the density function of the updated state $\tilde{\mu}$. Then

$$
\begin{aligned}
d\mu = \eta(x_1,x_2)\,dx_1dx_2,\ d\tilde{\mu} &= \tilde{\eta}(x_1,x_2)\,dx_1dx_2,\\
\rho(x_2) &= \int_{-\infty}^{\infty} \eta(x_1,x_2)\,dx_1,\ d\nu_f = \rho(x_2)\,dx_2,\\
\tilde{\rho}(x_2) &= \int_{-\infty}^{\infty} \tilde{\eta}(x_1,x_2)\,dx_1,\ d\tilde{\nu}_f = \tilde{\rho}(x_2)\,dx_2,\\
\frac{d\nu_{f,S}}{d\nu_f}(y_0) &= \lim_{\epsilon\to 0^+} \frac{\nu_{f,S}([y_0-\epsilon, y_0+\epsilon])}{\nu_f([y_0-\epsilon, y_0+\epsilon])}\\
&= \lim_{\epsilon\to 0^+} \frac{\int_{y_0-\epsilon}^{y_0+\epsilon}\int_{-\infty}^{+\infty} \chi_S(x_1,x_2)\eta(x_1,x_2)dx_1dx_2}{\int_{y_0-\epsilon}^{y_0+\epsilon} \rho(x_2)dx_2}\\
&= \lim_{\epsilon\to 0^+} \frac{\int_{-\infty}^{+\infty} \chi_S(x_1,y_0)\eta(x_1,y_0)\,dx_1 2\epsilon}{\rho(y_0)2\epsilon}\\
&= \frac{\int_{-\infty}^{+\infty} \chi_S(x_1,y_0)\eta(x_1,y_0)dx_1}{\rho(y_0)}.
\end{aligned}
$$

Then we find the updated state to be

$$
\begin{aligned}
\tilde{\mu}(S) &= \int_{\mathbb{R}} \frac{d\nu_{f,S}}{d\nu_f}\,d\tilde{\nu}_f\\
&= \int_{\mathbb{R}} \frac{\int_{-\infty}^{+\infty} \chi_S(x_1,x_2)\eta(x_1,x_2)\,dx_1}{\rho(x_2)} \tilde{\rho}(x_2)dx_2\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \chi_S(x_1,x_2)\eta(x_1,x_2)\frac{\tilde{\rho}(x_2)}{\rho(x_2)}dx_1dx_2.
\end{aligned}
$$

By the definition of $\tilde{\eta}$, for all $S \in \mathcal{B}(\mathbb{R}^2)$

$$
\tilde{\mu}(S) = \int_{\mathbb{R}}\int_{\mathbb{R}} \chi_S(x_1,x_2)\tilde{\eta}(x_1,x_2)\,dx_1dx_2,
$$

and therefore we have the relation $\eta(x_1,x_2)\frac{\tilde{\rho}(x_2)}{\rho(x_2)} = \tilde{\eta}(x_1,x_2)$.

2.3. Mathematical Formalism of Quantum Mechanics

2.3.1. Hilbert Space. Each system is associated to a seperable Hilbert space $\mathcal{H}$, which is not strictly analogous to the classical phase space, but plays roughly the same role in quantum mechanics as the phase space plays in classical mechanics. It is difficult to give a systematic theory of the Hilbert spaces for each particle type, so we will assume that they are known and focus on deriving the Hilbert space for a system of many particles.

Let $\mathcal{A}$ be the set of all particle types, such as 'electron', 'proton', 'neutron', 'Helium nucleus', etc. Also assume $\mathcal{A} = \mathcal{A}_{\text{boson}} \cup \mathcal{A}_{\text{fermion}}$. Assume that $\mathcal{H}_\alpha$ is the Hilbert space for a single particle of type $\alpha \in \mathcal{A}$. Let $\mathcal{N}$ be the set of all particles in the system. Then $\mathcal{N} = \cup_{\alpha \in \mathcal{A}} \mathcal{N}_\alpha$, where $\mathcal{N}_{\alpha_1} \cap \mathcal{N}_{\alpha_2} = \emptyset$ if $\alpha_1 \neq \alpha_2$, where $\mathcal{N}_\alpha$ is the set of all particles in $\mathcal{N}$ of type α.

Then we have the following rules:

(1) Suppose $\alpha \in \mathcal{A}_{\text{fermion}}$. Then the Hilbert space for the system of particles $\mathcal{N}_\alpha$ is $\mathcal{H}_{\mathcal{N}_\alpha} = \mathcal{H}_\alpha \overset{1}{\wedge} \cdots \overset{|\mathcal{N}_\alpha|-1}{\wedge} \mathcal{H}_\alpha$, where $\wedge$ is the Hilbert space wedge product.

(2) Suppose $\alpha \in \mathcal{A}_{\text{boson}}$, then the Hilbert space for the system of particles $\mathcal{N}_\alpha$ is $\mathcal{H}_{\mathcal{N}_\alpha} = \mathcal{H}_\alpha \overset{1}{\otimes}_S \cdots \overset{|\mathcal{N}_\alpha|-1}{\otimes}_S \mathcal{H}_\alpha$, where $\otimes_S$ is the Hilbert space symmetric tensor product.

(3) The Hilbert space for the entire system of particles $\mathcal{N}$ is $\mathcal{H}_{\mathcal{N}} = \mathcal{H}_{\mathcal{N}_{\alpha_1}} \otimes \cdots \otimes \mathcal{H}_{\mathcal{N}_{\alpha_k}}$, where $\mathcal{A} = \{\alpha_1, \cdots, \alpha_k\}$ and $\otimes$ is ordinary Hilbert space tensor product.

2.3.2. Logic.

Definition. A *bounded linear operator* $T : V \to W$ between two Banach spaces V, W satisfies the inequality

$$\|T(v)\|_W \leq C\|v\|_V, \forall v \in V,$$

where C is a constant independent of v.

Fact. For any bounded linear functional $l : \mathcal{H} \to \mathbb{R}$, where $\mathcal{H}$ is a Hilbert space, there is an unique $z \in \mathcal{H}$ such that $l(y) = (z, y)$ for all $y \in \mathcal{H}$, where (z, y) is the inner product of z and y. We write $l = z^\dagger$.

For any fixed $x \in \mathcal{H}$, a bounded linear operator P on Hilbert space $\mathcal{H}$ defines a linear functional $l(y) := (x, Py), \forall y \in \mathcal{H}$. By the above fact, for any fixed $x \in \mathcal{H}$, there is an unique $z \in \mathcal{H}$ such that $(x, Py) = l(y) = (z, y)$ for all $y \in \mathcal{H}$. We define the Hermitian conjugate (or adjoint) of P, denoted as $P^\dagger$, by the rule $P^\dagger x = z$. It is also a bounded linear operator on Hilbert space.

Definition. Let $\mathcal{H}$ be a Hilbert space. We define the set of orthogonal projection operators as $\mathcal{L} = \{\hat{P} : \mathcal{H} \to \mathcal{H} | \hat{P} \text{ is linear and bounded, } \hat{P} = \hat{P}^\dagger = \hat{P}^2\}$. The *Logic* $\mathcal{L}$ of a quantum mechanical system is the set of all linear bounded Hermitian projection operators on $\mathcal{H}$. We say that orthogonal projection operators $\hat{P}$ and $\hat{Q}$ are *perpendicular*, denoted as $\hat{P} \perp \hat{Q}$, if and only if $\text{range}(\hat{P}) \perp \text{range}(\hat{Q})$. For $\hat{P}, \hat{Q} \in \mathcal{L}$, we define $\hat{P} \leq \hat{Q}$ if $\text{range}(\hat{P}) \subset \text{range}(\hat{Q})$. We define $\hat{P}^\perp := \hat{I} - \hat{P}$.

Fact. For $\hat{P}, \hat{Q} \in \mathcal{L}$, $\hat{P} \leq \hat{Q}$ iff $\hat{Q}\hat{P} = \hat{P}$.

Fact. For $\hat{P}, \hat{Q} \in \mathcal{L}$, $\hat{P} \leq \hat{Q}^\perp$ iff $\hat{P} \perp \hat{Q}$ iff $\hat{Q}\hat{P} = 0$ iff $\hat{P}\hat{Q} = 0$.

Fact. For $\hat{P}, \hat{Q} \in \mathcal{L}$, if $\hat{P} \perp \hat{Q}$, then $\hat{P} + \hat{Q} \in \mathcal{L}$.

Fact. The set of orthogonal projection operators $\mathcal{L}$ on $\mathcal{H}$ has the following properties:

(1) The zero projection $\hat{0} : \mathcal{H} \to \{0\}$ is in $\mathcal{L}$;
(2) if $\hat{P} \in \mathcal{L}$, then $\hat{P}^\perp := \hat{I} - \hat{P}$ is in $\mathcal{L}$;
(3) If a sequence of projectors $\hat{P}_1, \hat{P}_2, \cdots \in \mathcal{L}$ s.t. $\hat{P}_i \perp \hat{P}_j$, $\forall i \neq j$, then $\sum_i \hat{P}_i$ is in $\mathcal{L}$.

These properties are analogous to those of a σ-algebra.

2.3.3. Observables. A physical quantity A relating to this system is called an *observable*, which is represented by a Hermitian unbounded linear operator $\hat{A}$:

$\mathrm{Dom}\hat{A} \to \mathcal{H}$, where $\mathrm{Dom}\hat{A}$ is the domain of $\hat{A}$. Observables are closed and densely defined in $\mathcal{H}$ (the following definitions of $\mathrm{Dom}\hat{A}$, observables closed and densely defined are found in [**19**]).

Thus the set of observables is $\mathcal{U} = \{\hat{A} : \mathrm{Dom}(\hat{A}) \subset \mathcal{H} \to \mathcal{H} | \mathrm{Dom}(\hat{A})$ is a dense subspace of $\mathcal{H}$, $\hat{A}$ is linear, $\mathrm{graph}(\hat{A})$ is closed in $\mathcal{H} \times \mathcal{H}$, and $\mathrm{Dom}(\hat{A}^\dagger) = \{x \in \mathcal{H} | \sup_{y \in \mathrm{Dom}(\hat{A}), y \neq 0} \frac{|(x, \hat{A}y)|}{\|y\|} < \infty\} = \mathrm{Dom}(\hat{A})$, and $(\hat{A}x, y) = (x, \hat{A}y)\, \forall x, y \in \mathrm{Dom}(\hat{A})\}$.

By spectral decomposition theorem [**19**], there is a spectral decomposition $F_{\hat{A}} : \mathcal{B}(\mathbb{R}) \to \mathcal{L}$ s.t.

$$\hat{A} = \int_{\mathbb{R}} \lambda \, dF_{\hat{A}}(\lambda).$$

Moreover, $F_{\hat{A}}$ has the following properties:

(1) $F_{\hat{A}}(\emptyset) = \hat{0}$, $F_{\hat{A}}(\mathbb{R}) = \hat{I}$.

(2) if $U, V \in \mathcal{B}(\mathbb{R})$ s.t. $U \cap V = \emptyset$, (i.e. $U \perp V$), then $F_{\hat{A}}(U) \perp F_{\hat{A}}(V)$.

(3) If $\{U_n\}_{n=1}^{\infty}$ is a sequence of sets in $\mathcal{B}(\mathbb{R})$ and $U_n \cap U_m = \emptyset, \forall n \neq m$, $U = \cup_{n=1}^{\infty} U_n$, then $F_{\hat{A}}(U) = \sum_{n=1}^{\infty} F_{\hat{A}}(U_n)$.

This shows that $F_{\hat{A}}$ is the quantum analog of the mapping $R \mapsto f^{-1}(R)$ when f is a classical observable.

Bounded observables form a Jordan-Lie algebra (Landsman [**23**]) with the following two binary operations. If $\hat{A}, \hat{B}$ are bounded observables on Hilbert space $\mathcal{H}$, then define $\hat{A} \circ \hat{B} = \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A})$, which is analogous to fg of classical observables f and g; also define $\{\hat{A}, \hat{B}\} = \frac{i}{\hbar}(\hat{A}\hat{B} - \hat{B}\hat{A})$, where $\hbar = 1.055 \times 10^{-34}$Js is the Planck's constant. $\{\hat{A}, \hat{B}\}$ is analogous to $\{f, g\}$, the Poisson bracket of classical observables f and g. The following properties hold:

$$\hat{A} \circ \hat{B} = \hat{B} \circ \hat{A}, \; \{\hat{A}, \hat{B}\} = -\{\hat{B}, \hat{A}\}$$

$$\hat{A} \circ (\beta_1 \hat{B}_1 + \beta_2 \hat{B}_2) = \beta_1 \hat{A} \circ \hat{B}_1 + \beta_2 \hat{A} \circ \hat{B}_2$$

$$\{\hat{A}, \beta_1 \hat{B}_1 + \beta_2 \hat{B}_2\} = \beta_1 \{\hat{A}, \hat{B}_1\} + \beta_2 \{\hat{A}, \hat{B}_2\}$$

$$\text{Leibnitz rule: } \{\hat{A}, \hat{B} \circ \hat{C}\} = \{\hat{A}, \hat{B}\} \circ \hat{C} + \hat{B} \circ \{\hat{A}, \hat{C}\}$$

$$\text{Jacobi identity: } \{\hat{A}, \{\hat{B}, \hat{C}\}\} = \{\{\hat{A}, \hat{B}\}, \hat{C}\} + \{\hat{B}, \{\hat{A}, \hat{C}\}\}$$

In the quantum case we have

$$(\hat{A} \circ \hat{B}) \circ \hat{C} - \hat{A} \circ (\hat{B} \circ \hat{C}) = \frac{\hbar^2}{4}\{\{\hat{A}, \hat{C}\}, \hat{B}\}$$

whereas in the classical case we have $(fg)h - f(gh) = 0$.

Fact. If $\hat{A}$ and $\hat{B}$ are bounded observables, and $F_{\hat{A}}$, $F_{\hat{B}}$ are their spectral decompositions, then $F_{\hat{A}}(U)F_{\hat{B}}(U) = F_{\hat{B}}(U)F_{\hat{A}}(U)$, $\forall U \in \mathcal{B}(\mathbb{R})$, iff $\hat{A}\hat{B} = \hat{B}\hat{A}$ [**40**]. We usually use $[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A} = 0$ to represent that $\hat{A}$ and $\hat{B}$ commute.

This fact give us a hint about how to extend the notion of commuting bounded operators to the context of unbounded operators.

Definition. Observables $\hat{A}_1, \cdots, \hat{A}_n$ are commuting if $\forall R \in \mathcal{B}(\mathbb{R})$,

$$F_{\hat{A}_i}(R)F_{\hat{A}_j}(R) = F_{\hat{A}_j}(R)F_{\hat{A}_i}(R), \forall 1 \leq i, j \leq n.$$

Theorem. (Simultaneous Diagonalization theorem [**19**])

Suppose $\hat{A}_1, \cdots, \hat{A}_N$ are commuting observables. Then there exists a projection valued measure $F_{(\hat{A}_1, \cdots, \hat{A}_N)} : \mathcal{B}(\mathbb{R}^N) \to \mathcal{L}$ such that

$$\hat{A}_j = \int_{\lambda \in \mathbb{R}^N} \lambda_j \, dF_{(\hat{A}_1, \cdots, \hat{A}_N)}(\lambda), \quad j = 1, \cdots, N$$

where $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}$.

Moreover, if $\hat{A}_j = \int_{-\infty}^{\infty} \lambda \, dF_{\hat{A}_j}(\lambda)$ then

$$F_{(\hat{A}_1, \cdots, \hat{A}_N)}((a_1, b_1] \times \cdots \times (a_N, b_N]) = F_{\hat{A}_1}((a_1, b_1]) \circ \cdots \circ F_{\hat{A}_N}((a_N, b_N])$$

2.3.4. State. A *state* of a quantum system is given by a mapping $\mu : \mathcal{L} \to [0, 1]$ s.t. $\mu(\hat{0}) = 0, \mu(\hat{I}) = 1$, and if $\{\hat{P}_j\}$ is a sequence of projection in $\mathcal{L}$, such that $\hat{P}_j \perp \hat{P}_k$, whenever $j \neq k$ then $\mu(\sum_j \hat{P}_j) = \sum_j \mu(\hat{P}_j)$. This concept is analogous to the classical state i.e. a probability measure.

Definition. A bounded operator $\hat{D}$ on the Hilbert space $\mathcal{H}$ is of *trace class* if, for any orthonormal basis $\{e_i\}$ of $\mathcal{H}$, the series

$$\sum_i |(e_i, \hat{D}e_i)| < \infty;$$

In this case the sum

$$\operatorname{tr}(\hat{D}) = \sum_i (e_i, \hat{D}e_i)$$

exists for any orthonormal basis $\{e_i\}$, and is independent of the basis used; it is called the trace of $\hat{D}$. [**40**]

Definition. A bounded operator $\hat{D}$ on the Hilbert space $\mathcal{H}$ is *non-negative* if $(h, \hat{D}h) \geq 0$ for all $h \in \mathcal{H}$.

Gleason's theorem [**40**] asserts that for every state μ there is a *density operator* $\hat{D}$, i.e. a Hermitian, non-negative, trace class operator with unit trace, such that

$$\mu(\hat{P}) = \operatorname{tr}(\hat{D}\hat{P}), \ \forall \hat{P} \in \mathcal{L}.$$

μ is identified with $\hat{D}$, so we usually write μ as $\mu_{\hat{D}}$. Moreover, $\hat{D}$ is analogous to the classical density function for an absolutely continuous state. The trace is analogous to integration with respect to Lebesgue measure $d^n\mathbf{p}d^n\mathbf{q}$.

A *pure state* is an extreme point of the convex set of all states. Gleason's Theorem also identifies all the pure states to be exactly those of the form $\hat{D} = ff^{\dagger}, \quad f \in \mathcal{H}, \quad \|f\| = 1$, where $f^{\dagger} : \mathcal{H} \to \mathbb{C} : \ g \mapsto (f, g)$. Let us check that $\hat{D} = ff^{\dagger}$ is a density operator:

(1) $\hat{D}g = ff^{\dagger}(g) = f(f, g)$ is a linear function of g with values in $\mathcal{H}$;

(2) $\|\hat{D}g\| = \|f\| \, |(f, g)| \leq \|f\| \, \|f\| \, \|g\| = \|g\|$, so $\hat{D}$ is bounded;

(3) $(\hat{D}g, h) = (f(f, g), h) = \overline{(f, g)}(f, h) = (g, f)(f, h) = (g, f(f, h))$
$= (g, \hat{D}h)$, so $\hat{D}$ is Hermitian;

(4) because $(g, \hat{D}g) = (g, f(f, g)) = (g, f)(f, g) = |(g, f)|^2 \geq 0$, so $\hat{D}$ is non-negative.

(5) $\mathrm{tr}(\hat{D}) = \sum_{n=1}^{\infty}(e_n, \hat{D}e_n) = \sum_{n=1}^{\infty}(e_n, f)(f, e_n) = \sum_{n=1}^{\infty}|(e_n, f)|^2 = \|f\|^2$
$= 1$, where $\{e_i\}$ is any orthonormal basis of $\mathcal{H}$.

Remark: For the pure state $\hat{D} = ff^{\dagger}$, $\hat{D}$ is not to be identified with the unit vector in $f \in \mathcal{H}$. To see, if $e^{i\theta} \in S^1$, then $fe^{i\theta}(fe^{i\theta})^{\dagger} = fe^{i\theta}e^{-i\theta}f^{\dagger} = ff^{\dagger}$. Thus $fe^{i\theta}$ and f determine the same pure state. For this reason, pure states are often defined as rays in $\mathcal{H}$.

2.3.5. State Evolution. The state of an isolated system with a known Hamiltonian operator $\hat{H}$ evolves according to the rule:

$$\hat{D}(t) = U(t)\hat{D}(0)U(t)^{-1},$$

where the unitary group $\{U(t)\}$ is the solution of the initial value problem for the differential equation(Schrodinger's equation)(Reference: converse of Stone's Theorem [**19**] for the existence of $U(t)$):

$$\begin{cases} \frac{d}{dt}U(t) = \frac{1}{i\hbar}\hat{H}U(t) \\ U(0) = I. \end{cases}$$

It is clear that $\hat{D}'(t) = -\frac{i}{\hbar}[\hat{H}, \hat{D}(t)] = -\{\hat{H}, \hat{D}(t)\}$ which is called Liouville-von Neumann equation, and hence this evolution is analogous to classical density function evolution. If $\hat{D}(0) = f_0 f_0^{\dagger}$, then $\hat{D}(t) = f_t f_t^{\dagger}$, where $f_t = U(t)f_0$, since $f_t^{\dagger}g = (f_t, g) = (U(t)f_0, g) = (f_0, U(t)^{\dagger}g) = f_0^{\dagger}U(t)^{-1}g$ for all $g \in \mathcal{H}$.

By spectral decomposition theorem ([**19**] p.270), $\hat{H}f(t) = \int_{-\infty}^{\infty} \lambda dF_{\hat{H}}(\lambda)f(t)$, and $f = \int_{-\infty}^{\infty} dF_{\hat{H}}(\lambda)f$. The Schrödinger's equation

$$\begin{aligned}
&i\hbar f'(t) = \hat{H}f(t) \\
&i\hbar\int_{-\infty}^{\infty} dF_{\hat{H}}(\lambda)f'(t) = \int_{-\infty}^{\infty} \lambda dF_{\hat{H}}(\lambda)f(t) \\
&i\hbar\frac{d}{dt}\left[dF_{\hat{H}}(\lambda)f(t)\right] - \lambda\left[dF_{\hat{H}}(\lambda)f(t)\right] = 0 \\
&\frac{d}{dt}\left[dF_{\hat{H}}(\lambda)f(t)\right] + \frac{i\lambda}{\hbar}\left[dF_{\hat{H}}(\lambda)f(t)\right] = 0
\end{aligned}$$

$$\frac{d}{dt}\left[e^{\frac{i\lambda t}{\hbar}} dF_{\hat{H}} f(t)\right] = 0$$

$$e^{\frac{i\lambda t}{\hbar}} dF_{\hat{H}} f(t) - dF_{\hat{H}} f(0) = 0$$

$$dF_{\hat{H}}(\lambda) f(t) = e^{-\frac{i\lambda t}{\hbar}} dF_{\hat{H}}(\lambda) f(0)$$

$$\Leftrightarrow f(t) = \int_{-\infty}^{\infty} e^{-\frac{i\lambda t}{\hbar}} dF_{\hat{H}}(\lambda) f(0) = U(t)[f(0)],$$

where $U(t) = \int_{-\infty}^{\infty} e^{-\frac{i\lambda t}{\hbar}} dF_{\hat{H}}(\lambda)$.

2.3.6. Predicting Measurement Outcomes.

Noiseless Measurement. Suppose the observable $\hat{A} = \int_{-\infty}^{\infty} \lambda \, dF_{\hat{A}}(\lambda)$ is to be measured and we wish to predict the distribution of the measured values of this observable when the system is in the state $\hat{D}$. If we suppose the measurement is *noiseless* then the probability that the measured values will be in the set $R \in \mathcal{B}(\mathbb{R})$ is $\nu_{\hat{A}}(R) = \text{tr}\,[F_{\hat{A}}(R)\hat{D}]$.

Suppose we wish to predict the outcome of simultaneous measurements of n commuting observables $\hat{A}_1, \cdots, \hat{A}_n$; denote $\hat{\mathbf{A}} = (\hat{A}_1, \cdots, \hat{A}_n)$. Let $F_{\hat{\mathbf{A}}} : \mathcal{B}(\mathbb{R}^n) \to \mathcal{L}$ be the associated spectral measure s.t. $\hat{A}_j = \int_{\lambda \in \mathbb{R}^n} \lambda_j \, dF_{(\hat{A}_1, \cdots, \hat{A}_n)}(\lambda), 1 \leq j \leq n$. The predicted joint distribution of measured values is

$$\nu_{\hat{\mathbf{A}}}(R) = \mu(F_{\hat{\mathbf{A}}}(R)) = \text{tr}(\hat{D} F_{\hat{\mathbf{A}}}(R))$$

for all $R \in \mathcal{B}(\mathbb{R}^n)$.

For instance the predicted mean of the measured values of observable $\hat{A}_j$ is

$$\begin{aligned} < \hat{A}_j > &= \int_{\lambda \in \mathbb{R}^n} \lambda_j \, d\nu_{\hat{\mathbf{A}}}(\lambda) \\ &= \int_{\lambda \in \mathbb{R}^n} \lambda_j \mu_{\hat{D}}(\, dF_{\hat{\mathbf{A}}}(\lambda)) \\ &= \int_{\lambda \in \mathbb{R}^n} \lambda_j \text{tr}(\hat{D} \, dF_{\hat{\mathbf{A}}}(\lambda)) \\ &= \text{tr}(\hat{D} \int_{\lambda \in \mathbb{R}^n} \lambda_j \, dF_{\hat{\mathbf{A}}}(\lambda)) \\ &= \text{tr}(\hat{D} \hat{A}_j) \end{aligned}$$

Noisy Measurement. If the measurement apparatus has noise with mean zero and variance ϵ^2 then by analogy with the classical expression we agree that the probability that the noisy measured values will be in the set $R \in \mathcal{B}(\mathbb{R})$ is given by

$$(\nu_{\hat{A}} * \alpha_\epsilon^2)(R) = \int_{-\infty}^{\infty} (\chi_R * \alpha_\epsilon^2)(\lambda)\, \nu_{\hat{A}}(\,d\lambda).$$

This prescription has been justified in [**30**]. Various equivalent expressions exist for this probability:

$$\begin{aligned}
\int_{-\infty}^{\infty} (\chi_R * \alpha_\epsilon^2)(\lambda)\, \nu_{\hat{A}}(\,d\lambda) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi_R(y)\alpha_\epsilon(\lambda - y)^2\, dy \text{ tr } [F_{\hat{A}}(\,d\lambda)\hat{D}] \\
&= \int_R \text{tr } [\int_{-\infty}^{\infty} \alpha_\epsilon(y-\lambda)^2 F_{\hat{A}}(\,d\lambda)\hat{D}]\, dy \\
&= \int_R \text{tr } [\alpha_\epsilon(y\hat{1} - \hat{A})^2\hat{D}]\, dy \\
&= \text{tr } [\int_R \alpha_\epsilon(y\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(y\hat{1} - \hat{A})\, dy] \\
&= \text{tr } [\int_R \alpha_\epsilon(y\hat{1} - \hat{A})^2\, dy\hat{D}].
\end{aligned}$$

Thus we predict that the noisily measured values will be continuously distributed on the real y line with the density function

$$\nu_{\hat{A},\epsilon}(y) = \text{tr } [\alpha_\epsilon(y\hat{1} - \hat{A})^2\hat{D}].$$

However, this continuous probability distribution will tend to $\nu_{\hat{A}}$, which can have probability atoms, in the limit as $\epsilon \to 0^+$.

2.3.7. Updating the state after the measurement. In quantum mechanics one must be much more careful about the process of measurement and its theoretical treatment than was the case in classical mechanics. Measurements of atoms, and in particular electrons, could only be explained if one admits that the measurement process might perturb the system in ways which may be impossible to completely control, and this may necessitate a fundamental revision in the theoretical treatment of physical quantities (such as position and momentum). However the insistence of most of the founders of quantum mechanics that the state of the system represents "all that can be known about the system" (i.e. "what really is") as opposed to "what an observer knows about the system at a particular level of physical theory and approximation" has lead to much murky philosophy (i.e the Copenhagen interpretation) and has helped give the entire subject a reputation of mystery and incomprehensibility.

Perhaps therefore it should not be surprising that in order to give an account of how to update the state after a measurement of an arbitrary observable $\hat{A} = \int_{-\infty}^{\infty} \lambda F_{\hat{A}}(\,d\lambda)$ one must also admit that all real measurements have some noise. The noiseless case represents an idealized limit which may fail to exist in some cases. In the following we will describe how to update the state after a *canonical measurement* [**30**] (M. Ozawa). A measurement is canonical if it can be modeled as a particularly simple type of interaction between the system and an apparatus *pointer*. This interaction is treated quantum mechanically. The Hilbert space of the apparatus pointer is $\mathcal{K} = L^2(\mathbb{R})$. If $\alpha(q)$ is a square-integrable complex-valued function of $q \in \mathbb{R}$ then $\alpha \in \mathcal{K}$. The pointer position observable is $\hat{Q}$, defined by $(\hat{Q}\alpha)(q) = q\alpha(q)$. The pointer momentum observable is $\hat{P}$, defined by $(\hat{P}\alpha)(q) = -i\hbar\alpha'(q)$. We assume the system (being measured) and the apparatus pointer interact for a time period of duration Δt. During this interaction we must treat the system and the apparatus as a composite system, with the Hilbert space $\mathcal{H} \otimes \mathcal{K} \cong L^2(\mathbb{R}, \mathcal{H})$. Most importantly, we assume that during this interaction the evolution is governed by Schrödinger's equation with the Hamiltonian $\hat{H}_{\text{int}} = (\Delta t)^{-1} \hat{A} \otimes \hat{P}$. This interaction Hamiltonian is not easy to justify on physical grounds, as might be expected since the actual measurement interaction is between the system and a complex apparatus which contains the pointer as a rather small part. This Hamiltonian was chosen (by von Neumann [**29**]) because it yields an explicitly solvable evolution of the composite system, and it leads to results which are consistent with experimental observations. Other choices of $\mathcal{K}$ and other interaction Hamiltonians have been proposed in various cases, but the choices we have listed above characterize canonical measurements. Canonical measurements seem to be adequate for the realm of atoms and molecules.

If the state of the system just prior to the beginning of the interaction between the system and the apparatus is $\mu(\hat{P}) = \text{tr}\ (\hat{D}\hat{P})$ for all $\hat{P} \in \mathcal{L}$, where $\hat{D}$ is a density operator, and the state of the apparatus pointer at this time is given by the density operator $\alpha_\epsilon \alpha_\epsilon^\dagger$ (where α_ϵ is defined in section §2.2.6), then the state of the composite system is $\hat{D} \otimes (\alpha_\epsilon \alpha_\epsilon^\dagger)$. If $\{U(t)\}$ is the unitary evolution group for the composite system under the Hamiltonian $\hat{H}_{\text{int}}$ then the state of the composite system at the end of the interaction is (c.f. [**30**] M. Ozawa)

$$U(\Delta t)[\hat{D} \otimes (\alpha_\epsilon \alpha_\epsilon^\dagger)]U(\Delta t)^\dagger$$

This state contains all the information about the correlations between the system and the pointer which exist at the end of the measurement interaction. If we ignore all information concerning the pointer we can obtain a state $\hat{D}'$ of the system at the end of the measurement:

$$\hat{D}' = \text{tr}_{\mathcal{K}}\{U(\Delta t)[\hat{D} \otimes (\alpha_\epsilon \alpha_\epsilon^\dagger)]U(\Delta t)^\dagger\}.$$

We may think of the partial trace $\text{tr}_{\mathcal{K}}$ as the operation of "averaging over the pointer degrees of freedom". If $\{e_n\}_{n=1}^{\infty}$ is a complete orthonormal set in $\mathcal{K}$ then the partial trace of a trace-class operator $\hat{B}$ in $\mathcal{H} \otimes \mathcal{K}$ is an operator in $\mathcal{H}$, which when applied to $\psi \in \mathcal{H}$ yields

$$[\text{tr}_{\mathcal{K}}\hat{B}]\psi = \sum_{n=1}^{\infty}(\hat{1} \otimes e_n^\dagger)B(\psi \otimes e_n).$$

It is shown in [**30**] (M. Ozawa) that

$$\hat{D}' = \int_{-\infty}^{\infty} \alpha_\epsilon(q\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger \, dq.$$

In the above the operator $\alpha_\epsilon(q\hat{1} - \hat{A})$ is computed using the spectral theorem:

$$\alpha_\epsilon(q\hat{1} - \hat{A}) = \int_{-\infty}^{\infty} \alpha_\epsilon(q - \lambda)\, F_{\hat{A}}(d\lambda).$$

Clearly, in this case we have $\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger = \alpha_\epsilon(q\hat{1} - \hat{A})$. $\hat{D}'$ would be the state of the system after the (canonical) measurement interaction if we were to throw away all the measured data. (Remember, the state encodes the observer's knowledge of the system!) The fact that undergoing a measurement interaction causes a state change from $\hat{D}$ to $\hat{D}'$ is a marked contrast with the classical case, and underlines the point that in quantum mechanics we cannot ignore the effect of the measurement interaction on the system. The transition $\hat{D} \to \hat{D}'$ is often called the *dynamical state change*.

Now we come to the issue of observing the pointer. Since we have treated the pointer as a quantum system, must we hypothesize another apparatus to measure it? To make this unnecessary we assume that the pointer is heavy enough to be accurately treated as a classical system. (The nature and accuracy of this approximation is discussed in [**35**] and in Balian [**9**].) Hence we hypothesize a noiseless measurement (for the composite system) of the position of the pointer which incurs no perturbation of the state of the composite system. The observable in the Hilbert space $\mathcal{H} \otimes \mathcal{K}$ corresponding to the pointer position is $\hat{1} \otimes \hat{Q}$. If $R \in \mathcal{B}(\mathbb{R})$ then the probability that the outcome of this noiseless measurement will lie in R is

$$\mathrm{tr}_{\mathcal{H}\otimes\mathcal{K}}\big\{[\hat{1} \otimes F_{\hat{Q}}(R)]\{U(\Delta t)[\hat{D} \otimes (\alpha_\epsilon \alpha_\epsilon^\dagger)]U(\Delta t)^\dagger\}\big\}.$$

Using the identities presented in [**30**] (M. Ozawa) this probability can be reexpressed as:

$$\mathrm{tr}_{\mathcal{H}}\left[\int_R \alpha_\epsilon(q\hat{1} - \hat{A})^2 \, dq \hat{D}\right] = \int_R \nu_{\hat{A},\epsilon}(q)\, dq.$$

This is exactly the probability that we gave in section §2.3.6 that a noisy (variance ϵ^2) measurement of $\hat{A}$ gives values in the set R when the system is in the state $\hat{D}$ before the measurement. That formula we derived by analogy with the classical case, but we see now that it also arises from the formalism of canonical measurement where the state of the pointer before the measurement is $\alpha_\epsilon \alpha_\epsilon^\dagger$. This result bolsters our confidence in the canonical measurement formalism.

Now suppose that the measured values are observed to be distributed according to the density function $\tilde{\nu}(q)$ instead of according to the predicted density function $\nu_{\hat{A},\epsilon}(q)$. How should the state of the system be updated to take this new information into account? By analogy with the classical case we note that the state $\hat{D}'$ of the system after the measurement can be written

$$\hat{D}' = \int_{-\infty}^{\infty} \frac{\alpha_\epsilon(q\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger}{\text{tr }[\alpha_\epsilon(q\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger]}\nu_{\hat{A},\epsilon}(q)\, dq.$$

The quotient

$$\hat{D}'_q = \frac{\alpha_\epsilon(q\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger}{\text{tr }[\alpha_\epsilon(q\hat{1} - \hat{A})\hat{D}\alpha_\epsilon(q\hat{1} - \hat{A})^\dagger]}$$

we interpret as a *conditional density operator*, analogous to the fibre measure in the classical case. We have assumed that the process of observing the pointer does not perturb the composite system at all, hence we have every reason to believe that the conditional density operators $\hat{D}'_q$ should be unaffected by the observation of the pointer. Hence, as in the classical case, we define the updated state to be:

$$\tilde{\hat{D}} = \int_{-\infty}^{\infty} \hat{D}'_q \tilde{\nu}(q)\, dq,$$

and $\tilde{\mu}(P) = \text{tr }(\tilde{D}P)$ for all $P \in \mathcal{L}$. The transition $\hat{D}' \to \tilde{\hat{D}}$ is often called *state reduction.*

In the very special case where $\hat{A} = \sum_{i=1}^m \lambda_i A_i$, where $A_i = F_{\hat{A}}(\{\lambda_i\})$, we can make sense of the noiseless limit $\epsilon \to 0^+$ of these results. In this case $\hat{D}'$ becomes

$$\hat{D}' = \sum_{i=1}^m \sum_{j=1}^m \int_{-\infty}^{\infty} \alpha_\epsilon(q - \lambda_i)\alpha_\epsilon(q - \lambda_j)\, dq A_i \hat{D} A_j.$$

When ϵ is small when compared to $\min_{i \neq j}|\lambda_i - \lambda_j|$ the off-diagonal $(i \neq j)$ coefficient $\int_{-\infty}^{\infty} \alpha_\epsilon(q - \lambda_i)\alpha_\epsilon(q - \lambda_j)\, dq$ is very small, and hence the noiseless limit of $\hat{D}'$ is given by:

$$\hat{D}'_{\text{NL}} = \sum_{i=1}^m A_i \hat{D} A_i.$$

"NL" stands for *Neumann-Lüders*, since von Neumann (and later Lüders) derived this formula for the dynamical state change from various auxiliary postulates about the measurement process, such as its *repeatability*. If $p_i = \text{tr}\,[A_i \hat{D} A_i]$ then then noiseless limit of $\nu_{\hat{A}} * \alpha_\epsilon^2$ is $\nu_{\hat{A}} = \sum_{i=1}^m p_i \delta_{\lambda_i}$. If $\tilde{\nu} << \nu_{\hat{A}}$ then we must have $\tilde{\nu} = \sum_{i=1}^m \tilde{p}_i \delta_{\lambda_i}$, for some nonnegative numbers $\tilde{p}_i$, where $\sum_{i=1}^m \tilde{p}_i = 1$. Also if for some i we have $p_i = 0$ then we must also have $\tilde{p}_i = 0$. Thus the noiseless limit of the reduced state $\tilde{\hat{D}}$ should be:

$$\tilde{\hat{D}}_{\text{NL}} = \sum_{i=1}^m \frac{A_i \hat{D} A_i}{p_i} \tilde{p}_i.$$

In order to show that this is indeed the limit as $\epsilon \to 0^+$ of our general formula for the reduced state we must hypothesize that the observed distribution $\tilde{\nu}(q)\,dq$ varies with ϵ (the noise level of the apparatus), and tends to $\tilde{\nu} = \sum_{i=1}^m \tilde{p}_i \delta_{\lambda_i}$ as $\epsilon \to 0^+$. For example it is reasonable to suppose that

$$\tilde{\nu}(q) = \sum_{i=1}^m \tilde{p}_i \alpha_\epsilon (q - \lambda_i)^2.$$

Then the reduced state becomes

$$\tilde{\hat{D}} = \int_{-\infty}^{\infty} \frac{\sum_{i,j} \alpha_\epsilon(q-\lambda_i)\alpha_\epsilon(q-\lambda_j) A_i \hat{D} A_j}{\sum_i \alpha_\epsilon(q-\lambda_i)^2 p_i} \sum_k \tilde{p}_k \alpha_\epsilon(q-\lambda_k)^2\, dq,$$

The quotient in the above integral is the conditional density operator $\hat{D}'_q$, and it is not difficult to see that $\lim_{\epsilon\to 0^+} \hat{D}'_q$ is a piecewise constant function of q:

$$\lim_{\epsilon\to 0^+} \hat{D}'_q = \frac{A_i \hat{D} A_i}{p_i}, \qquad \text{if } |q-\lambda_i| < |q-\lambda_j| \text{ for all } j \neq i.$$

Thus the desired convergence is now clear.

Suppose the state $\hat{D}$ before the measurement is arbitrary and for some $1 \leq i \leq m$ we have that A_i is a projection onto a one-dimensional subspace of $\mathcal{H}$ spanned by the unit vector ψ_i, and $p_i = (\psi_i, \hat{D}\psi_i) > 0$. If the system is a large ensemble of “identical” subsystems, and the apparatus is designed to select only those subsystems with the measured value λ_i, discarding the other subsystems, then we have $\tilde{p}_i = 1$ for the selected subensemble. The result of the measurement is that the state of the selected subensemble has been changed to $\tilde{\hat{D}}_{\text{NL}} = \psi_i \psi_i^\dagger$. Thus the (noiseless) measurement process (of a family of commuting observables) can be used to *prepare* a system in a specific pure state.

If $\hat{A} = \sum_{i=1}^{\infty} \lambda_i F_{\hat{A}}(\{\lambda_i\})$, where the eigenvalues λ_i have an accumulation point then the noiseless limit is more delicate. Examples of this case, as well as that of observables with continuous spectrum, will only be briefly discussed at §3.2.9.

Chapter 3

Examples of Quantum Systems

In this chapter, we apply the quantum formalism to several microscopic systems from simple to complicate: the spin system, the artificially spinless H atom and H_2^+ ion, and the real H_2 and H_3 molecule systems.

3.1. Spin Systems

Comments on the Idea of Spin If a system is composed of multiple subsystems capable of relative motion and if those subsystems rotate around one another then the system gains angular momentum. Spin is the name for angular momentum possessed by a system without any attempt to attribute it to a rotary motion of subsystems about an axis. In the case of an electron there are no detectable subsystems, and yet there is a detectable angular momentum. If the subsystems were charged then their rotary motion would create a magnetic field. Spin angular momentum is also associated with a magnetic field created by the system, and experimental detection of spin is usually related to how this field affects overall motion of the system.

(Adapted from [**46**]) In 1921, Otto Stern and Walter Gerlach performed an experiment which showed the quantization of electron spin into two orientations. This made a major contribution to the development of the quantum theory of the atom.

The actual experiment was carried out with a beam of silver atoms from a hot oven because they could be readily detected using a photographic emulsion. The silver atoms allowed Stern and Gerlach to study the magnetic properties of a single electron because these atoms have a single outer electron which moves in the Coulomb potential caused by the 47 protons of the nucleus shielded by the 46 inner electrons. Since this electron has zero orbital angular momentum (orbital quantum number $l = 0$), one would expect there to be no interaction with an external magnetic field.

Stern and Gerlach directed the beam of silver atoms into a region of nonuniform magnetic field. A magnetic dipole moment will experience a force proportional to the field gradient since the two "poles" will be subject to different fields. Classically one would expect all possible orientations of the dipoles so that a continuous smear would be produced on the photographic plate, but they found that the field separated the beam into two distinct parts, indicating just two possible orientations of the magnetic moment of the electron.

But how does the electron obtain a magnetic moment if it has zero orbital angular momentum and therefore produces no "current loop" to produce a magnetic moment? In 1925, Samuel A. Goudsmit and George E. Uhlenbeck postulated that the electron had an intrinsic angular momentum, independent of its orbital characteristics. In classical terms, a ball of charge could have a magnetic moment if it were spinning such that the charge at the edges produced an effective current loop. This kind of reasoning led to the use of "electron spin" to describe the intrinsic angular momentum.

3.1.1. Hilbert Space. The internal angular momentum of a particle is called its *spin*. For a system of n spins i.e. n-particles with spin and with unknown positions and momenta, the Hilbert space is $\mathcal{H} = \mathbb{C}^{2^n} \cong \mathbb{C}^2 \overset{1}{\otimes} \cdots \overset{n-1}{\otimes} \mathbb{C}^2$.

3.1.2. Logic. Logic is $\mathcal{L} = \{\hat{P} \in \mathbb{C}^{2^n \times 2^n} | \hat{P}^\dagger = \hat{P} = \hat{P}^2\}$, where $\hat{P}^\dagger$ is the complex conjugate transpose of $\hat{P}$.

3.1.3. Observables. The observables are members of the set $\mathcal{U} = \{\hat{A} \in \mathbb{C}^{2^n \times 2^n} | \hat{A}^\dagger = \hat{A}\}$.

By spectral decomposition theorem [**19**], for $\hat{A} \in \mathcal{U}$, we have $\hat{A}X = X\Lambda$, where $X = (\mathbf{x}_1, \cdots, \mathbf{x}_N)$ is unitary, $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_N), N = 2^n$, and $\lambda_i \in \mathbb{R}$ is the eigenvalue of $\hat{A}$ with corresponding eigenvector $\mathbf{x}_i \in \mathbb{C}^N$. We define $\sigma(\hat{A}) = \{\lambda_1, \cdots, \lambda_N\}$ to be the spectrum of $\hat{A}$, and $F_{\hat{A}}$ to be the atomic measure supported on $\sigma(\hat{A})$ defined as follows.

$$F_{\hat{A}}(R) := \sum_{\lambda \in R} F_{\hat{A}}(\{\lambda\}), \text{ where } R \in \mathcal{B}(\mathbb{R}),$$

$$F_{\hat{A}}(\{\lambda\}) := \sum_{1 \le i \le N, \lambda_i = \lambda} \mathbf{x}_i \mathbf{x}_i^\dagger, \lambda \in \mathbb{R}.$$

Therefore

$$\hat{A} = \sum_{\lambda \in \sigma(\hat{A})} \lambda F_{\hat{A}}(\{\lambda\}) = \sum_{i=1}^{N} \lambda_i \mathbf{x}_i \mathbf{x}_i^\dagger.$$

Recall that the *Pauli matrices* are $\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Definition. For a n-spin system consisting n electrons, let $\hat{S}_j(i)$ be the jth component of the spin of the ith electron, where $j = 1, 2, 3$,

$$\hat{S}_j(i) := I \overset{1}{\otimes} \cdots \overset{i-1}{\otimes} \frac{\hbar}{2}\hat{\sigma}_j \overset{i}{\otimes} \cdots \overset{n-1}{\otimes} I.$$

$\hat{S}_j := \sum_{i=1}^{n} \hat{S}_j(i)$, which is the jth component of the spin of the n-spin system.
$\hat{\mathbf{S}}^2 := \hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_3^2$ is the total squared spin of the n spin system.

Here we will work out the explicit formulae for the joint spectral decompositions of $\hat{\mathbf{S}}^2$ and $\hat{S}_3$ for the cases of $n = 1, 2, 3$ respectively. See [**32**] for general n.

Example. For a single spin, $n = 1$, the members of the Hilbert space are spin functions $s : \{0, 1\} \to \mathbb{C}$, which we can identify with vectors $\begin{pmatrix} s(0) \\ s(1) \end{pmatrix}$ in $\mathbb{C}^2$.

(1) The x-component of the spin is $\hat{S}_1 = \frac{\hbar}{2}\hat{\sigma}_1$;

(2) The y-component of the spin is $\hat{S}_2 = \frac{\hbar}{2}\hat{\sigma}_2$;

(3) The z-component of the spin is $\hat{S}_3 = \frac{\hbar}{2}\hat{\sigma}_3$.

Then

$$
\begin{aligned}
\hat{\mathbf{S}}^2 &= \hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_3^2 \\
&= (\frac{\hbar}{2})^2 \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \\
&= \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}
$$

$\hat{S}_1$ has two eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, and their corresponding eigenvectors are $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively. Then the spectral decomposition of $\hat{S}_1$ is

$$
\begin{aligned}
\hat{S}_1 &= \frac{\hbar}{2} F_{\hat{S}_1}(\{\frac{\hbar}{2}\}) - \frac{\hbar}{2} F_{\hat{S}_1}(\{-\frac{\hbar}{2}\}) \\
&= \frac{\hbar}{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^\dagger - \frac{\hbar}{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}^\dagger \\
&= \frac{\hbar}{2} \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - \frac{\hbar}{2} \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
&= \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
&= \frac{\hbar}{2} \hat{\sigma}_1
\end{aligned}
$$

$\hat{S}_2$ has two eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, and their corresponding eigenvectors are $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$ and $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$ respectively. Then the spectral decomposition of $\hat{S}_2$ is

$$\begin{aligned}\hat{S}_2 &= \frac{\hbar}{2}F_{\hat{S}_2}(\{\frac{\hbar}{2}\}) - \frac{\hbar}{2}F_{\hat{S}_2}(\{-\frac{\hbar}{2}\})\\
&= \frac{\hbar}{2}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}^{\dagger} - \frac{\hbar}{2}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}^{\dagger}\\
&= \frac{\hbar}{2}\frac{1}{2}\begin{pmatrix}1 & -i\\ i & 1\end{pmatrix} - \frac{\hbar}{2}\frac{1}{2}\begin{pmatrix}1 & i\\ -i & 1\end{pmatrix}\\
&= \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\\
&= \frac{\hbar}{2}\hat{\sigma}_2\end{aligned}$$

$\hat{S}_3$ has two eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, and their corresponding eigenvectors are $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ respectively. The spectral decomposition of $\hat{S}_3$ is

$$\begin{aligned}\hat{S}_3 &= \frac{\hbar}{2}F_{\hat{S}_3}(\{\frac{\hbar}{2}\}) - \frac{\hbar}{2}F_{\hat{S}_3}(\{\frac{-\hbar}{2}\})\\
&= \frac{\hbar}{2}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}^{\dagger} - \frac{\hbar}{2}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}^{\dagger}\\
&= \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\\
&= \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\end{aligned}$$

$$= \frac{\hbar}{2}\hat{\sigma}_3$$

$\hat{\mathbf{S}}^2$ has only one eigenvalue $\frac{3\hbar^2}{4}$, there are two corresponding eigenvectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then the spectral decomposition of $\hat{\mathbf{S}}^2$ is

$$\begin{aligned}
\hat{\mathbf{S}}^2 &= \frac{3\hbar^2}{4} F_{\hat{\mathbf{S}}^2}(\{\frac{3\hbar^2}{4}\}) \\
&= \frac{3\hbar^2}{4}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\dagger} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}^{\dagger}\right] \\
&= \frac{3\hbar^2}{4}\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] \\
&= \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \frac{3\hbar^2}{4} I
\end{aligned}$$

It is easy to check that $\hat{S}_3\hat{\mathbf{S}}^2 = \hat{\mathbf{S}}^2\hat{S}_3$, so we have the joint spectral decomposition of $\hat{S}_3$ and $\hat{\mathbf{S}}^2$. For $X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ we have

$$\begin{aligned}
\hat{\mathbf{S}}^2 X &= X \cdot \mathrm{diag}(\frac{3\hbar^2}{4}, \frac{3\hbar^2}{4}) \\
\hat{S}_3 X &= X \cdot \mathrm{diag}(\frac{\hbar}{2}, -\frac{\hbar}{2})
\end{aligned}$$

Therefore, the joint spectral decomposition of $\hat{S}_3$ and $\hat{\mathbf{S}}^2$ is

$$\begin{aligned}
F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(\frac{3\hbar^2}{4}, \frac{\hbar}{2})\}) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\dagger} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\frac{3\hbar^2}{4},-\frac{\hbar}{2})\}) = \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}^\dagger = \begin{pmatrix}0&0\\0&1\end{pmatrix}$$

EXAMPLE. For $n=2$, the members of the Hilbert space are spin functions: $s:\{0,1\}^2\to\mathbb{C}$, which we can identify with vectors $\begin{pmatrix}s(0,0)\\s(0,1)\\s(1,0)\\s(1,1)\end{pmatrix}$ in $\mathbb{C}^4$. This is because

$$\begin{pmatrix}s_1(0)\\s_1(1)\end{pmatrix}\otimes\begin{pmatrix}s_2(0)\\s_2(1)\end{pmatrix} = \begin{pmatrix}s_1(0)s_2(0)\\s_1(0)s_2(1)\\s_1(1)s_2(0)\\s_1(1)s_2(1)\end{pmatrix} = \begin{pmatrix}(s_1\otimes s_2)(0,0)\\(s_1\otimes s_2)(0,1)\\(s_1\otimes s_2)(1,0)\\(s_1\otimes s_2)(1,1)\end{pmatrix}.$$

The x-component of the first spin is

$$\hat{S}_1(1) = \frac{\hbar}{2}\hat{\sigma}_1\otimes I = \tfrac{\hbar}{2}\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}.$$

The x-component of the second spin is

$$\hat{S}_1(2) = I\otimes\frac{\hbar}{2}\hat{\sigma}_1 = \tfrac{\hbar}{2}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}.$$

Then the x-component of the 2-spin system is

$$\hat{S}_1 = \hat{S}_1(1)+\hat{S}_1(2) = \tfrac{\hbar}{2}\begin{pmatrix}0&1&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\end{pmatrix}.$$

Therefore

$$\hat{S}_1^2 = \tfrac{\hbar^2}{4}\begin{pmatrix}0&1&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\end{pmatrix}\begin{pmatrix}0&1&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\end{pmatrix} = \tfrac{\hbar^2}{2}\begin{pmatrix}1&0&0&1\\0&1&1&0\\0&1&1&0\\1&0&0&1\end{pmatrix}.$$

The y-component of the first spin is

$$\hat{S}_2(1) = \frac{\hbar}{2}\hat{\sigma}_2\otimes I = \tfrac{\hbar}{2}\begin{pmatrix}0&0&-i&0\\0&0&0&-i\\i&0&0&0\\0&i&0&0\end{pmatrix}.$$

The y-component of the second spin is

$$\hat{S}_2(2) = I\otimes\frac{\hbar}{2}\hat{\sigma}_2 = \tfrac{\hbar}{2}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix}.$$

Then the y-component of the 2-spin system is

$$\hat{S}_2 = \hat{S}_2(1)+\hat{S}_2(2) = \tfrac{\hbar i}{2}\begin{pmatrix}0&-1&-1&0\\1&0&0&-1\\1&0&0&-1\\0&1&1&0\end{pmatrix}.$$

Therefore

$$\hat{S}_2^2 = \frac{-\hbar^2}{4}\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$= \frac{\hbar^2}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

The z-component of the first spin is

$$\hat{S}_3(1) = \frac{\hbar}{2}\hat{\sigma}_3 \otimes I = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The z-component of the second spin is

$$\hat{S}_3(2) = I \otimes \frac{\hbar}{2}\hat{\sigma}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The z-component of the 2-spin system is

$$\hat{S}_3 = \hat{S}_3(1) + \hat{S}_3(2) = \hbar\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Therefore

$$\hat{S}_3^2 = \frac{\hbar^2}{2}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Finally

$$\hat{\mathbf{S}}^2 = \hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_3^2 = \hbar^2\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

It is easy to check that $\hat{S}_3\hat{\mathbf{S}}^2 = \hat{\mathbf{S}}^2\hat{S}_3$. We can find $\tilde{X} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ such that

$$\hat{\mathbf{S}}^2\tilde{X} = \hbar^2\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \hbar^2\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}\hbar^2$$

$$= \tilde{X} \cdot \mathrm{diag}(0, 2\hbar^2, 2\hbar^2, 2\hbar^2)$$

$$\hat{S}_3\tilde{X} = \hbar\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \hbar\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \hbar$$

$$= \tilde{X} \cdot \mathrm{diag}(0, \hbar, 0, -\hbar)$$

$\hat{\mathbf{S}}^2$'s eigenvalues can be parameterized by $\hbar^2 S(S+1)$, where $S \in \{0, 1\}$, and $\hat{S}_3$'s eigenvalues can be parameterized by $\hbar m_S$, where $|m_S| \leq S, m_S \in \{-1, 0, 1\}$.

The column vectors of X, i.e. joint normalized eigenvectors of $\hat{\mathbf{S}}^2$ and $\hat{S}_3$ are:

(1) $\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} [\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)], S = 0,$ $m_S = 0$;

(2) $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $S = 1, m_S = 1$;

(3) $\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} [\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)]$, $S = 1, m_S = 0$;

(4) $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \otimes \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, $S = 1, m_S = -1$.

We call the state $\mathbf{x}\mathbf{x}^T$ represented by the first column vector $\mathbf{x}$ the *singlet state*, because there is only one allowed value of m_S i.e. 0. We call the states represented by the last three column vectors *triplet states*, because there are three allowed values of m_S i.e. $-1, 0, 1$, with the same value of S, i.e. $S = 1$.

Therefore, the joint spectral decomposition for $\hat{\mathbf{S}}^2$ and $\hat{S}_3$ is

$$F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(0,0)\}) = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}^\dagger = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(2\hbar^2, \hbar)\}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}^\dagger = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(2\hbar^2, 0)\}) = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}^\dagger = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(2\hbar^2, -\hbar)\}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Given the eigenvalue and corresponding eigenvector(s), we construct $\hat{S}_3$ and $\hat{\mathbf{S}}^2$ by spectral decomposition theorem as follows.

$$F_{\hat{\mathbf{S}}^2}(\{0\}) = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}^\dagger$$

$$= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{\hat{\mathbf{S}}^2}(\{2\hbar^2\}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}^\dagger + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}^\dagger + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^\dagger$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Leftrightarrow \hat{\mathbf{S}}^2 = 0\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + 2\hbar^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= 2\hbar^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$F_{\hat{S}_3}(\{0\}) = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}^\dagger + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}^\dagger$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{\hat{S}_3}(\{\hbar\}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}^\dagger$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F_{\hat{S}_3}(\{-\hbar\}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^\dagger$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Leftrightarrow \hat{S}_3 = \hbar\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - \hbar\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \hbar\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Now we are able to illustrate the relationship $F_{(\hat{A}_1,\hat{A}_2)}(R \times S) = F_{\hat{A}_1}(R) \circ F_{\hat{A}_2}(S)$ as follows.

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(0,0)\}) = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$
\begin{aligned}
&= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \end{pmatrix} \\
&= F_{\hat{\mathbf{S}}^2}(\{0\})F_{\hat{S}_3}(\{0\}), \\
F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(2\hbar^2,\hbar)\}) &= \begin{pmatrix} 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \\
&= \begin{pmatrix} 1&0&0&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \\
&= F_{\hat{\mathbf{S}}^2}(\{2\hbar^2\})F_{\hat{S}_3}(\{\hbar\}), \\
F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(2\hbar^2,0)\}) &= \frac{1}{2}\begin{pmatrix} 0&0&0&0 \\ 0&1&1&0 \\ 0&1&1&0 \\ 0&0&0&0 \end{pmatrix} \\
&= \begin{pmatrix} 1&0&0&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 0&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \end{pmatrix} \\
&= F_{\hat{\mathbf{S}}^2}(\{2\hbar^2\})F_{\hat{S}_3}(\{0\}), \\
F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(2\hbar^2,-\hbar)\}) &= \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \end{pmatrix} \\
&= \begin{pmatrix} 1&0&0&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&\frac{1}{2}&\frac{1}{2}&0 \\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \end{pmatrix} \\
&= F_{\hat{\mathbf{S}}^2}(\{2\hbar^2\})F_{\hat{S}_3}(\{-\hbar\}).
\end{aligned}
$$

EXAMPLE. For $n = 3$, through a similar process as in the previous example, we can construct

(1) the x-component of the first spin is

$$
\hat{S}_1(1) = \hat{\sigma}_1 \otimes I \otimes I = \frac{\hbar}{2}\begin{pmatrix} 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \end{pmatrix};
$$

(2) the x-component of the second spin is

$$
\hat{S}_1(2) = I \otimes \hat{\sigma}_1 \otimes I = \frac{\hbar}{2}\begin{pmatrix} 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \end{pmatrix};
$$

(3) the x-component of the third spin is

$$\hat{S}_1(3) = I \otimes I \otimes \hat{\sigma}_1 = \frac{\hbar}{2}\begin{pmatrix} 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \end{pmatrix};$$

Then the total x-component of the three spin system is

$$\hat{S}_1 = \hat{S}_1(1) + \hat{S}_1(2) + \hat{S}_1(3) = \frac{\hbar}{2}\begin{pmatrix} 0&1&1&0&1&0&0&0 \\ 1&0&0&1&0&1&0&0 \\ 1&0&0&1&0&0&1&0 \\ 0&1&1&0&0&0&0&1 \\ 1&0&0&0&0&1&1&0 \\ 0&1&0&0&1&0&0&1 \\ 0&0&1&0&1&0&0&1 \\ 0&0&0&1&0&1&1&0 \end{pmatrix}.$$

(1) the y-component of the first spin is

$$\hat{S}_2(1) = \hat{\sigma}_2 \otimes I \otimes I = \frac{\hbar}{2}\begin{pmatrix} 0&0&0&0&-i&0&0&0 \\ 0&0&0&0&0&-i&0&0 \\ 0&0&0&0&0&0&-i&0 \\ 0&0&0&0&0&0&0&-i \\ i&0&0&0&0&0&0&0 \\ 0&i&0&0&0&0&0&0 \\ 0&0&i&0&0&0&0&0 \\ 0&0&0&i&0&0&0&0 \end{pmatrix};$$

(2) the y-component of the second spin is

$$\hat{S}_2(2) = I \otimes \hat{\sigma}_2 \otimes I = \frac{\hbar}{2}\begin{pmatrix} 0&0&-i&0&0&0&0&0 \\ 0&0&0&-i&0&0&0&0 \\ i&0&0&0&0&0&0&0 \\ 0&i&0&0&0&0&0&0 \\ 0&0&0&0&0&0&-i&0 \\ 0&0&0&0&0&0&0&-i \\ 0&0&0&0&i&0&0&0 \\ 0&0&0&0&0&i&0&0 \end{pmatrix};$$

(3) the y-component of the third spin is

$$\hat{S}_2(3) = I \otimes I \otimes \hat{\sigma}_2 = \frac{\hbar}{2}\begin{pmatrix} 0&-i&0&0&0&0&0&0 \\ i&0&0&0&0&0&0&0 \\ 0&0&0&-i&0&0&0&0 \\ 0&0&i&0&0&0&0&0 \\ 0&0&0&0&0&-i&0&0 \\ 0&0&0&0&i&0&0&0 \\ 0&0&0&0&0&0&0&-i \\ 0&0&0&0&0&0&i&0 \end{pmatrix};$$

Then the total y-component of the three spin system is

$$\hat{S}_2 = \hat{S}_2(1) + \hat{S}_2(2) + \hat{S}_2(3) = \frac{\hbar}{2}\begin{pmatrix} 0&-i&-i&0&-i&0&0&0 \\ i&0&0&-i&0&-i&0&0 \\ i&0&0&-i&0&0&-i&0 \\ 0&i&i&0&0&0&0&-i \\ i&0&0&0&0&-i&-i&0 \\ 0&i&0&0&i&0&0&-i \\ 0&0&i&0&i&0&0&-i \\ 0&0&0&i&0&i&i&0 \end{pmatrix}.$$

(1) the z-component of the first spin is

$$\hat{S}_3(1) = \hat{\sigma}_3 \otimes I \otimes I = \frac{\hbar}{2}\begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&-1&0&0&0 \\ 0&0&0&0&0&-1&0&0 \\ 0&0&0&0&0&0&-1&0 \\ 0&0&0&0&0&0&0&-1 \end{pmatrix};$$

(2) the z-component of the second spin is

$$\hat{S}_3(2) = I \otimes \hat{\sigma}_3 \otimes I = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

(3) the z-component of the third spin is

$$\hat{S}_3(3) = I \otimes I \otimes \hat{\sigma}_3 = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

Then the total z-component of the three spin system is

$$\hat{S}_3 = \hat{S}_3(1) + \hat{S}_3(2) + \hat{S}_3(3) = \frac{\hbar}{2} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 \end{pmatrix}.$$

$$\hat{\mathbf{S}}^2 = \hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_3^2 = \frac{\hbar^2}{4} \begin{pmatrix} 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 4 & 0 & 4 & 0 & 0 & 0 \\ 0 & 4 & 7 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 7 & 4 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 \end{pmatrix}$$

The eigenvalues of $\hat{S}_3$ are $3, 1, 1, 1, -1, -1, -1, -3$ multiplied by $\frac{\hbar}{2}$ respectively. The eigenvalues of $\hat{\mathbf{S}}^2$ are $15, 15, 15, 15$ (which correspond to $S = \frac{3}{2}$) $3, 3, 3, 3$ (which correspond to $S = \frac{1}{2}$) multiplied by $\frac{\hbar^2}{4}$ respectively.

We can find $\tilde{X} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ such that

$$\hat{\mathbf{S}}^2 \tilde{X} = \tilde{X} \cdot \operatorname{diag}(15, 15, 15, 15, 3, 3, 3, 3) \times \frac{\hbar^2}{4}$$

$$\hat{S}_3 \tilde{X} = \tilde{X} \cdot \operatorname{diag}(3, -3, 1, -1, 1, 1, -1, -1) \times \frac{\hbar}{2}$$

as we show at Figure 1. The joint spectral decomposition for $\hat{\mathbf{S}}^2$ and$\hat{S}_3$ is:

$$F_{(\hat{\mathbf{S}}^2, \hat{S}_3)}(\{(\tfrac{\hbar^2 15}{4}, \tfrac{3\hbar}{2})\}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}^{\dagger} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$\hat{\mathbf{S}}^2$'s eigenvalue:$\hbar^2 S(S+1)$	$\hat{S}_3$'s eigenvalue:$m_S\hbar$	Unnormalized Simultaneous Eigenvector
$\frac{15\hbar^2}{4}, S=\frac{3}{2}$	$\frac{3\hbar}{2}, m_S=\frac{3}{2}$	$(1,0,0,0,0,0,0,0)^T$
$\frac{15\hbar^2}{4}, S=\frac{3}{2}$	$\frac{-3\hbar}{2}, m_S=-\frac{3}{2}$	$(0,0,0,0,0,0,0,1)^T$
$\frac{15\hbar^2}{4}, S=\frac{3}{2}$	$\frac{\hbar}{2}, m_S=\frac{1}{2}$	$(0,1,1,0,1,0,0,0)^T$
$\frac{15\hbar^2}{4}, S=\frac{3}{2}$	$\frac{-\hbar}{2}, m_S=-\frac{1}{2}$	$(0,0,0,1,0,1,1,0)^T$
$\frac{3\hbar^2}{4}, S=\frac{1}{2}$	$\frac{\hbar}{2}, m_S=\frac{1}{2}$	$(0,0,1,0,-1,0,0,0)^T$
$\frac{3\hbar^2}{4}, S=\frac{1}{2}$	$\frac{\hbar}{2}, m_S=\frac{1}{2}$	$(0,1,0,0,-1,0,0,0)^T$
$\frac{3\hbar^2}{4}, S=\frac{1}{2}$	$\frac{-\hbar}{2}, m_S=-\frac{1}{2}$	$(0,0,0,1,0,0,-1,0)^T$
$\frac{3\hbar^2}{4}, S=\frac{1}{2}$	$\frac{-\hbar}{2}, m_S=-\frac{1}{2}$	$(0,0,0,0,0,1,-1,0)^T$

FIGURE 1. Simultaneous Transposed Eigenvectors of $\hat{\mathbf{S}}^2$ and $\hat{S}_3$ of a 3-spin electron system.

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\frac{\hbar^2 15}{4},-\frac{3\hbar}{2})\}) = \begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\1\end{pmatrix}\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\1\end{pmatrix}^{\dagger} = \begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&1\end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\frac{\hbar^2 15}{4},\frac{\hbar}{2})\}) = \frac{1}{\sqrt{3}}\begin{pmatrix}0\\1\\1\\0\\1\\0\\0\\0\end{pmatrix}\frac{1}{\sqrt{3}}\begin{pmatrix}0\\1\\1\\0\\1\\0\\0\\0\end{pmatrix}^{\dagger} = \frac{1}{3}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&1&1&0&1&0&0&0\\0&1&1&0&1&0&0&0\\0&0&0&0&0&0&0&0\\0&1&1&0&1&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\frac{\hbar^2 15}{4},-\frac{\hbar}{2})\}) = \frac{1}{\sqrt{3}}\begin{pmatrix}0\\0\\0\\1\\0\\1\\1\\0\end{pmatrix}\frac{1}{\sqrt{3}}\begin{pmatrix}0\\0\\0\\1\\0\\1\\1\\0\end{pmatrix}^{\dagger} = \frac{1}{3}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&1&0&1&1&0\\0&0&0&0&0&0&0&0\\0&0&0&1&0&1&1&0\\0&0&0&1&0&1&1&0\\0&0&0&0&0&0&0&0\end{pmatrix}$$

$$F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\frac{\hbar^2 3}{4},\frac{\hbar}{2})\}) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\0\\-1\\0\\0\\0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\0\\-1\\0\\0\\0\end{pmatrix}^{\dagger} + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\0\\-1\\0\\0\\0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\0\\-1\\0\\0\\0\end{pmatrix}^{\dagger}$$

$$= \frac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&1&0&-1&0&0&0\\0&0&0&0&0&0&0&0\\0&0&-1&0&1&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&1&0&0&-1&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&-1&0&0&1&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&1&0&0&-1&0&0&0\\0&0&1&0&-1&0&0&0\\0&0&0&0&0&0&0&0\\0&-1&-1&0&2&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix}$$

$$
\begin{aligned}
F_{(\hat{\mathbf{S}}^2,\hat{S}_3)}(\{(\tfrac{\hbar^2 3}{4},-\tfrac{\hbar}{2})\}) &= \tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\1\\0\\0\\-1\\0\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\1\\0\\0\\-1\\0\end{pmatrix}^{\dagger}+\tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\0\\0\\1\\-1\\0\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\0\\0\\1\\-1\\0\end{pmatrix}^{\dagger}\\
&= \tfrac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&1&0&0&-1&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&-1&0&0&1&0\\0&0&0&0&0&0&0&0\end{pmatrix}+\tfrac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&1&-1&0\\0&0&0&0&0&-1&1&0\\0&0&0&0&0&0&0&0\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&1&0&0&-1&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&1&-1&0\\0&0&0&-1&0&-1&2&0\\0&0&0&0&0&0&0&0\end{pmatrix}
\end{aligned}
$$

3.1.4. State. The set of states is $\mathcal{S} = \{\hat{D} \in \mathbb{C}^{2^n\times 2^n} | \hat{D}^\dagger = \hat{D}, y^\dagger D y \geq 0, \forall y \in \mathcal{H}, \operatorname{tr}(\hat{D}) = 1\}$. The condition $y^\dagger \hat{D} y \geq 0, \forall y \in \mathcal{H}$ can be rephrased for Hermitian matrices to say that all eigenvalues of D are greater than or equal to 0.

Pure states can be represented by $\hat{D} = yy^\dagger$, where $y \in \mathbb{C}^{2^n}, \|y\| = 1$.

3.1.5. State Evolution. $\hat{H} = 0, \hat{U}(t) = I$ so $\hat{D}(t) = \hat{D}(0)$ unless the system is measured. This is because in most atoms and molecules (except in strong magnetic fields) spins do not interact except very weakly.

3.1.6. Predicting Measurement Outcomes.

Example. When $n = 1$, assume that we intend to noiselessly measure the z-component of the spin for a single electron in the pure state $\hat{D} = yy^\dagger, \|y\| = 1$ s.t. $\hat{S}_1 y = \frac{\hbar}{2} y$. We want to predict the distribution of the measured values of $\hat{S}_3$ on $\hat{D}$ as follows.

A solution for y is $y = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$, then $\hat{D} = yy^\dagger = \begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2}&\frac{1}{2}\end{pmatrix}$.

The probabilities that the outcome of the measurement of $\hat{S}_3 = \frac{\hbar}{2}\hat{\sigma}_3$ are $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ respectively are

$$
\begin{aligned}
\nu(\{\frac{\hbar}{2}\}) &= \operatorname{tr}(\hat{D} F_{\hat{S}_3}(\{\frac{\hbar}{2}\})\\
&= \operatorname{tr}(\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\ \frac{1}{2}&\frac{1}{2}\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix})
\end{aligned}
$$

$$= \frac{1}{2},$$

$$\nu(\{\frac{-\hbar}{2}\}) = \mathrm{tr}(\hat{D}F_{\hat{S}_3}(\{-\frac{\hbar}{2}\})$$

$$= \mathrm{tr}(\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

$$= \frac{1}{2}$$

Example. When $n = 2$, consider the pure state $D = yy^\dagger$, where y satisfies $\|y\| = 1$, $\hat{S}_3 y = 0$, and $\hat{\mathbf{S}}^2 y = 0$; a solution for y is $\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$, and

$$D = yy^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The probabilities that the outcome of the measurement of $\hat{S}_3(1)$ are $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ respectively are

$$\nu(\{\frac{\hbar}{2}\}) = \mathrm{tr}(\hat{D}F_{\hat{S}_3(1)}(\{\frac{\hbar}{2}\}))$$

$$= \mathrm{tr}(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix})$$

$$= \mathrm{tr}(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix})$$

$$= \frac{1}{2},$$

$$\nu(\{\frac{-\hbar}{2}\}) = \mathrm{tr}(\hat{D}F_{\hat{S}_3(1)}(\{-\frac{\hbar}{2}\}))$$

$$= \mathrm{tr}(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix})$$

$$= \mathrm{tr}(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix})$$

$$= \frac{1}{2}.$$

3.1.7. Updating the state after the measurement. As before, we take the measurement as noiseless. We define the state $\tilde{D}_{\mathrm{NL}}$ after the measurement in the discrete case to be $\tilde{D}_{\text{NL}} = \sum_{\lambda\in\sigma(\hat{A})} \hat{D}'_\lambda \tilde{\nu}(\{\lambda\})$, where $\hat{D}'_\lambda = \frac{F_{\hat{A}}(\{\lambda\})\hat{D}F_{\hat{A}}(\{\lambda\})}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))}$ and $\tilde{\nu}$ is the observed distribution of the measured values. We need to check that it is a density operator.

(1) To check $\hat{D}'_\lambda$ is of unit trace.

$$\begin{aligned}
\text{tr}(\hat{D}'_\lambda) &= \frac{\text{tr}(F_{\hat{A}}(\{\lambda\})\hat{D}F_{\hat{A}}(\{\lambda\}))}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \\
&= \frac{\text{tr}\left(\hat{D}F_{\hat{A}}(\{\lambda\})^2\right)}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \\
&= \frac{\text{tr}\left(\hat{D}F_{\hat{A}}(\{\lambda\})\right)}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \\
&= 1
\end{aligned}$$

(2) To check $\hat{D}'_\lambda$ is Hermitian.

$$\begin{aligned}
(\tilde{D}'_\lambda)^\dagger &= \frac{F_{\hat{A}}(\{\lambda\})^\dagger \hat{D}^\dagger F_{\hat{A}}(\{\lambda\})^\dagger}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \\
&= \frac{F_{\hat{A}}(\{\lambda\})\hat{D}F_{\hat{A}}(\{\lambda\})}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \\
&= \hat{D}'_\lambda
\end{aligned}$$

(3) To check $\hat{D}'_\lambda$ is nonnegative.

$$x^\dagger \hat{D}'_\lambda x = \frac{x^\dagger F_{\hat{A}}(\{\lambda\})\hat{D}(F_{\hat{A}}(\{\lambda\})x)}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} = \frac{(F_{\hat{A}}(\{\lambda\})x)^\dagger \hat{D}(F_{\hat{A}}(\{\lambda\})x)}{\text{tr}(\hat{D}F_{\hat{A}}(\{\lambda\}))} \geq 0,\ \forall x \in \mathbb{C}^{2^n}.$$

□

Example. We continue the the previous example in §3.1.6. Suppose that $\hat{S}_3(1)$ is measured to be $\frac{\hbar}{2}$, we want to find the updated state right after the measurement, it is enough to find the conditional density operator as follows.

$$D'_{\frac{\hbar}{2}} = \frac{F_{\hat{S}_3(1)}(\{\frac{\hbar}{2}\})\hat{D}F_{\hat{S}_3(1)}(\{\frac{\hbar}{2}\})}{\text{tr}(\hat{D}F_{\hat{S}_3(1)}(\{\frac{\hbar}{2}\}))}$$

$$= \frac{\begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix} \begin{pmatrix} 0&0&0&0\\ 0&\frac{1}{2}&0&0\\ 0&-\frac{1}{2}&0&0\\ 0&0&0&0 \end{pmatrix}}{\frac{1}{2}}$$

$$= \begin{pmatrix} 0&0&0&0\\ 0&1&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}$$

3.2. Spinless One Electron Systems: H and H_2^+

Because that the ratio of the mass of the proton and that of the electron is

$$\frac{\text{mass of proton}}{\text{mass of electron}} = \frac{1.6726 \cdot 10^{-27}\text{kg}}{9.1095 \cdot 10^{-31}\text{kg}} = 1.836 \cdot 10^3$$

we can treat the proton as classical object which is fixed in space. We will model electrons by the quantum formalism.

H, H_2^+ are systems with only one electron, and the spin of the electron does not affect any of the other observables, so we take the electron as spinless.

3.2.1. Hilbert Space. Consider one spinless electron in $\mathbb{R}^3$. This physical system is represented by the Hilbert space

$$\mathcal{H} = \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}),$$

where $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ is the set of equivalence classes of square integrable complex valued functions w.r.t. Lebesgue measure in $\mathbb{R}^3$, where the equivalence relation is equality almost everywhere.

3.2.2. Logic. $\mathcal{L} = \{\hat{P} \in \mathcal{B}(\mathcal{H}) | \hat{P}^\dagger = \hat{P} = \hat{P}^2\}$.

3.2.3. Observables. By the spectral theorem [**19**], for each $\hat{A} \in \mathcal{U}$, $\hat{A} = \int_{\mathbb{R}} \lambda \, dF_{\hat{A}}(\lambda)$, where $F_{\hat{A}} : \mathcal{B}(\mathbb{R}) \to \mathcal{L}$ s.t.

(1) $F_{\hat{A}}(\emptyset) = \hat{0}, F_{\hat{A}}(\mathbb{R}) = \hat{I}$;

(2) if $R_1 \cap R_2 = \emptyset$, then $F_{\hat{A}}(R_1) \perp F_{\hat{A}}(R_2)$;

(3) if $R = \cup_{j=1}^{\infty} R_j, R_j \cap R_k = \emptyset, j \neq k$, then $F_{\hat{A}}(\cup_{j=1}^{\infty} R_j) = \sum_{j=1}^{\infty} F_{\hat{A}}(R_j)$.

The observables of the electron systems of H and H_2^+ are the same except for the energy observable, they are:

(1) **position observables**

$$(\hat{x}_j f)(\mathbf{y}) = y_j f(\mathbf{y}),$$

where $\mathbf{y} = (y_1, y_2, y_3)^T \in \mathbb{R}^3,\ j = 1, 2, 3,\ f \in \mathcal{H}$;

(2) **momentum observables**

$$[\hat{p}_j(f)]\,(\mathbf{y}) = -i\hbar \frac{\partial f}{\partial y_j}(\mathbf{y}),$$

where $\mathbf{y} = (y_1, y_2, y_3)^T \in \mathbb{R}^3,\ j = 1, 2, 3,\ f \in \mathcal{H}$;

(3) **angular momentum observables**

$$\hat{\mathbf{J}} := \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2 \\ \hat{J}_3 \end{pmatrix} = \hat{\mathbf{x}} \times \hat{\mathbf{p}} = \begin{pmatrix} \hat{x}_2\hat{p}_3 - \hat{x}_3\hat{p}_2 \\ \hat{x}_3\hat{p}_1 - \hat{x}_1\hat{p}_3 \\ \hat{x}_1\hat{p}_2 - \hat{x}_2\hat{p}_1 \end{pmatrix},$$

where $\hat{\mathbf{x}} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix}$ and $\hat{\mathbf{p}} = \begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \hat{p}_3 \end{pmatrix}$. We also define $\hat{\mathbf{J}}^2 = (\hat{J}_1)^2 + (\hat{J}_2)^2 + (\hat{J}_3)^2$.

(4) **Energy observable of H (fixed nucleus approximation)** Assume that the coordinates of the nucleus is $(0, 0, 0)$,

$$(\widehat{H} f)(\mathbf{y}) = -\frac{\hbar^2}{2m}\triangle f(\mathbf{y}) - KZe^2 \frac{1}{(y_1^2 + y_2^2 + y_3^2)^{\frac{1}{2}}} f(\mathbf{y}),$$

where $f \in \mathcal{H}$, $\mathbf{y} \in \mathbb{R}^3$, $\triangle := (\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial y_3^2})$ and $-\frac{\hbar^2}{2m}\triangle$ is the operator of kinetic energy of the electron, $-KZe^2 \frac{1}{(y_1^2+y_2^2+y_3^2)^{\frac{1}{2}}}$ is the operator of potential energy of the electron due to its attraction to the nucleus; Ze is the charge on the nucleus; $K = \frac{1}{4\pi\varepsilon_0}, \varepsilon_0 = 8.854\cdots \times 10^{-12}\mathrm{F/m}$.

For the H_2^+ system, we assume that the two protons are at the positions $(0, 0, R/2), (0, 0, -R/2)$, where R is the distance between the two protons.

Energy observable of H_2^+ (fixed nucleus approximation)

$$(\widehat{H} f)(\mathbf{y}) = -\frac{\hbar^2}{2m}\triangle f(\mathbf{y}) - [\frac{Ke^2}{(y_1^2 + y_2^2 + (y_3 - R/2)^2)^{\frac{1}{2}}}$$

$$+\frac{Ke^2}{(y_1^2+y_2^2+(y_3+R/2)^2)^{\frac{1}{2}}}]f(\mathbf{y})+\frac{Ke^2}{R}f(\mathbf{y}),$$

where $f\in\mathcal{H}$, $\mathbf{y}\in\mathbb{R}^3$, $-\frac{\hbar^2}{2m}\triangle$ is the operator of kinetic energy of the electron; $-\left[\frac{Ke^2}{(y_1^2+y_2^2+(y_3-R/2)^2)^{\frac{1}{2}}}+\frac{Ke^2}{(y_1^2+y_2^2+(y_3+R/2)^2)^{\frac{1}{2}}}\right]$ is the operator of potential energy of the electron due to its attraction to both protons; $\frac{Ke^2}{R}$ is due to proton-proton repulsion.

Example. We want to check the spectral decomposition theorem for the position observables. Suppose: $\hat{A}=\hat{x}_j$, for some $j=1,2,3$. We claim that for all $R\in\mathcal{B}(\mathbb{R})$, we have $[F_{\hat{x}_j}(R)f](\mathbf{y})=\chi_R(y_j)f(\mathbf{y})$, where $\mathbf{y}=(y_1,y_2,y_3)\in\mathbb{R}^3$ and $f\in\mathcal{H}$.

To see that this is correct:

$$\begin{aligned}\left[\left(\int_{-\infty}^{\infty}\lambda\,dF_{\hat{A}}(\lambda)\right)f\right](\mathbf{y}) &= \int_{-\infty}^{\infty}\lambda\left[(\,dF_{\hat{x}_j})(\lambda)f\right](\mathbf{y})\\ &= \lim\sum_{i=-\infty}^{\infty}\lambda_i\{F_{\hat{x}_j}((\lambda_{i-1},\lambda_i])f\}(\mathbf{y})\\ &= \lim\sum_{i=-\infty}^{\infty}\lambda_i\chi_{(\lambda_{i-1},\lambda_i]}(y_j)f(\mathbf{y})\\ &= \lim\lambda_{i(y_j)}f(\mathbf{y})\\ &= y_jf(\mathbf{y})\\ &= (\hat{x}_jf)(\mathbf{y}),\end{aligned}$$

where $\{\lambda_i\}_{i\in\mathbb{Z}}$ determines a partition of $\mathbb{R}$, s.t. $\lambda_i<\lambda_{i+1}$; lim is performed in the sense that $\max_i(\lambda_{i+1}-\lambda_i)\to 0^+$; $i(y_j)$ is the index of the interval $(\lambda_{i-1},\lambda_i]$ which contains y_j.

Example. For another example of the spectral decomposition, consider the momentum observable $\hat{p}_j=-i\hbar\frac{\partial}{\partial y_j}$. We need the *Fourier transform* $\mathcal{F}:\mathcal{H}\to\mathcal{H}$ defined as

$$(\mathcal{F}f)(\mathbf{k})=\int_{\mathbb{R}^3}f(\mathbf{y})e^{-i\mathbf{k}\cdot\mathbf{y}}\,d^3\mathbf{y}$$

where $\mathbf{k} \in \mathbb{R}^3$, $d^3\mathbf{y} = dy_1\, dy_2\, dy_3$.

We are going to use the three facts about the Fourier transform [**33**] listed as below.

(1) $(\mathcal{F}f)(\mathbf{k})$ exists for almost every $\mathbf{k} \in \mathbb{R}^3$ and $\mathcal{F}f \in \mathcal{H}$;

(2) $\mathcal{F}(\hat{p}_j f)(\mathbf{k}) = \hbar k_j (\mathcal{F}f)(\mathbf{k})$;

(3) if $(\mathcal{F}^{-1}g)(\mathbf{y}) = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} g(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{y}}\, d^3\mathbf{k} = \mathcal{F}^{-1}\{g(\mathbf{y})\}$, then $\mathcal{F}^{-1}\mathcal{F}$ is the identity mapping on $\mathcal{H}$.

Definition. The spectral decomposition of the momentum observable is

$$[F_{\hat{p}_j}(R)f] := \mathcal{F}^{-1}\{\chi_R(\hbar k_j)(\mathcal{F}f)(\mathbf{k})\},$$

where $R \in \mathcal{B}(\mathbb{R}), f \in \mathcal{H}$. In different notation

$$\mathcal{F}[F_{\hat{p}_j}(R)f](\mathbf{k}) := \chi_R(\hbar k_j)(\mathcal{F}f)(\mathbf{k}).$$

To check that the above definition is correct, we need to show that

$$\hat{p}_j = \int_{-\infty}^{\infty} \lambda\, dF_{\hat{p}_j}(\lambda).$$

On one hand, $(\hat{p}_j f)(\mathbf{y}) = -i\hbar\frac{\partial f(\mathbf{y})}{\partial y_j}$ by the definition of momentum operator; on the other hand, by the property of Fourier transform on $\hat{p}_j$, we have:

$$\begin{aligned}
[\mathcal{F}(\hat{p}_j f)](\mathbf{k}) &= \hbar k_j (\mathcal{F}f)(\mathbf{k}) \\
&= \lim \sum_{i=-\infty}^{\infty} \lambda_i \chi_{(\lambda_{i-1},\lambda_i]}(\hbar k_j)(\mathcal{F}f)(\mathbf{k}) \\
&= \lim \sum_{i=-\infty}^{\infty} \lambda_i \mathcal{F}\{F_{\hat{p}_j}((\lambda_{i-1},\lambda_i])f\}(\mathbf{k}) \\
&= \mathcal{F}\{\lim \sum_{i=-\infty}^{\infty} \lambda_i F_{\hat{p}_j}((\lambda_{i-1},\lambda_i])f\}(\mathbf{k}) \\
&= \mathcal{F}\left[\int_{-\infty}^{\infty} \lambda\, dF_{\hat{p}_j}(\lambda) f\right](\mathbf{k}),
\end{aligned}$$

where $\{\lambda_i\}_{i\in\mathbb{Z}}$ determines a partition of $\mathbb{R}$, s.t. $\lambda_i < \lambda_{i+1}$, and lim is performed in the sense that $\max_i(\lambda_{i+1} - \lambda_i) \to 0^+$. So we have $\hat{p}_j f = \int_{-\infty}^{\infty} \lambda \, dF_{\hat{p}_j}(\lambda) f$.

3.2.4. Joint Spectral Decomposition of $\hat{H}, \hat{\mathbf{J}}^2, \hat{J}_3$ for H Atom. In spherical coordinates (r, ϕ, θ) [**6**], for H atom,

$$\hat{J}_3 = -i\hbar\frac{\partial}{\partial\theta}$$
$$\hat{\mathbf{J}}^2 = -\hbar^2[\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}(\sin\phi\frac{\partial}{\partial\phi}) + \frac{1}{\sin^2\phi}\frac{\partial^2}{\partial\phi^2}]$$
$$\hat{H} = -\frac{\hbar^2}{2m}(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{r^2\hbar}\hat{\mathbf{J}}^2) - \frac{KZe^2}{r}$$

Clearly we have

$$\hat{J}_3\hat{H} = \hat{H}\hat{J}_3$$
$$\hat{\mathbf{J}}^2\hat{H} = \hat{H}\hat{\mathbf{J}}^2$$
$$\hat{J}_3\hat{\mathbf{J}}^2 = \hat{\mathbf{J}}^2\hat{J}_3$$

By the simultaneous diagonalization theorem, there is $F_{(\hat{H},\hat{\mathbf{J}}^2,\hat{J}_3)} : \mathcal{B}(\mathbb{R}^3) \to \mathcal{L}$ such that

(1) $\hat{H} = \int_{\boldsymbol{\lambda}\in\mathbb{R}^3} \lambda_1 dF_{(\hat{H},\hat{\mathbf{J}}^2,\hat{J}_3)}(\boldsymbol{\lambda})$;
(2) $\hat{\mathbf{J}}^2 = \int_{\boldsymbol{\lambda}\in\mathbb{R}^3} \lambda_2 dF_{(\hat{H},\hat{\mathbf{J}}^2,\hat{J}_3)}(\boldsymbol{\lambda})$;
(3) $J_3 = \int_{\boldsymbol{\lambda}\in\mathbb{R}^3} \lambda_3 dF_{(\hat{H},\hat{\mathbf{J}}^2,\hat{J}_3)}(\boldsymbol{\lambda})$;

where $\boldsymbol{\lambda} = \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$.

According to [**6**] p.355, the following three quantum numbers index the the triples $\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$ on which $F_{(\hat{H},\hat{\mathbf{J}}^2,\hat{J}_3)}$ is concentrated:

principle quantum number $n = 1, 2, 3, \cdots$;

angular momentum quantum number $l = 0, 1, \cdots, n-1$

magnetic quantum number $m_l = -l, -l+1, \cdots, 0, \cdots, l-1, l$

l	m_l	Y_{l,m_l}
0	0	$(\frac{1}{4\pi})^{\frac{1}{2}}$
1	0	$(\frac{3}{4\pi})^{\frac{1}{2}}\cos\theta$
1	±1	$\mp(\frac{3}{8\pi})^{\frac{1}{2}}\sin\theta e^{\pm i\phi}$
2	0	$(\frac{5}{16\pi})^{\frac{1}{2}}(3\cos^2\theta-1)$
2	±1	$\mp(\frac{15}{8\pi})^{\frac{1}{2}}\cos\theta\sin\theta e^{\pm i\phi}$
2	±2	$(\frac{15}{32\pi})^{\frac{1}{2}}\sin^2\theta e^{\pm 2i\phi}$

TABLE 1. The Spherical Harmonics $Y_{l,m_l}(\theta,\phi)$

orbital	n	l	$R_{n,l}$
1s	1	0	$2(\frac{Z}{a_0})^{\frac{3}{2}}e^{-\frac{1}{2}\rho}$
$2s$	2	0	$\frac{1}{2\sqrt{2}}(\frac{Z}{a_0})^{\frac{3}{2}}(2-\rho)e^{-\frac{1}{2}\rho}$
$2p$	2	1	$\frac{1}{2\sqrt{6}}(\frac{Z}{a_0})^{\frac{3}{2}}\rho e^{-\frac{1}{2}\rho}$
3s	3	0	$\frac{1}{9\sqrt{3}}(\frac{Z}{a_0})^{\frac{3}{2}}(6-6\rho+\rho^2)e^{-\frac{1}{2}\rho}$
$3p$	3	1	$\frac{1}{9\sqrt{6}}(\frac{Z}{a_0})^{\frac{3}{2}}(4-\rho)\rho e^{-\frac{1}{2}\rho}$
$3d$	3	2	$\frac{1}{30\sqrt{3}}(\frac{Z}{a_0})^{\frac{3}{2}}\rho^2 e^{-\frac{1}{2}\rho}$

TABLE 2. Hydrogenic Radial Wavefunctions

The energy eigenvalue is related to n by

$$E_n = \frac{-Z^2e^2}{8n^2\pi\epsilon_0 a}$$

where $a = \frac{4\pi\epsilon_0\hbar^2}{me^2}$ is the Born radius; Ze is the charge on nucleus.

The eigenvalue of the squared orbital angular momentum operator $\hat{\mathbf{J}}^2$ is $l(l+1)\hbar^2$ and the eigenvalue of $\hat{J}_3$ is $m_l\hbar$.

The wavefunctions are products of radial and angular components:

$$\psi_{n,l,m_l} = R_{n,l}(r)Y_{l,m_l}(\theta,\phi)$$

Assume that $\|\psi_{n,l,m_l}\| = 1$; the spectral measure is

$$F_{(H,J^2,J_3)}(\{(E_n, l(l+1)\hbar^2, m_l\hbar)\}) = \psi_{n,l,m_l}\psi^\dagger_{n,l,m_l}.$$

The angular wavefunctions Y are the spherical harmonics (see table 1) and the radial wavefunctions R are the associated Laguerre functions (see table 2).

3.2.5. Joint Spectral Decomposition of $\hat{H}, \hat{J}_3$ for H_2^+ Ion. We define the prolate spherical coordinates of the electron (see [**1**], p.752) (ϕ, λ, μ) as follows. Suppose that coordinates of the protons of H_2^+ are $(0, 0, \frac{R}{2})$ and $(0, 0, -\frac{R}{2})$ respectively, and those of the electron are (x, y, z). Let ϕ be the angle between the part of the xz plane where $x > 0$ and the plane determined by the electron and the two protons of H_2^+.

Define $r_1 := \sqrt{x^2 + y^2 + (z + \frac{R}{2})^2}$, i.e. the distance between the electron and the proton at $(0, 0, -\frac{R}{2})$; define $r_2 := \sqrt{x^2 + y^2 + (z - \frac{R}{2})^2}$, i.e. the distance between the electron and the proton at $(0, 0, \frac{R}{2})$. Define

$$\lambda := \frac{r_1 + r_2}{R} = \frac{\sqrt{x^2 + y^2 + (z + \frac{R}{2})^2} + \sqrt{x^2 + y^2 + (z - \frac{R}{2})^2}}{R},$$

$$\mu := \frac{r_1 - r_2}{R} = \frac{\sqrt{x^2 + y^2 + (z + \frac{R}{2})^2} - \sqrt{x^2 + y^2 + (z - \frac{R}{2})^2}}{R},$$

then $\lambda \geq 1$ and $-1 \leq \mu \leq 1$.

It can be shown that

$$x = \frac{R}{2}\sqrt{(\lambda^2 - 1)(1 - \mu^2)}\cos\phi,$$
$$y = \frac{R}{2}\sqrt{(\lambda^2 - 1)(1 - \mu^2)}\sin\phi,$$
$$z = \frac{R}{2}\lambda\mu;$$

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{4}{R^2(\lambda^2 - \mu^2)}\{\frac{\partial}{\partial\lambda}[(\lambda^2 - 1)\frac{\partial}{\partial\lambda}]$$
$$+ \frac{\partial}{\partial\mu}[(1 - \mu^2)\frac{\partial}{\partial\mu}] + \frac{\partial}{\partial\phi}[(\frac{1}{\lambda^2 - 1} + \frac{1}{1 - \mu^2})\frac{\partial}{\partial\phi}]\} + \frac{Ke^2}{R}\{-\frac{4\lambda}{\lambda^2 - \mu^2} + 1\},$$
$$\hat{J}_3 = -i\hbar\frac{\partial}{\partial\phi}.$$

It is clear that $\hat{H}\hat{J}_3 = \hat{J}_3\hat{H}$. Thus we have the joint spectral decomposition of $(\hat{H}, \hat{J}_3)$ in the bound state sector, which is supported on the set $\{(E(n,l,m;R), m\hbar) \in \mathbb{R}^2 | n,l,m \in \mathbb{Z}, n > l \geq |m|\}$. $E(n,l,m;R)$ are well-studied and tabulated functions, see Bates et.al.[**10**]. For each $(n,l,m) \in \mathbb{Z}^3$, s.t. $n > l \geq |m|$, there are functions $\psi(n,l,m,R;\lambda,\mu,\phi) \in \mathcal{L}^2(\mathbb{R}^3,\mathbb{C})$ of the form $\psi = \Lambda(n,l,m,R;\lambda)M(n,l,m,R;\mu)e^{im\phi}$, s.t. $\|\psi\|^2 = 1$ and where the functions Λ and M are also well-studied and tabulated. We have

$$F_{(\hat{H},\hat{J}_3)}(\{(E, m\hbar)\}) = \sum_{n>l\geq|m|, E(n,l,m;R)=E} \psi(n,l,m;R)\psi(n,l,m;R)^\dagger$$

3.2.6. State. (The theory is the same as that in section 2.3.4)

3.2.7. State Evolution. (The theory is the same as that in section 2.3.5.)

3.2.8. Predicting Measurement Outcomes.

Example. Suppose the system is in the pure state $\hat{D} = ff^\dagger$, $\|f\| = 1$ and we noiselessly measure position $\hat{x}_j$, $j = 1,2,3$. Then the predicted probability that the measured values will be in the interval $(a,b]$ is:

$$\begin{aligned}
\nu_{\hat{x}_j}((a,b]) &= \mathrm{prob}(\hat{x}_j \in (a,b] | D) \\
&= \mathrm{tr}(\hat{D}F_{\hat{x}_j}((a,b])) \\
&= \sum_{n=1}^{\infty}(e_n, \hat{D}F_{\hat{x}_j}((a,b])e_n) \\
&= \sum_{n=1}^{\infty}(e_n, f(f, F_{\hat{x}_j}((a,b])e_n)) \\
&= \sum_{n=1}^{\infty}(e_n, f)\int_{\mathbb{R}^3} \overline{f(\mathbf{y})}\chi_{(a,b]}(y_j)e_n(\mathbf{y})\,d^3\mathbf{y} \\
&= \sum_{n=1}^{\infty}(e_n, f)\int_{y_j\in(a,b]} \overline{f(\mathbf{y})}e_n(\mathbf{y})\,d^3\mathbf{y} \\
&= (f,f)\int_{y_j\in(a,b]} \overline{f(\mathbf{y})}f(\mathbf{y})\,d^3\mathbf{y} + \sum_{n=2}^{\infty}(e_n, e_1)\int_{y_j\in(a,b]} \overline{f(\mathbf{y})}e_n(\mathbf{y})\,d^3\mathbf{y},
\end{aligned}$$

where we set $e_1 = f$.

$$= \int_{y_j \in (a,b]} \overline{f(\mathbf{y})} f(\mathbf{y})\, d^3\mathbf{y},$$

where $\{e_n\}_{n=1}^{\infty}$ is any orthonormal basis of the Hilbert space $L^2(\mathbb{R})$ with $e_1 = f$.

Using similar technique as above, we find the joint probability of simultaneous measurement

$$\begin{aligned}
&\text{prob}(\hat{x}_1 \in (a_1, b_1], \hat{x}_2 \in (a_2, b_2], \hat{x}_3 \in (a_3, b_3]|D) \\
&= \text{tr}\left[\hat{D} F_{(\hat{x}_1, \hat{x}_2, \hat{x}_3)}((a_1, b_1] \times (a_2, b_2] \times (a_3, b_3])\right] \\
&= \int_{y_1 \in (a_1, b_1]} \int_{y_2 \in (a_2, b_2]} \int_{y_3 \in (a_3, b_3]} |f(\mathbf{y})|^2\, dy_3\, dy_2\, dy_1
\end{aligned}$$

This expresses the Born interpretation of Schrödinger's wave function: If the normalized wavefunction of a particle is ψ, then the probability of finding the particle in an infinitesimal volume $d\tau = dxdydz$ at the point $\mathbf{r}$ is $|\psi(\mathbf{r})|^2\, d\tau$.

Example. Suppose the system is in the pure state $\hat{D} = ff^{\dagger}$ and we noiselessly measure $\hat{p}_j$, $j = 1, 2, 3$. Then the predicted probability that the measured values will be in the interval $(a, b]$ is:

$$\begin{aligned}
\nu_{\hat{p}_j}((a, b]) &= \text{prob}(\hat{p}_j \in (a, b]|\hat{D}) \\
&= \text{tr}(\hat{D} F_{\hat{p}_j}((a, b])) \\
&= \sum_{n=1}^{\infty} (e_n, \hat{D} F_{\hat{p}_j}((a, b,])e_n) \\
&= \sum_{n=1}^{\infty} (e_n, f(f, F_{\hat{p}_j}((a, b])e_n)) \\
&= \sum_{n=1}^{\infty} (e_n, f)(f, F_{\hat{p}_j}((a, b])e_n) \\
&= \sum_{n=1}^{\infty} (e_n, f) \frac{1}{(2\pi)^3} (\mathcal{F}f, \mathcal{F}\{F_{\hat{p}_j}((a, b])e_n\})
\end{aligned}$$

$$= (f, f)\frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\overline{(\mathcal{F}f)(\mathbf{k})}\chi_{(a,b]}(\hbar k_j)(\mathcal{F}f)(\mathbf{k})\, d^3\mathbf{k}$$

$$= \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}|(\mathcal{F}f)(\mathbf{k})|^2\chi_{(a,b]}(\hbar k_j)\, d^3\mathbf{k},$$

where $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of the Hilbert space $L^2(\mathbb{R})$ with $e_1 = f$.

Similarly

$$\text{prob}(\hat{p}_1 \in (a_1, b_1], \hat{p}_2 \in (a_2, b_2], \hat{p}_3 \in (a_3, b_3]|\hat{D})$$

$$= \frac{1}{(2\pi)^3}\int_{\hbar k_1\in(a_1,b_1]}\int_{\hbar k_2\in(a_2,b_2]}\int_{\hbar k_3\in(a_3,b_3]}|\mathcal{F}f(\mathbf{k})|^2\, dk_3\, dk_2\, dk_1.$$

So $\frac{1}{(2\pi)^3}|\mathcal{F}f(\mathbf{k})|^2\, dk_3\, dk_2\, dk_1$ can be interpreted as the probability that the measured value of the momentum $\mathbf{P}$ will be such that $\frac{\mathbf{P}}{\hbar}$ will be in a small region around $\mathbf{k}$ of size $dk_3\, dk_2\, dk_1$.

Example. Consider a hydrogen atom whose electron is in a pure state $D = \psi\psi^\dagger$ where $\psi(r, \phi, \theta) = ce^{-ar^2}$, $c, a > 0$, $\|\psi\| = 1$. Suppose we noiselessly measure the total energy observable $\hat{H}$. What is the predicted distribution of the outcomes of this measurement?

The spectral measure in the bound state sector for $\hat{H}$ is as follows. for $n = 1, 2, \ldots$

$$F_{\hat{H}}(\{E_n\}) = \sum_{l=0}^{n-1}\sum_{m_l=-l}^{l}\psi_{n,l,m_l}\psi_{n,l,m_l}^\dagger.$$

Let $\nu_{\hat{H}}(\{E_n\})$ denote the predicted probability that the energy measurement yields the value E_n, $n \geq 1$. Then

$$\nu_{\hat{H}}(\{E_n\}) = \text{tr}(\hat{D}F_{\hat{H}}(\{E_n\}))$$

$$= \sum_{j=1}^{\infty}(e_j, \hat{D}F_{\hat{H}}(\{E_n\})e_j)$$

$$= \sum_{j=1}^{\infty}(e_j, \psi)(\psi, \sum_{l=0}^{n-1}\sum_{m_l=-l}^{l}\psi_{n,l,m_l}\psi_{n,l,m_l}^\dagger e_j)$$

$$= \sum_{j=1}^{\infty}(e_j, \psi)\sum_{l=0}^{n-1}\sum_{m_l=-l}^{l}(\psi, \psi_{n,l,m_l})(\psi_{n,l,m_l}, e_j).$$

Since ψ depends only on r, and the spherical harmonics $Y_{l,m_l}(\theta,\phi)$ are orthogonal, and $Y_{0,0}(\theta,\phi) = \frac{1}{\sqrt{4\pi}}$, we have that $(\psi, \psi_{n,l,m_l}) = 0$ unless $l = 0$ and $m_l = 0$. The orthonormal basis $\{e_j\}_{j=1}^{\infty}$ can be chosen arbitrarily, so let $e_1 = \psi$, so that $(e_j, \psi) = 0$ for all $j \geq 2$. Thus

$$\nu_{\hat{H}}(\{E_n\}) = |(\psi, \psi_{n,0,0})|^2, \qquad n \geq 1.$$

3.2.9. Updating the state after the measurement. The theory behind this discussion is in section §2.2.7. Here we will illustrate that theory in the simplest infinite dimensional case.

Example. Here we resume the study of an example from the previous section, namely that of an energy measurement of the electron of a hydrogen atom in the state $D = \psi\psi^\dagger$ prior to the measurement. Suppose as a result of the measurement the energy is found to be E_1 with probability 1, i.e. $\tilde{\nu} = \delta_{E_1}$. What is the updated state of the electron?

Since the energy operator $\hat{H}$ has an infinite discrete spectrum $\{E_n \mid n \geq 1\}$ accumulating at 0 and a continuous spectrum $[0,\infty)$, we must take the noiseless limit of our formulae for the updated state after a noisy measurement. Let ϵ^2 be the variance of the noise and $\delta_{E_1} * \alpha_\epsilon^2$ the observed distribution of noisily measured values. The conditional density operator at energy E is

$$\begin{aligned} D'_E &= \frac{\alpha_\epsilon(E\hat{1} - \hat{H})\psi\psi^\dagger\alpha_\epsilon(E\hat{1} - \hat{H})^\dagger}{\operatorname{tr}\left[\alpha_\epsilon(E\hat{1} - \hat{H})\psi\psi^\dagger\alpha_\epsilon(E\hat{1} - \hat{H})^\dagger\right]} \\ &= \frac{\alpha_\epsilon(E\hat{1} - \hat{H})\psi[\alpha_\epsilon(E\hat{1} - \hat{H})\psi]^\dagger}{\|\alpha_\epsilon(E\hat{1} - \hat{H})\psi\|^2} \end{aligned}$$

In the above we have used the identity $\operatorname{tr}\phi\phi^\dagger = \|\phi\|^2$. The vector $\alpha_\epsilon(E\hat{1} - \hat{H})\psi$ is computed as follows.

$$\begin{aligned} \alpha_\epsilon(E\hat{1} - \hat{H})\psi &= \sum_{n=1}^{\infty} \alpha_\epsilon(E - E_n) F_{\hat{H}}(\{E_n\})\psi + \int_0^\infty \alpha_\epsilon(E - E') F_{\hat{H}}(\,dE')\psi \\ &= \alpha_\epsilon(E - E_1)\psi_{1,0,0}(\psi_{1,0,0}, \psi) + \psi_{E,\epsilon}, \end{aligned}$$

where

$$\psi_{E,\epsilon} = \sum_{n=2}^{\infty} \alpha_\epsilon(E - E_n)\psi_{n,0,0}(\psi_{n,0,0}, \psi) + \int_0^\infty \alpha_\epsilon(E - E')F_{\hat{H}}(\,dE')\psi.$$

Its norm can be computed as follows.

$$\begin{aligned}\|\alpha_\epsilon(E\hat{1} - \hat{H})\psi\|^2 &= \sum_{n=1}^{\infty} \alpha_\epsilon(E - E_n)^2 \|F_{\hat{H}}(\{E_n\})\psi\|^2 + \int_0^\infty \alpha_\epsilon(E - E')^2 \|F_{\hat{H}}(\,dE')\psi\|^2 \\ &= \alpha_\epsilon(E - E_1)^2 |(\psi_{1,0,0}, \psi)|^2 + \|\psi_{E,\epsilon}\|^2.\end{aligned}$$

The updated state after the noisy measurement is:

$$\begin{aligned}\tilde{D} &= \int_{-\infty}^{\infty} D'_E \alpha_\epsilon(E - E_1)^2\, dE \\ &= \int_{-\infty}^{\infty} \frac{\alpha_\epsilon(E - E_1)^2 |(\psi_{1,0,0}, \psi)|^2 \psi_{1,0,0}\psi_{1,0,0}^\dagger}{\alpha_\epsilon(E - E_1)^2 |(\psi_{1,0,0}, \psi)|^2 + \|\psi_{E,\epsilon}\|^2} \alpha_\epsilon(E - E_1)^2\, dE \\ &\quad + \int_{-\infty}^{\infty} \frac{\alpha_\epsilon(E - E_1)(\psi_{1,0,0}, \psi)[\psi_{1,0,0}\psi_{E,\epsilon}^\dagger + \psi_{E,\epsilon}\psi_{1,0,0}^\dagger] + \psi_{E,\epsilon}\psi_{E,\epsilon}^\dagger}{\alpha_\epsilon(E - E_1)^2 |(\psi_{1,0,0}, \psi)|^2 + \|\psi_{E,\epsilon}\|^2} \alpha_\epsilon(E - E_1)^2\, dE.\end{aligned}$$

Both of these integrals should be split over the intervals $(-\infty, \frac{1}{2}(E_1+E_2)]$ and $(\frac{1}{2}(E_1+E_2), \infty)$. If $E \leq \frac{1}{2}(E_1 + E_2)$ then $\|\psi_{E,\epsilon}\| \to 0$ as $\epsilon \to 0^+$. If $E > \frac{1}{2}(E_1 + E_2)$ then $\|\psi_{E,\epsilon}\|$ remains bounded and $\alpha_\epsilon(E - E_1) \to 0$ as $\epsilon \to 0^+$. Thus the noiseless limit of $\tilde{D}$ is $\psi_{1,0,0}\psi_{1,0,0}^\dagger$. This result agrees with the Neumann-Lüder's formulae derived in section 2.2.7:

$$\tilde{D}_{NL} = \frac{F_{\hat{H}}(\{E_1\})DF_{\hat{H}}(\{E_1\})}{\mathrm{tr}(DF_{\hat{H}}(\{E_1\}))} = \frac{\psi_{1,0,0}\psi_{1,0,0}^\dagger \psi\psi^\dagger \psi_{1,0,0}\psi_{1,0,0}^\dagger}{|(\psi, \psi_{1,0,0})|^2} = \psi_{1,0,0}\psi_{1,0,0}^\dagger.$$

Thus measurement of the total electronic energy followed by a selection of the subsystems which yield the answer E_1 is an effective way of preparing an ensemble of hydrogen atoms whose electrons are in the pure state $\psi_{1,0,0}\psi_{1,0,0}^\dagger$, i.e. the "ground state".

3.3. Two Electron System: H_2

H_2 contains 2 electrons. The spins of these electrons will affect their spatial distribution, so we need to consider the spins of the system.

3.3.1. Hilbert Space. Let $\mathcal{H}_e := \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \otimes \mathbb{C}^2$ be the Hilbert space of a single electron with spin. Then the Hilbert space of two electron system of H_2 is represented by

$$\mathcal{H} = \mathcal{H}_e \wedge \mathcal{H}_e$$

Elements $f \in \mathcal{H}_e$ are (equivalence classes of) mappings $f : \mathbb{R}^3 \times \{0, 1\} \to \mathbb{C}$ such that $\sum_{\sigma=0}^{1} \int_{\mathbb{R}^3} |f(\mathbf{y}, \sigma)|^2 \, d^3\mathbf{y} < \infty$.

Elements $f \in \mathcal{H} = \mathcal{H}_e \wedge \mathcal{H}_e$ are (equivalence classes of) mappings $f : \mathbb{R}^3 \times \mathbb{R}^3 \times \{0, 1\}^2 \to \mathbb{C}$ such that $\sum_{\sigma_1=0}^{1} \sum_{\sigma_2=0}^{1} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2)|^2 \, d^3\mathbf{y}_1 d^3\mathbf{y}_2 < \infty$ and $f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) = -f(\mathbf{y}_2, \mathbf{y}_1, \sigma_2, \sigma_1)$.

3.3.2. Logic. $\mathcal{L} = \{\hat{P} \in \mathcal{B}(\mathcal{H}) | \hat{P}^\dagger = \hat{P} = \hat{P}^2\}$.

3.3.3. Observables. The observables of the H_2 system include:

(1) **position observables**

$$[\hat{x}_j(i) f](\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) = y_j^i f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2),$$

where for $i = 1, 2$, $\mathbf{y}_i = (y_1^i, y_2^i, y_3^i)^T \in \mathbb{R}^3$ is the position of the ith electron; $f \in \mathcal{H}$.

(2) **momentum observables**

$$[\hat{p}_j(i) f](\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) = -i\hbar \frac{\partial f}{\partial y_j^i}(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2).$$

(3) **angular momentum observables**

$$\hat{\mathbf{J}}(i) := \begin{pmatrix} \hat{J}_1(i) \\ \hat{J}_2(i) \\ \hat{J}_3(i) \end{pmatrix} := \hat{\mathbf{x}}(i) \times \hat{\mathbf{p}}(i) = \begin{pmatrix} \hat{x}_2(i)\hat{p}_3(i) - \hat{x}_3(i)\hat{p}_2(i) \\ \hat{x}_3(i)\hat{p}_1(i) - \hat{x}_1(i)\hat{p}_3(i) \\ \hat{x}_1(i)\hat{p}_2(i) - \hat{x}_2(i)\hat{p}_1(i) \end{pmatrix},$$

where for $i=1,2$, $\hat{\mathbf{x}}(i) := (\hat{x}_1(i), \hat{x}_2(i), \hat{x}_3(i))^T$, $\hat{\mathbf{p}}(i) := (\hat{p}_1(i), \hat{p}_2(i), \hat{p}_3(i))^T$.

$$\hat{\mathbf{J}}_j := \hat{\mathbf{J}}_j(1) + \hat{\mathbf{J}}_j(2), j = 1,2,3.$$

(4) **Spin observables** We represent $f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2)$ in vector form as $\begin{pmatrix} f(\mathbf{y}_1,\mathbf{y}_2,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,1,1) \end{pmatrix}$, i.e. a $\mathbb{C}^4$-valued function of $(\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{R}^3 \times \mathbb{R}^3$. All the spin operators $\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{\mathbf{S}}^2$ can be written as 4×4 matrices as in the $n = 2$ example of §3.1.3. For example $\hat{\mathbf{S}}^2 = \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Thus $\hat{\mathbf{S}}^2 f$ in vector form is

$$\begin{pmatrix} \hat{\mathbf{S}}^2 f(\mathbf{y}_1,\mathbf{y}_2,0,0) \\ \hat{\mathbf{S}}^2 f(\mathbf{y}_1,\mathbf{y}_2,0,1) \\ \hat{\mathbf{S}}^2 f(\mathbf{y}_1,\mathbf{y}_2,1,0) \\ \hat{\mathbf{S}}^2 f(\mathbf{y}_1,\mathbf{y}_2,1,1) \end{pmatrix} = \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} f(\mathbf{y}_1,\mathbf{y}_2,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,1,1) \end{pmatrix} = \hbar^2 \begin{pmatrix} 2f(\mathbf{y}_1,\mathbf{y}_2,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,0,1)+f(\mathbf{y}_1,\mathbf{y}_2,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,0,1)+f(\mathbf{y}_1,\mathbf{y}_2,1,0) \\ 2f(\mathbf{y}_1,\mathbf{y}_2,1,1) \end{pmatrix}.$$

(5) **Energy observable (fixed nucleus approximation)**

$$\begin{aligned} (\widehat{H} f)(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) = & -\frac{\hbar^2}{2m}(\triangle_1 + \triangle_2) f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) \\ & + \frac{Ke^2}{\|\mathbf{y}_1 - \mathbf{y}_2\|} f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) \\ & + Ke^2 \frac{1}{\|\mathbf{R}_1 - \mathbf{R}_2\|} f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2) \\ & - Ke^2 \sum_{i,j=1}^{2} \frac{1}{\|\mathbf{R}_i - \mathbf{y}_j\|} f(\mathbf{y}_1, \mathbf{y}_2, \sigma_1, \sigma_2), \end{aligned}$$

where $\mathbf{R}_1 = \begin{pmatrix} 0 \\ 0 \\ R/2 \end{pmatrix}, \mathbf{R}_2 = \begin{pmatrix} 0 \\ 0 \\ -R/2 \end{pmatrix} \in \mathbb{R}^3$ and $-\frac{\hbar^2}{2m}(\triangle_1 + \frac{\hbar^2}{2m}\triangle_2)$ is the kinetic energy operator; for $i = 1,2$ representing the 1st and the 2nd electron, $\triangle_i := (\frac{\partial^2}{\partial(y_1^i)^2} + \frac{\partial^2}{\partial(y_2^i)^2} + \frac{\partial^2}{\partial(y_3^i)^2})$ is the Laplacian for the ith electron's position; $Ke^2 \frac{1}{\|\mathbf{R}_1 - \mathbf{R}_2\|}$ is due to proton-proton repulsion; $\frac{Ke^2}{\|\mathbf{y}_1 - \mathbf{y}_2\|}$ is due to electron-electron repulsion; $-Ke^2 \sum_{i,j=1}^{2} \frac{1}{\|\mathbf{R}_i - \mathbf{y}_j\|}$ is the potential energy operator of the electrons due to their attraction to protons.

FACT. For H_2 system,

$$[\widehat{H}, \hat{J}_3] = 0, [\hat{J}_3, \hat{S}_3] = 0, [\hat{J}_3, \hat{\mathbf{S}}^2] = 0$$

3.3.4. State. (The theory is the same as that in section 2.3.4.)

In numerical approximation of the electronic ground state, fix $R > 0$, so that $\hat{H}$ is a fixed operator. We want to find the smallest eigenvalue E of $\hat{H}$.

$$E = \min_{\psi \in \mathcal{H}, \psi \neq 0} \frac{(\psi, \hat{H}\psi)}{(\psi, \psi)}$$

The minimum is attained at some eigenfunction ψ associated with E.

We use the idea of ansatz to do the approximation. An *ansatz* for ψ in $\mathcal{H}$ is a subset $S \subset \mathcal{H}$ where elements of S can be more explicitly described than the general element in $\mathcal{H}$.

Given an ansatz $S \subset \mathcal{H}$, the approximate eigenvalue of the ground state is

$$E_S = \min_{\psi \in S, \psi \neq 0} \frac{(\psi, \hat{H}\psi)}{(\psi, \psi)}$$

The approximate eigenfunction ψ_S is an element of S when the minimum is attained. If S is a "big enough" subset of $\mathcal{H}$, then $E_S \approx E, \psi_S \approx \psi$. Computationally S should be parameterized by a finite list of parameters, so that S is a finite dimensional manifold in the infinite dimensional space $\mathcal{H}$.

For instance, we define the generalized valence bond ansatz as

$$S_{\text{GVB}} = \{\psi \in \mathcal{H}_e \wedge \mathcal{H}_e | \psi = \mathcal{A}[\varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\Sigma(\sigma_1, \sigma_2)],$$
$$\text{where } \varphi_1, \varphi_2 \in V, \Sigma \in \mathbb{C}^4\},$$

where V is a finite dimensional subspace of $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$; φ_1 is the spatial orbital for the first electron; φ_2 is the spatial orbital for the second electron space; $\sigma_1, \sigma_2 \in \{0, 1\}$; Σ can be expanded in joint eigenfunction of $\hat{\mathbf{S}}^2, \hat{S}_3$ and represents the spin state as a resonance of the four possible joint eigenstates.

Another example of an ansatz is the Hartree-Fock ansatz, denoted S_{HF}, where ψ can be explicitly represented as

$$\psi(\mathbf{x}_1, \sigma, \mathbf{x}_2, \sigma) = \det \begin{pmatrix} \varphi_1(\mathbf{x}_1)\alpha(\sigma) & \varphi_1(\mathbf{x}_2)\alpha(\sigma) \\ \varphi_2(\mathbf{x}_1)\beta(\sigma) & \varphi_2(\mathbf{x}_2)\beta(\sigma) \end{pmatrix}$$

where $\alpha(\sigma) = \begin{cases} 1, & \text{if } \sigma = 0 \\ 0, & \text{if } \sigma = 1 \end{cases}, \beta(\sigma) = \begin{cases} 0, & \text{if } \sigma = 0 \\ 1, & \text{if } \sigma = 1 \end{cases}.$

3.3.5. State Evolution. (The theory is the same as that in section 2.3.5.)

3.3.6. Predicting Measurement Outcomes. (The theory is the same as that in section 2.3.6.)

3.3.7. Updating the state after the measurement. (The theory is the same as that in section 2.3.7.)

3.4. Three Electron System H_3

3.4.1. Hilbert Space. H_3 contains 3 electrons. The electron system is represented by Hilbert space

$$\mathcal{H} = \mathcal{H}_e \wedge \mathcal{H}_e \wedge \mathcal{H}_e$$

Elements $f \in \mathcal{H} = \mathcal{H}_e \wedge \mathcal{H}_e \wedge \mathcal{H}_e$ are (equivalence classes of) mappings $f : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \{0,1\}^3 \to \mathbb{C}$ such that

$$\sum_{\sigma_1=0}^{1} \sum_{\sigma_2=0}^{1} \sum_{\sigma_3=0}^{1} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |f(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3)|^2 \, d^3\mathbf{y}_1 d^3\mathbf{y}_2 d^3\mathbf{y}_3 < \infty$$

and

$$f(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3) = f(\mathbf{y}_3, \mathbf{y}_1, \mathbf{y}_2, \sigma_3, \sigma_1, \sigma_2) = f(\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_1, \sigma_2, \sigma_3, \sigma_1)$$

$$= -f(\mathbf{y}_2, \mathbf{y}_1, \mathbf{y}_3, \sigma_2, \sigma_1, \sigma_3) = -f(\mathbf{y}_1, \mathbf{y}_3, \mathbf{y}_2, \sigma_1, \sigma_3, \sigma_2) = -f(\mathbf{y}_3, \mathbf{y}_2, \mathbf{y}_1, \sigma_3, \sigma_2, \sigma_1).$$

3.4.2. Logic. $\mathcal{L} = \{\hat{P} \in \mathcal{B}(\mathcal{H}) | \hat{P}^\dagger = \hat{P} = \hat{P}^2\}$.

3.4.3. Observables. The observable of the H_3 system includes:

(1) **position observables**

$$[\hat{x}_j(i)f](\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3) = y_j^i f(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3),$$

where for $i = 1, 2, 3$, $\mathbf{y}_i = (y_1^i, y_2^i, y_3^i)^T \in \mathbb{R}^3$ is the position of the ith electron; $f \in \mathcal{H}$.

(2) **momentum observables**

$$[\hat{p}_j(i)f](\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3) = -i\hbar \frac{\partial f}{\partial y_j^i}(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3).$$

(3) **angular momentum observables**

$$\hat{\mathbf{J}}(i) := \begin{pmatrix} \hat{J}_1(i) \\ \hat{J}_2(i) \\ \hat{J}_3(i) \end{pmatrix} := \hat{\mathbf{x}}(i) \times \hat{\mathbf{p}}(i) = \begin{pmatrix} \hat{x}_2(i)\hat{p}_3(i) - \hat{x}_3(i)\hat{p}_2(i) \\ \hat{x}_3(i)\hat{p}_1(i) - \hat{x}_1(i)\hat{p}_3(i) \\ \hat{x}_1(i)\hat{p}_2(i) - \hat{x}_2(i)\hat{p}_1(i) \end{pmatrix}$$

where for $i = 1, 2, 3$, $\hat{\mathbf{x}}(i) := (\hat{x}_1(i), \hat{x}_2(i), \hat{x}_3(i))^T$, $\hat{\mathbf{p}}(i) := (\hat{p}_1(i), \hat{p}_2(i), \hat{p}_3(i))^T$.

(4) **Spin observables** We represent $f(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \sigma_1, \sigma_2, \sigma_3)$ in vector form as $\begin{pmatrix} f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,1) \end{pmatrix}$, i.e. a $\mathbb{C}^8$-valued function of $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$.
All the spin operators $\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{\mathbf{S}}^2$ can be written as 8×8 matrices as in the $n = 3$ example of §3.1.3. For example $\hat{\mathbf{S}}^2 = \frac{\hbar^2}{4} \begin{pmatrix} 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 4 & 0 & 4 & 0 & 0 & 0 \\ 0 & 4 & 7 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 7 & 4 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 \end{pmatrix}$.

Thus $\hat{\mathbf{S}}^2 f$ in vector form is

$$\begin{pmatrix} (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,0) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ (\hat{\mathbf{S}}^2 f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,1) \end{pmatrix}$$

$$=\frac{\hbar^2}{4}\begin{pmatrix} 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 4 & 0 & 4 & 0 & 0 & 0 \\ 0 & 4 & 7 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 7 & 4 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 \end{pmatrix}\begin{pmatrix} f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,1) \end{pmatrix}$$

$$=\frac{\hbar^2}{4}\begin{pmatrix} 15f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,0) \\ 7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ 4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1)+7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ 7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ 4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,0,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,0)+7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,0) \\ 4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1)+7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ 4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,0,1,1)+4f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,0,1)+7f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,0) \\ 15f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,1,1,1) \end{pmatrix}.$$

(5) **Energy observable(fixed nucleus approximation)**

$$\begin{aligned}
&(\widehat{H}f)(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,\sigma_1,\sigma_2,\sigma_3) \\
=&-\frac{\hbar^2}{2m}(\triangle_1+\triangle_2+\triangle_3)f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,\sigma_1,\sigma_2,\sigma_3) \\
&-Ke^2\sum_{i,j=1}^{3}\frac{1}{\|\mathbf{R}_i-\mathbf{y}_j\|}f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,\sigma_1,\sigma_2,\sigma_3) \\
&+Ke^2(\frac{1}{\|\mathbf{R}_1-\mathbf{R}_2\|}+\frac{1}{\|\mathbf{R}_2-\mathbf{R}_3\|}+\frac{1}{\|\mathbf{R}_3-\mathbf{R}_1\|})f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,\sigma_1,\sigma_2,\sigma_3) \\
&+Ke^2(\frac{1}{\|\mathbf{y}_1-\mathbf{y}_2\|}+\frac{1}{\|\mathbf{y}_2-\mathbf{y}_3\|}+\frac{1}{\|\mathbf{y}_3-\mathbf{y}_1\|})f(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3,\sigma_1,\sigma_2,\sigma_3),
\end{aligned}$$

where $-\frac{\hbar^2}{2m}(\triangle_1+\triangle_2+\triangle_3)$ are the kinetic energy operator; for $i=1,2,3$ representing the 1st, the 2nd and the 3rd electrons, $\triangle_i := (\frac{\partial^2}{\partial(y_1^i)^2}+\frac{\partial^2}{\partial(y_2^i)^2}+\frac{\partial^2}{\partial(y_3^i)^2})$ are the Laplacian for the ith electron's position; $-Ke^2\sum_{i,j=1}^{3}\frac{1}{\|\mathbf{R}_i-\mathbf{y}_j\|^{\frac{1}{2}}}$ is the potential energy operator of the electrons due to attraction to protons; $Ke^2(\frac{1}{\|\mathbf{R}_1-\mathbf{R}_2\|}+\frac{1}{\|\mathbf{R}_2-\mathbf{R}_3\|}+\frac{1}{\|\mathbf{R}_3-\mathbf{R}_1\|})$ is due to proton-proton repulsion; $Ke^2(\frac{1}{\|\mathbf{y}_1-\mathbf{y}_2\|}+\frac{1}{\|\mathbf{y}_2-\mathbf{y}_3\|}+\frac{1}{\|\mathbf{y}_3-\mathbf{y}_1\|})$ is due to electron-electron repulsion.

FACT.

$$[\widehat{H},\hat{\mathbf{S}}^2]=0,[\widehat{H},\hat{S}_3]=0$$

3.4.4. State. (The theory is the same as that in section 2.3.4.)

3.4.5. State Evolution. (The theory is the same as that in section 2.3.5.)

3.4.6. Predicting Measurement Outcomes. (The theory is the same as that in section 2.3.6.)

3.4.7. Updating the state after the measurement. (The theory is the same as that in section 2.3.7.)

Chapter 4

Symmetry Groups of Molecular Systems

4.1. Preparation: Rotation and Reflection Matrices

We develop two lemmas in this section, which show that any rotation matrix A can be identified with a 3×3 matrix $R(e^{i\theta}, \mathbf{u})$, for some $\theta \in \mathbb{R}$, some unit vector $\mathbf{u} \in \mathbb{R}^3$.

Definition. We define

$$\mathrm{SO}(3) := \{A \text{ is a } 3 \times 3 \text{ real matrix } | A^T A = I, \det A = 1\}$$

where I is the 3×3 identity matrix. We call $A \in \mathrm{SO}(3)$ a *rotation matrix.*

Lemma. If $A \in \mathrm{SO}(3)$ then there exists a matrix $U \in \mathrm{SO}(3)$ such that $AU = UR$, where $R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $0 \le \theta \le \pi$. The third column of the matrix U is the axis of the right-handed rotation through the angle θ performed by A in $\mathbb{R}^3$.

Proof. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ then $(A\mathbf{x})^T A\mathbf{y} = \mathbf{x}^T A^T A\mathbf{y} = \mathbf{x}^T\mathbf{y}$, so in particular $\|A\mathbf{x}\| = \|\mathbf{x}\|$ and the angle between $A\mathbf{x}$ and $A\mathbf{y}$ is the same as the angle between $\mathbf{x}$ and $\mathbf{y}$. The cubic polynomial $\det(\lambda I - A)$ has real coefficients, so A has at least one real eigenvalue, and any nonreal eigenvalues must form a complex conjugate pair. If λ is any eigenvalue of A with eigenvector $\mathbf{x}$, then $\|A\mathbf{x}\| = \|\mathbf{x}\|$ implies $|\lambda| = 1$. Thus the real eigenvalues are from the set $\{1, -1\}$. If -1 is the only real eigenvalue, it cannot occur with algebraic multiplicity two, since the other eigenvalue would have to be real, and yet could not be 1 or -1. Since $\det(A)$ is the product of the eigenvalues, we see that the product of the eigenvalues is 1. If -1 has multiplicity one, then there must be a nonreal complex conjugate pair $e^{i\theta}, e^{-i\theta}$ of eigenvalues. But since the product of $e^{i\theta}$ and its complex conjugate is 1, we obtain the contradiction that $(-1)e^{i\theta}e^{-i\theta} = 1$. If -1 has multiplicity three then we obtain the contradiction $(-1)^3 = 1$. Thus 1 must be an eigenvalue. Let $\tilde{\mathbf{u}}_3$ be a normalized eigenvector of A belonging to the eigenvalue 1, and let $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2$ be an orthonormal basis of the plane perpendicular to $\tilde{\mathbf{u}}_3$, so that $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \tilde{\mathbf{u}}_3)$ forms a positively oriented frame of $\mathbb{R}^3$. Define $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \begin{cases} (\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \tilde{\mathbf{u}}_3) & \tilde{\mathbf{u}}_2^T A \tilde{\mathbf{u}}_1 \geq 0, \\ (\tilde{\mathbf{u}}_2, \tilde{\mathbf{u}}_1, -\tilde{\mathbf{u}}_3) & \tilde{\mathbf{u}}_2^T A \tilde{\mathbf{u}}_1 < 0. \end{cases}$ Clearly $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ is a positively oriented orthonormal basis. In the first case above we clearly have $\mathbf{u}_2^T A \mathbf{u}_1 \geq 0$. In the second case we claim that $\mathbf{u}_2^T A \mathbf{u}_1 > 0$. To see this, let $P = \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2\}$. A maps P into itself. The ordered pairs $(\mathbf{u}_1, \mathbf{u}_2)$ and $(A\mathbf{u}_1, A\mathbf{u}_2)$ determine the same orientation of P, i.e. they are related by a 2×2 matrix with positive determinant. (To see this note that $A(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = (A\mathbf{u}_1, A\mathbf{u}_2, A\mathbf{u}_3) = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \left(\begin{smallmatrix} a_{11} & a_{21} & 0 \\ a_{12} & a_{22} & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Now take the determinant of both sides of this equation.) Let the plane P be coordinatized by the components with respect to the vectors $(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$. Then the second case is characterized by the inequality $\tilde{\mathbf{u}}_2^T A \tilde{\mathbf{u}}_1 < 0$, which means that $A\mathbf{u}_2 = A\tilde{\mathbf{u}}_1$ is in the third or fourth quadrant. Hence $A\mathbf{u}_1 = A\tilde{\mathbf{u}}_2$ is in the first or fourth quadrant, and hence the angle between $\mathbf{u}_2$ and $A\mathbf{u}_1$ is less than $\pi/2$, as claimed. Now define $0 \leq \theta \leq \pi$ such that $\cos\theta = \mathbf{u}_1^T A \mathbf{u}_1$. It follows that $A\mathbf{u}_1 = \mathbf{u}_1 \cos\theta + \mathbf{u}_2 \sin\theta$ and $A\mathbf{u}_2 = \mathbf{u}_1(-\sin\theta) + \mathbf{u}_2 \cos\theta$. Setting $U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ we get the result. □

Definition. We define

$$R(e^{i\theta}, \mathbf{u}) = \mathbf{u}\mathbf{u}^T + [I - \mathbf{u}\mathbf{u}^T]\cos\theta + [\mathbf{u}\times]\sin\theta,$$

where θ is real, $\mathbf{u} = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right) \in \mathbb{R}^3$ is a unit vector and $[\mathbf{u}\times] = \left(\begin{smallmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{smallmatrix}\right)$. For $\theta \in (-\pi, \pi]$, we simplify the notation $R(e^{i\theta}, \mathbf{u})$ as $R(\theta, \mathbf{u})$.

Lemma. If $A \in \mathrm{SO}(3)$, then there exists a pair $(e^{i\theta}, \mathbf{u})$, uniquely determined (when $A \neq I$) up to reflection $(e^{-i\theta}, -\mathbf{u})$, where θ is real and $\mathbf{u} \in \mathbb{R}^3$ is a unit vector, such that

$$A = R(e^{i\theta}, \mathbf{u})$$

and hence for all $\mathbf{x} \in \mathbb{R}^3$ we have $A\mathbf{x} = \mathbf{u}(\mathbf{u}\cdot\mathbf{x}) + [\mathbf{x} - \mathbf{u}(\mathbf{u}\cdot\mathbf{x})]\cos\theta + (\mathbf{u}\times\mathbf{x})\sin\theta$.

Proof. We define $\mathbf{u}$ to be the third column vector $\mathbf{u}_3$ of the matrix U in the above lemma. Now let $\mathbf{x} \in \mathbb{R}^3$ be given. Since $U \in \mathrm{SO}(3)$ we have $U^{-1} = U^T$. Thus

$$\begin{aligned}
A\mathbf{x} = URU^T\mathbf{x} &= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T\mathbf{x} \\ \mathbf{u}_2^T\mathbf{x} \\ \mathbf{u}^T\mathbf{x} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u} \end{pmatrix} \begin{pmatrix} \cos\theta\mathbf{u}_1^T\mathbf{x} - \sin\theta\mathbf{u}_2^T\mathbf{x} \\ \sin\theta\mathbf{u}_1^T\mathbf{x} + \cos\theta\mathbf{u}_2^T\mathbf{x} \\ \mathbf{u}^T\mathbf{x} \end{pmatrix} \\
&= \mathbf{u}_1(\cos\theta\mathbf{u}_1^T\mathbf{x} - \sin\theta\mathbf{u}_2^T\mathbf{x}) + \mathbf{u}_2(\sin\theta\mathbf{u}_1^T\mathbf{x} + \cos\theta\mathbf{u}_2^T\mathbf{x}) + \mathbf{u}\mathbf{u}^T\mathbf{x} \\
&= \mathbf{u}\mathbf{u}^T\mathbf{x} + (\mathbf{u}_1\mathbf{u}_1^T\mathbf{x} + \mathbf{u}_2\mathbf{u}_2^T\mathbf{x})\cos\theta + (-\mathbf{u}_1\mathbf{u}_2^T\mathbf{x} + \mathbf{u}_2\mathbf{u}_1^T\mathbf{x})\sin\theta.
\end{aligned}$$

Since $I = UU^T = \mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T + \mathbf{u}\mathbf{u}^T$ we have that $\mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T = I - \mathbf{u}\mathbf{u}^T$ and therefore $\mathbf{u}_1\mathbf{u}_1^T\mathbf{x} + \mathbf{u}_2\mathbf{u}_2^T\mathbf{x} = \mathbf{x} - \mathbf{u}\mathbf{u}^T\mathbf{x}$. Also $\mathbf{u}\times\mathbf{x} = \mathbf{u}\times(\mathbf{u}_1\mathbf{u}_1^T\mathbf{x} + \mathbf{u}_2\mathbf{u}_2^T\mathbf{x} + \mathbf{u}\mathbf{u}^T\mathbf{x}) = (\mathbf{u}\times\mathbf{u}_1)\mathbf{u}_1^T\mathbf{x} + (\mathbf{u}\times\mathbf{u}_2)\mathbf{u}_2^T\mathbf{x} = \mathbf{u}_2\mathbf{u}_1^T\mathbf{x} - \mathbf{u}_1\mathbf{u}_2^T\mathbf{x}$. This demonstrates the existence of the pair $(e^{i\theta}, \mathbf{u})$ with the desired properties.

Clearly if the pair $(e^{i\theta}, \mathbf{u})$ works then so does $(e^{-i\theta}, -\mathbf{u})$. The vector $\mathbf{u}$ must be an eigenvector of A associated to the eigenvalue 1, and this eigenvalue cannot have algebraic or geometric multiplicity two, since then the other eigenvalue would have to be -1, contradicting the fact that $\det A = 1$. (The geometric multiplicity is equal to the algebraic multiplicity since A is clearly diagonalizable.) If the multiplicity is 3 then $A = I$. If $A \neq I$ then the multiplicity is 1, and hence the eigenspace of 1 contains only two real unit eigenvectors $\mathbf{u}$ and $-\mathbf{u}$. Suppose $\mathbf{x}$ is a unit vector perpendicular to $\mathbf{u}$. Then $A\mathbf{x} = \mathbf{x}\cos\theta + \mathbf{u}\times\mathbf{x}\sin\theta$ is an expansion in an orthonormal basis $\{\mathbf{u}, \mathbf{x}, \mathbf{u}\times\mathbf{x}\}$, and hence $\cos\theta = \mathbf{x}\cdot A\mathbf{x}$ and $\sin\theta = (\mathbf{u}\times\mathbf{x})\cdot A\mathbf{x}$. If $A = R(e^{i\theta}, \mathbf{u}) = R(e^{i\theta'}, \mathbf{u})$, then $A\mathbf{x} = \mathbf{x}\cos\theta + \mathbf{u}\times\mathbf{x}\sin\theta = \mathbf{x}\cos\theta' + \mathbf{u}\times\mathbf{x}\sin\theta' \Rightarrow \cos\theta = \cos\theta', \sin\theta = \sin\theta' \Rightarrow e^{i\theta} = e^{i\theta'}$. The same values of $\cos\theta$ and $\sin\theta$ are obtained independently of the choice of $\mathbf{x}$. Thus both $\cos\theta$ and $\sin\theta$ are determined by $\mathbf{u}$. This proves the uniqueness claim. □

Definition. Consider a plane which is perpendicular to the axis $\mathbf{u}$ and passes through $\mathbf{0}$. The matrix of reflection through this plane is $R_\sigma = I - 2\mathbf{u}\mathbf{u}^T$, where I is the identify matrix. (if it does not cause confusion, later we will not mention that the plan passes through $\mathbf{0}$.)

4.2. Representations of Molecular Symmetry Groups

4.2.1. Point symmetries and Conjugacy.

Definition. Let V be a finite dimensional complex vector space. We define

$$GL(V) := \{A : V \to V | A \text{ is linear and invertible } \};$$

$$GL(\mathbb{C}^n) := \text{the set of all } n \times n \text{ complex invertible matrices };$$

$$GL(\mathbb{R}^n) := \text{the set of all } n \times n \text{ real invertible matrices };$$

$$O(3) := \{A \in \mathbb{R}^{3\times 3} | A^T A = I\}.$$

Fact. $GL(V)$ is a group under composition of linear transformations.

Definition. An n dimensional complex representation of a group G is a group homomorphism $\rho : G \to GL(V)$, where V is n-dimensional complex vector space.

According to [**6**], a distance preserving transformation of space that leaves an object looking the same after it has been applied is called a *symmetry operation* of the object. There is a corresponding *symmetry element* for each symmetry operation; this is the point, line, or plane with respect to which the symmetry operation is performed. If a particular point is chosen as the *center of symmetry* then the symmetry operations which fix this point are called *point symmetries*. The *identity* I means doing nothing; the corresponding symmetry element is the entire space. An *n-fold rotation* C_n (the operation) about an *n-fold axis of symmetry* (the corresponding symmetry element) is a rotation through $2\pi/n$. The *principal axis* of a molecule is an n-fold axis of symmetry where n is as large as possible. A reflection σ (the operation) in a *plane of symmetry* or a *mirror plane* (the symmetry element) may be either parallel or perpendicular to a principle axis of a molecule. If the plane is parallel to the principal axis, it is called *vertical* and denoted σ_v. When the plane of symmetry is perpendicular to the principal axis it is called *horizontal* and denoted σ_h. A vertical mirror plane that bisects the angle between two C_2 axes is called a *dihedral plane* and denoted σ_d. In an inversion i (the operation) through the center of symmetry (the symmetry element) we imagine taking each point in a molecule, moving it to its center, and then moving it out the same distance on the other side. An *improper rotation* or *rotary-reflection* S_n (the operation) about an *axis of improper rotation* or a *rotary-reflection axis* (the symmetry element) consists of rotation through $2\pi/n$ about an n-fold rotation axis followed by a horizontal reflection.

DEFINITION. Let G be a group, $g_1, g_2 \in G$. If there is $x \in G$ such that $g_1 = x^{-1}g_2x$, then we say that g_1 is *conjugate* to g_2, denoted as $g_1 \sim g_2$. $\sim$ is an equivalence relation in G. A $\sim$ equivalence class in G is called a *conjugacy class*.

For the purpose of point symmetries we may identify space with a 3-dimensional $\mathbb{R}$-inner product space X where the zero vector $\mathbf{0} \in X$ represents the center of symmetry. Define

$$O(X) := \{g : X \to X \text{ is } \mathbb{R}\text{-linear and inner product preserving i.e. isometric}\}.$$

There is a natural left action of $O(X)$ on X.

Let $\mathcal{M} \subset X$ be a finite set of vectors. Define the point symmetry group $G(\mathcal{M})$ as

$$G(\mathcal{M}) := \{g \in O(X) | g \cdot \mathbf{m} \in \mathcal{M}, \forall \mathbf{m} \in \mathcal{M}\}.$$

It is clear that $G(\mathcal{M})$ is a subgroup of $O(X)$.

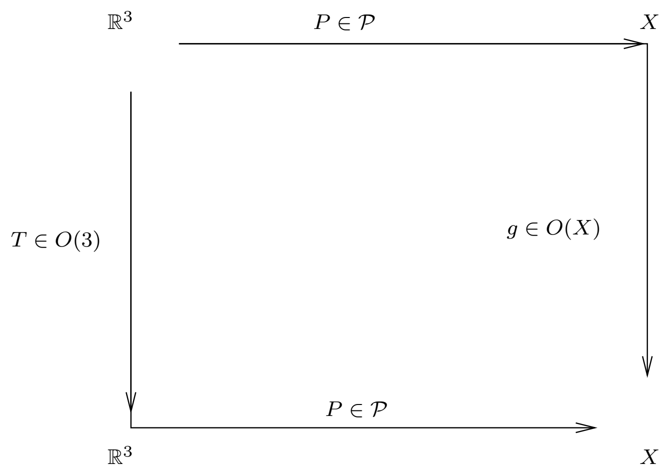

FIGURE 2. Given $P \in \mathcal{P}$, there exists a group isomorphism between $O(X)$ and $O(3)$ defined as $\rho_P(g) := P^{-1} \circ g \circ P$.

Define

$$\mathcal{P} := \{P : \mathbb{R}^3 \to X | P \text{ is a } \mathbb{R}\text{-linear isomorphism which is inner product preserving}\}.$$

$\mathcal{P}$ is called the set of *poses*. $P \in \mathcal{P}$ can be identified with the ordered 3-tuple $(P(\hat{\mathbf{e}}_1), P(\hat{\mathbf{e}}_2), P(\hat{\mathbf{e}}_3))$, where $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ is the canonical orthonormal basis of $\mathbb{R}^3$. $O(X)$ acts on $\mathcal{P}$ on the left by the rule

$$gP = (gP(\hat{\mathbf{e}}_1), gP(\hat{\mathbf{e}}_2), gP(\hat{\mathbf{e}}_3)), \text{i.e. } g \circ P : \mathbb{R}^3 \xrightarrow{P} X \xrightarrow{g} X, \text{ where } g \in O(X).$$

$O(3)$ acts on $\mathcal{P}$ on the right by the rule

$$PT = P \circ T : \mathbb{R}^3 \xrightarrow{T} \mathbb{R}^3 \xrightarrow{P} X, \text{ where } T \in O(3).$$

Given $P \in \mathcal{P}$, define a mapping $\rho_P : O(X) \to O(3)$ as $\rho_P(g) := P^{-1} \circ g \circ P$. This mapping is clearly a group isomorphism (see Figure 2). When we restrict the domain of ρ_P from $O(X)$ to $G(\mathcal{M})$, the above isomorphism gives a mapping $\rho_P : G(\mathcal{M}) \to O(3)$.

FACT. If $G(\mathcal{M})$ is a group of point symmetries of a molecule, then we have an injective (faithful) 3-dimensional real orthogonal representation: $\rho_P : G(\mathcal{M}) \to O(3)$.

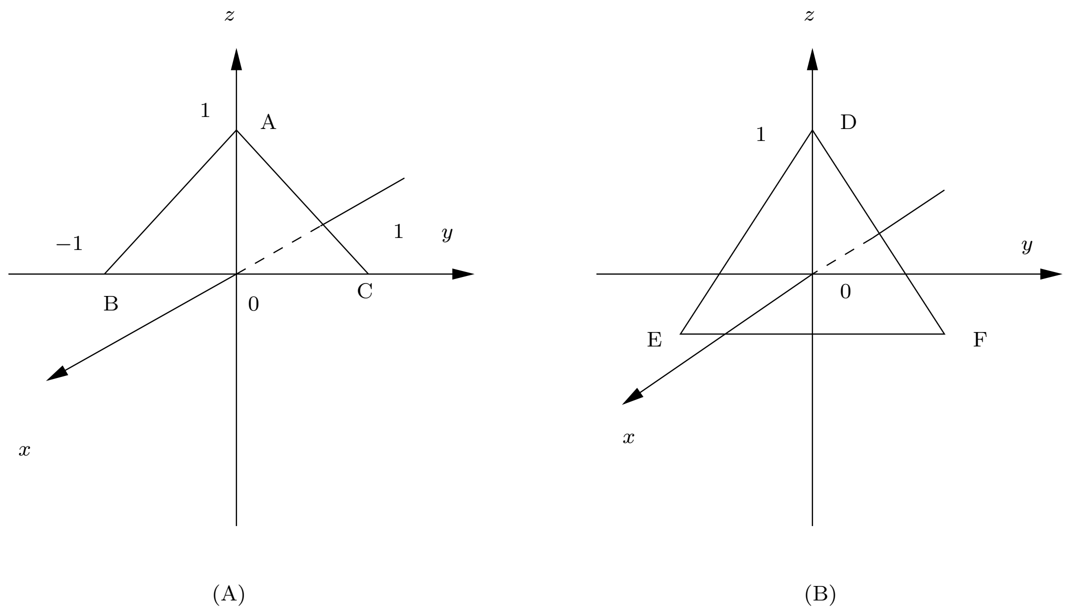

Figure 3. (A). The isosceles triangle group is $C_{2v} = \{E, C_2, \sigma_v, \sigma'_v\}$. (B). The equilateral triangle group is $D_{3h} = \{E, \sigma_h, 2C_3, 2S_3, 3C'_2, 3\sigma_v\}$. Note that in these figures a definite pose $P \in \mathcal{P}$ is given and fixed, so we can identify symmetry operations with matrices in $O(3)$.

Suppose $P_i \in \mathcal{P}, i = 1, 2$ are poses and $T_i = (P_i)^{-1} \circ g \circ P_i$, then $P_i \circ T_i \circ (P_i)^{-1} = g$ and $T_1 = P_1^{-1} \circ g \circ P_1 = P_1^{-1} \circ (P_2 \circ T_2 \circ (P_2)^{-1}) \circ P_1 = (P_2^{-1} \circ P_1)^{-1} \circ T_2 \circ (P_2^{-1} \circ P_1)$. Note that $P_2^{-1} \circ P_1 \in O(3)$, and thus $T_1 \sim T_2$ are conjugate in $O(3)$ and correspond to the same symmetry operation g with respect to two different poses P_1 and P_2. This is the geometric meaning of conjugacy.

4.2.2. Examples. From now on we will use the symmetry group names $E, C_{2v}, D_{3h}, T_d, \sigma_v$ etc.; they follow standard chemical nomenclature [**6**].

Example. We consider the symmetry group $C_{2v} = \{E, C_2, \sigma_v, \sigma'_v\}$ of the isosceles triangle $\triangle ABC$, where $A = (0, 0, 1), B = (0, -1, 0), C = (0, 1, 0)$ (see Figure 3 (A)). Recall that the rotation matrix $R(\theta, \mathbf{u})$ represents the rotation of a vector about the axis $\mathbf{u}$ through an angle of $\theta \in (-\pi, \pi]$, and the reflection matrix R_σ represents the reflection of a vector through the plane which is orthogonal to $\mathbf{u}$ and passes though the origin of the coordinates.

We denote the matrix representation of a group element $g \in G$ as $\rho(g)$. By the meaning of conjugacy, we classify the symmetry group elements for the isosceles triangle by conjugacy classes as follows.

(1) The identity E is represented by $\rho(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(2) The 180^0 rotation about the axis $\mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is represented by

$$\begin{aligned} \rho(C_2) &= R(180^0, \mathbf{u}) \\ &= \mathbf{u}\mathbf{u}^T + (I - \mathbf{u}\mathbf{u}^T)\cos(180^0) + [\mathbf{u}\times]\sin(180^0) \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} + (I - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix})(-1) \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

(3) The reflection through the plane perpendicular to the axis $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is represented by

$$\begin{aligned} \rho(\sigma_v) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

The reflection through the plane perpendicular to the axis $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ is represented by

$$\begin{aligned} \rho(\sigma_v') &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

EXAMPLE. We consider the symmetry group D_{3h} of the equilateral triangle $\triangle DEF$, where $D = (0, 0, 1), E = (0, -\frac{\sqrt{3}}{2}, -\frac{1}{2}), F = (0, \frac{\sqrt{3}}{2}, -\frac{1}{2})$ (see Figure 3 (B)). Then the elements of the group classified by conjugacy classes are as follows.

(1) The identity E is represented by $\rho'(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(2) The 180^0 rotation about the axis $\mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is represented by

$$\rho'(C_2(D, O)) = R(180^0, \mathbf{u}_1)$$

$$= \mathbf{u}_1\mathbf{u}_1^T + (I - \mathbf{u}_1\mathbf{u}_1^T)\cos(180^0) + [\mathbf{u}_1\times]\sin(180^0)$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The 180^0 rotation about the axis $\mathbf{u}_2 = \begin{pmatrix} 0 \\ -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}$ is represented by

$$\rho'(C_2(E,O)) = R(180^0, \mathbf{u}_2)$$

$$= \mathbf{u}_2\mathbf{u}_2^T + (I - \mathbf{u}_2\mathbf{u}_2^T)\cos(180^0) + [\mathbf{u}_2\times]\sin(180^0)$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

The 180^0 rotation about the axis $\mathbf{u}_3 = \begin{pmatrix} 0 \\ \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}$ is represented by

$$\rho'(C_2(F,O)) = R(180^0, \mathbf{u}_3)$$

$$= \mathbf{u}_3\mathbf{u}_3^T + (I - \mathbf{u}_3\mathbf{u}_3^T)\cos(180^0) + [\mathbf{u}_3\times]\sin(180^0)$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

(3) The 120^0 rotation around the axis $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is represented by

$$\rho'(C_3(+1)) = R(120^0, \mathbf{u})$$

$$= \mathbf{u}\mathbf{u}^T + (I - \mathbf{u}\mathbf{u}^T)\cos(120^0) + [\mathbf{u}\times]\sin(120^0)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

The 240^0 rotation around the axis $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is represented by

$$\rho'(C_3(-1)) = R(240^0, \mathbf{u})$$

$$= \mathbf{u}\mathbf{u}^T + (I - \mathbf{u}\mathbf{u}^T)\cos(240^0) + [\mathbf{u}\times]\sin(240^0)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

(4) The reflection through the plane perpendicular to $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, i.e. "horizontal" reflection, is represented by

$$\begin{aligned} \rho'(\sigma_h) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

(5) The reflection through the plane perpendicular to $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ is represented by

$$\begin{aligned} \rho'(\sigma_{v1}) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

The reflection through the plane perpendicular to $\mathbf{u} = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$ is represented by

$$\begin{aligned} \rho'(\sigma_{v2}) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

The reflection through the plane perpendicular to $\mathbf{u} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$ is represented by

$$\begin{aligned} \rho'(\sigma_{v3}) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

(6) The 120^0 rotation around the x axis followed by a horizontal reflection is represented by :

$$\begin{aligned} \rho'(S_{31}) &= \rho'(\sigma_h)\rho'(C_3(+1)) \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \end{aligned}$$

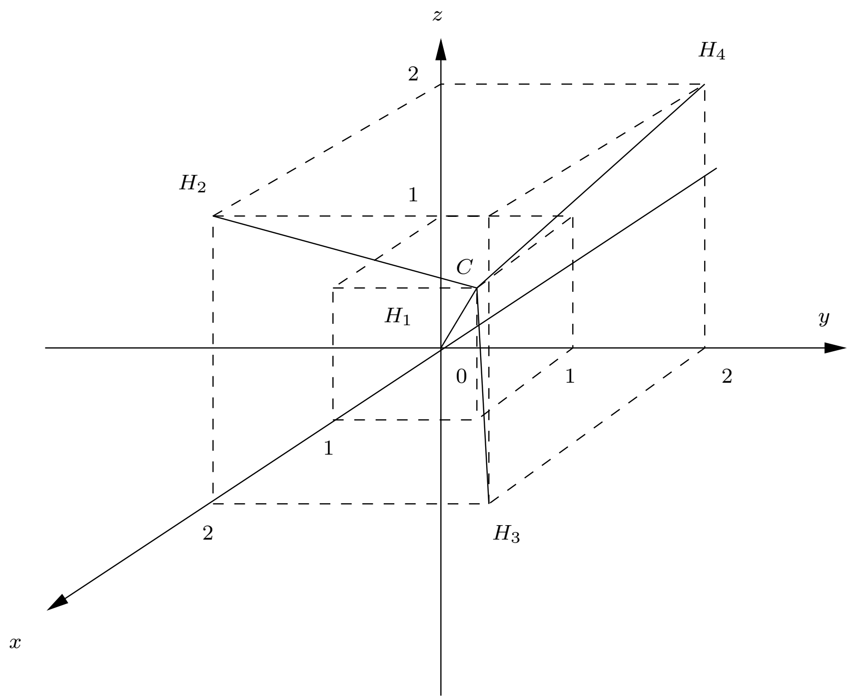

FIGURE 4. The T_d group on the molecule CH_4.

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

The 240^0 rotation around the x axis followed by a horizontal reflection is represented by :

$$\begin{aligned} \rho'(S_{32}) &= \rho'(\sigma_h)\rho'(C_3(-1)) \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

EXAMPLE. We want to construct a faithful 3 dimensional orthogonal representation of the symmetry group T_d of the Tetrahedron, and we will also classify the following 24 elements of T_d into conjugacy classes according to their geometric meanings:

$$\begin{aligned}
&\{E; \\
&C_3(1,+1), C_3(1,-1), C_3(2,+1), C_3(2,-1), \\
&C_3(3,+1), C_3(3,+1), C_3(4,-1), C_3(4,-1); \\
&C_2(1,2) = C_2(3,4), C_2(1,3) = C_2(2,4), C_2(1,4) = C_2(2,3); \\
&\sigma_d(1,2), \sigma_d(1,3), \sigma_d(1,4), \sigma_d(2,3), \sigma_d(2,4), \sigma_d(3,4); \\
&S_4(1,2), S_4(1,3), S_4(1,4), S_4(2,3), S_4(2,4), S_4(3,4)\}.
\end{aligned}$$

To simplify the notation, we will consider T_d on a model of the molecule CH_4 (see Figure 4).

The 24 elements of T_d classified by conjugacy classes by the geometric viewpoint are as follows.

(1) The identity E.

(2) The 120^0 rotations about the axis $C - H_i, i = 1, 2, 3, 4$, denoted as $C_3(i, +1)$ and the -120^0 rotations about the axis $C - H_i, i = 1, 2, 3, 4$, denoted as $C_3(i, -1)$. They are $\{C_3(1,+1), C_3(1,-1), C_3(2,+1), C_3(2,-1),$ $C_3(3,+1), C_3(3,-1), C_3(4,+1), C_3(4,-1)\}$.

(3) The 180^0 rotations about the axis bisecting the $H_i - C - H_j$ angle, denoted as $C_2(i,j)$. They are $\{C_2(1,2) = C_2(3,4), C_2(1,3) = C_2(2,4),$ $C_2(1,4) = C_2(2,3)\}$.

(4) The reflections through the plane of $H_i - C - H_j$, denoted as $\sigma_d(i,j)$. They are $\{\sigma_d(1,2), \sigma_d(1,3), \sigma_d(1,4), \sigma_d(2,3), \sigma_d(2,4), \sigma_d(3,4)\}$.

(5) The rotations of 90^0 about the axis bisecting $H_i - C - H_j$, followed by the reflection through the plane perpendicular to this axis, denoted as $S_4(i,j)$. They are $\{S_4(1,2), S_4(1,3), S_4(1,4), S_4(2,3), S_4(2,4), S_4(3,4)\}$.

Suppose the positions of the atoms in CH_4 are : $H_1(0,0,0), H_2(2,0,2), H_3(2,2,0)$, $H_4(0,2,2), C(1,1,1)$, which is the center of the symmetry. We will show how to compute $\rho''(C_3(2,+1))$ –the matrix of the rotation around the axis $C-H_2$ by 120^0.

The axis $C-H_2$ is represented by unit vector $\mathbf{u} = \frac{1}{\sqrt{3}}(\begin{pmatrix} 1\\0\\0 \end{pmatrix} + \begin{pmatrix} 0\\0\\1 \end{pmatrix} - \begin{pmatrix} 0\\1\\0 \end{pmatrix}) = \frac{1}{\sqrt{3}}\begin{pmatrix} 1\\-1\\1 \end{pmatrix}$.

Then we have $\mathbf{u}\mathbf{u}^T = \frac{1}{3}\begin{pmatrix} 1\\-1\\1 \end{pmatrix}\begin{pmatrix} 1\\-1\\1 \end{pmatrix}^T = \frac{1}{3}\begin{pmatrix} 1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & 1 \end{pmatrix}$.

$$\begin{aligned}\rho''(C_3(2,+1)) =& \mathbf{u}\mathbf{u}^T + (I-\mathbf{u}\mathbf{u}^T)\cos\theta + [\mathbf{u}\times]\sin\theta\\ =& \frac{1}{3}\begin{pmatrix} 1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & 1 \end{pmatrix} + \{\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & 1 \end{pmatrix}\}\cos 120^0\\ &+ \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & -1 & -1\\ 1 & 0 & -1\\ 1 & 1 & 0 \end{pmatrix}\frac{\sqrt{3}}{2}\\ =& \begin{pmatrix} 0 & -1 & 0\\ 0 & 0 & -1\\ 1 & 0 & 0 \end{pmatrix}.\end{aligned}$$

In similar way we can compute the matrices representation of the other tetrahedral group members as follows.

$\rho''(C_3(1,+1)) = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$,

$\rho''(C_3(1,-1)) = -\begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} = -\rho''(C_3(1,+1))^T$,

$\rho''(C_3(2,+1)) = \begin{pmatrix} 0 & -1 & 0\\ 0 & 0 & -1\\ 1 & 0 & 0 \end{pmatrix}$, (as computed above)

$\rho''(C_3(2,-1)) = \begin{pmatrix} 0 & 0 & -1\\ 1 & 0 & 0\\ 0 & 1 & 0 \end{pmatrix} = -\rho''(C_3(2,+1))^T$.

The six matrices above belong to one conjugacy class, the elements of which elements represent the $\pm 120^0$ rotations about the axis $C-H_i, i=1,2,3,4$.

$\rho''(C_2(1,2)) = \begin{pmatrix} -1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1 \end{pmatrix} = \rho''(C_2(3,4))$,

$\rho''(C_2(1,3)) = \begin{pmatrix} -1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1 \end{pmatrix} = \rho''(C_2(2,4))$,

$\rho''(C_2(1,4)) = \begin{pmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1 \end{pmatrix} = \rho''(C_2(2,3))$.

The six matrices above belong to one conjugacy class, the elements of which represent the 180^0 rotations about the axis bisecting the $H_i - C - H_j$ angle.

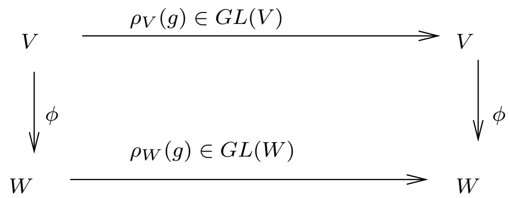

FIGURE 5. ρ_V is isomorphic with ρ_W.

$$\rho''(\sigma_d(1,2)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \rho''(\sigma_d(1,3)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\rho''(\sigma_d(1,4)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \rho''(\sigma_d(2,3)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix},$$
$$\rho''(\sigma_d(2,4)) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \rho''(\sigma_d(3,4)) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

The six matrices above belong to one conjugacy class, the elements of which represent the reflections through the plane of $H_i - C - H_j$.

$$\rho''(S_4(1,2)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \rho''(S_4(1,3)) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$
$$\rho''(S_4(1,4)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \rho''(S_4(2,3)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix},$$
$$\rho''(S_4(2,4)) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \rho''(S_4(3,4)) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The six matrices above belong to one conjugacy class, the elements of which represent the rotations of 90^0 about the axis bisecting $H_i - C - H_j$, followed by the reflection through the plane perpendicular to this axis.

4.2.3. Isomorphic Representations.

DEFINITION. Let $\rho_V : G \to GL(V)$, $\rho_W : G \to GL(W)$ be two group representations of group G, we say that ρ_V *is isomorphic with* ρ_W if there is a $\mathbb{C}$-linear isomorphism $\phi : V \to W$ such that for any given $g \in G$, $[\phi\rho_V(g)](v) = [\rho_W(g)\phi](v)$ or $\rho_V(g)(v) = [\phi^{-1}\rho_W(g)\phi](v)$ for all $v \in V$ (as illustrated at Figure 5). Particularly, if $V = W = \mathbb{C}^n$, then $\phi \in GL(\mathbb{C}^n)$.

EXAMPLE. A molecule having the symmetries of D_{3h} is PCl_5 (Phosphorus Pentachloride) , we will first use it as our model to compute a group representation ρ''' of D_{3h}. And then, we will show that this representation ρ''' for PCl_5 is isomorphic to the earlier representation ρ' of D_{3h} for the equilateral triangle $\triangle DEF$.

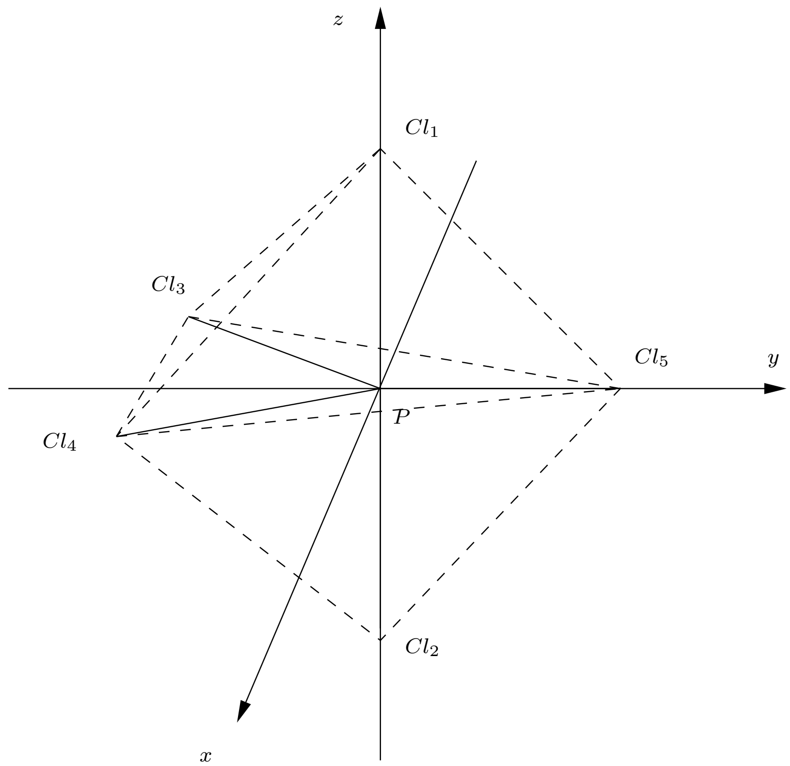

FIGURE 6. The group D_{3h} for the molecule PCl_5

Suppose the positions for the atoms of PCl_5 are as follows. $P(0,0,0), Cl_1(0,0,1)$, $Cl_2(0,0,-1)$, $Cl_3(\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0)$, $Cl_4(-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0)$, $Cl_5(0,1,0)$ (see Figure 6). $D_{3h} = \{E, \sigma_h, 2C_3, 2S_2, 3C_2', 3\sigma_v\}$. To distinguish the matrix representation for PCl_5 from the previous one for equilateral triangle, we denote the matrix representation for PCl_5 of $g \in D_{3h}$ as $\rho'''(g)$.

(1) The identity E is represented by $\rho'''(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$;

(2) For σ_h, $\mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$,

$$\begin{aligned} \rho'''(\sigma_h) &= I - 2\mathbf{u}\mathbf{u}^T \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

(3) For $C_3(+1), \mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \theta = 120°$,

$$\rho'''(C_3(+1)) = \mathbf{u}\mathbf{u}^T + (I - \mathbf{u}\mathbf{u}^T)\cos 120^\circ + [\mathbf{u}\times]\sin 120^\circ$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\sqrt{3}}{2}\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For $C_3(-1), \mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \theta = 120^\circ,$

$$\rho'''(C_3(-1)) = \mathbf{u}\mathbf{u}^T + (I - \mathbf{u}\mathbf{u}^T)\cos 120^\circ + [u\times]\sin 120^\circ$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}(-\frac{1}{2}) + \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\frac{\sqrt{3}}{2}$$

$$= \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(4) For $C_2(3,4) : \mathbf{u} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \theta = 180^\circ,$

$$\rho'''(C_2(3,4)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

For $C_2(4,5) : \mathbf{u} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, \theta = 180^\circ,$

$$\rho'''(C_2(4,5)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

For $C_2(5,3) : \mathbf{u} = \begin{pmatrix} \frac{-\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, \theta = 180^\circ,$

$$\rho'''(C_2(5,3)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(5) For $\sigma_v(1,5), \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$

$$\rho'''(\sigma_v(1,5)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For $\sigma_v(1,4), \mathbf{u} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}$,

$$\rho'''(\sigma_v(1,4)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For $\sigma_v(1,3), \mathbf{u} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}$,

$$\rho'''(\sigma_v(1,3)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(6) For S_3 and S_3',

$$\begin{aligned} \rho'''(S_3) &= \rho'''(\sigma_h)\rho'''(C_3) \\ &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}, \\ \rho'''(S_3) &= \rho'''(\sigma_h)\rho'''(C_3) \\ &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Now we introduce $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\phi(x,y,z) = (z,x,y)$. Then ϕ corresponds to a rotation matrix $R(-120^0, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, and hence ϕ^{-1} corresponds to $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. We need to check as follows that $\rho_V'(g) = \phi^{-1}\rho_V'''(g)\phi$, $\forall g \in D_{3h}$, where ρ_V', ρ_V''' denote the previous and the present D_{3h} matrix representation respectively:

(1) For the identity E,

$$\begin{aligned} \phi^{-1}\rho'(E)\phi &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \rho'''(E). \end{aligned}$$

(2) For σ_h,

$$\phi^{-1}\rho'(\sigma_h)\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(\sigma_h).$$

(3) For $2C_3$,

$$\phi^{-1}\rho'(C_3(+1))\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \rho'''(C_3(+1)),$$

$$\phi^{-1}\rho'(C_3(-1))\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \rho'''(C_3(-1)).$$

(4) For $2S_3$,

$$\phi^{-1}\rho'(S_{31})\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(S_3),$$

$$\phi^{-1}\rho'(S_{32})\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(S_3').$$

(5) For $3C_2$,

$$\phi^{-1}\rho'(C_2(D,O))\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(C_2(3,4)),$$

$$\phi^{-1}\rho'(C_2(E,O))\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(C_2(4,5)),$$

$$\phi^{-1}\rho'(C_2(F,O))\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$= \rho'''(C_2(5,3)).$$

(6) For the $3\sigma_v$,

$$\phi^{-1}\rho'(\sigma_{v1})\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \rho'''(\sigma_v(1,5)),$$

$$\phi^{-1}\rho'(\sigma_{v2})\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \rho'''(\sigma_v(1,3)),$$

$$\phi^{-1}\rho'(\sigma_{v3})\phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \rho'''(\sigma_v(1,4)).$$

Therefore ρ' and ρ''' are isomorphic representations.

4.2.4. Characters.

DEFINITION. If $\rho : G \to GL(V)$ is a representation of G, its *character* χ_ρ is the complex-valued function on the group defined by

$$\chi_\rho(g) = \mathrm{tr}(\rho(g)),$$

i.e. the trace of g on V. We sometimes write χ_V instead of χ_ρ if there is no possible confusion.

In particular, for any $g, h \in G$, we have

$$\chi_\rho(hgh^{-1}) = \chi_\rho(g),$$

so that χ_ρ is constant on a conjugacy classes of G; a complex-valued function on G which is constant on each conjugacy class of G is called a *class function*. Denote $\mathbb{C}_{\text{class}}(G) = \{\text{class functions on } G\}$. Note that $\chi_\rho(e) = \dim V$, where e is the identity element in G.

EXAMPLE. Consider the character of the matrix representation of $C_{2v} = \{E, C_2, \sigma_v, \sigma_v'\}$ by the model of isosceles triangle $\triangle ABC$ (see Figure 3).

(1) $\rho(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_\rho(I) = \mathrm{tr}(\rho(E)) = 3$.
(2) $\rho(C_2) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_\rho(C_2) = \mathrm{tr}(\rho(C_2)) = -1$.
(3) $\rho(\sigma_v) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_\rho(\sigma_v) = \mathrm{tr}(\rho(\sigma_v)) = 1$.
$\rho(\sigma_v') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_\rho(\sigma_v') = \mathrm{tr}(\rho(\sigma_v')) = 1$.

EXAMPLE. Consider the character of the matrix representation ρ' of D_{3h} by the model of equilateral triangle $\triangle DEF$ (see Figure 3).

(1) $\rho'(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'}(I) = \mathrm{tr}(\rho'(E)) = 3$.
(2) $\rho'(\sigma_h) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'}(\sigma_h) = \mathrm{tr}(\rho'(\sigma_h)) = 1$.

(3) $\rho'(C_3(+1)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(C_3(+1)) = \text{tr}(\rho'(C_3(+1))) = 0$.

$\rho'(C_3(-1)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(C_3(-1)) = \text{tr}(\rho'(C_3(-1))) = 0$.

(4) $\rho'(S_{31}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(S_{31}) = \text{tr}(\rho'(S_{31})) = -2$.

$\rho'(S_{32}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(S_{32}) = \text{tr}(\rho'(S_{32})) = -2$.

(5) $\rho'(C_2(D,O)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then

$\chi_{\rho'}(C_2(D,O)) = \text{tr}(\rho'(C_2(D,O))) = -1$.

$\rho'(C_2(E,O)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then

$\chi_{\rho'}(C_2(E,O) = \text{tr}(\rho'(C_2(E,O))) = -1$.

$\rho'(C_2(F,O)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then

$\chi_{\rho'}(C_2(F,O)) = \text{tr}(\rho'(C_2(F,O))) = -1$.

(6) $\rho'(\sigma_{v1}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'}(\sigma_{v1}) = \text{tr}(\rho'(\sigma_{v1})) = 1$.

$\rho'(\sigma_{v2}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(\sigma_{v2}) = \text{tr}(\rho'(\sigma_{v2})) = 1$.

$\rho'(\sigma_{v3}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$. Then $\chi_{\rho'}(\sigma_{v3}) = \text{tr}(\rho'(\sigma_{v3})) = 1$.

Example. Consider the character of the matrix representation ρ''' of $D_{3h} = \{E, \sigma_h, 2C_3, 2S_3, 3C_2', 3\sigma_v\}$ for PCl_5.

(1) $\rho'''(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(I) = \text{tr}(\rho'''(E)) = 3$.

(2) $\rho'''(\sigma_h) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(\sigma_h) = \text{tr}(\rho'''(\sigma_h)) = 1$.

(3) $\rho'''(C_3(+1)) = \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(C_3(+1)) = \text{tr}(\rho'''(C_3(+1))) = 0$.

$\rho'''(C_3(-1)) = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(C_3(-1)) = \text{tr}(\rho'''(C_3(-1))) = 0$.

(4) $\rho'''(S_3) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(S_3) = \text{tr}(\rho'''(S_3)) = -2$.

$\rho'''(S_3') = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(S_3') = \text{tr}(\rho'''(S_3')) = -2$.

(5) $\rho'''(C_2(3,4)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(C_2(3,4)) = \mathrm{tr}(\rho'''(C_2(3,4))) = -1$.

$\rho'''(C_2(4,5)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(C_2(4,5)) = \mathrm{tr}(\rho'''(C_2(4,5))) = -1$.

$\rho'''(C_2(5,3)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then $\chi_{\rho'''}(C_2(5,3)) = \mathrm{tr}(\rho'''(C_2(5,3))) = -1$.

(6) $\rho'''(\sigma_v(1,5)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(\sigma_v(1,5)) = \mathrm{tr}(\rho'''(\sigma_v(1,5))) = 1$.

$\rho'''(\sigma_v(1,4)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(\sigma_v(1,4)) = \mathrm{tr}(\rho'''(\sigma_v(1,4))) = 1$.

$\rho'''(\sigma_v(1,3)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\chi_{\rho'''}(\sigma_v(1,3)) = \mathrm{tr}(\rho'''(\sigma_v(1,3))) = 1$.

Remark: These characters of C_{2v} and D_{3h} are seen to be class functions. Through the example of D_{3h}, we see an illustration of the fact that the isomorphic representations have the same character.

Definition. A representation $\rho : G \to GL(V)$ is called *irreducible* if whenever $W \subseteq V$ is a vector subspace such that $\rho(g)W \subseteq W, \forall g \in G$, then either $W = \{\mathbf{0}\}$ or $W = V$. ρ is *reducible* if it is not irreducible.

Definition. Define the Hermitian inner product of α and β in $\mathbb{C}_{\text{class}}(G)$ by:

$$(\alpha, \beta) = \frac{1}{|G|} \sum_{g \in G} \overline{\alpha(g)} \beta(g) = \frac{1}{|G|} \sum_{c \text{ is a conjugacy class}} |c| \overline{\alpha(c)} \beta(c).$$

Theorem. Consider the set $\mathcal{I}_G$ of all equivalence classes $[V]$ of irreducible finite dimensional representations V of a finite group G, with the equivalence relation being isomorphism of representations. Then $|\mathcal{I}_G|$ equals the number of conjugacy classes of G. Furthermore if $\mathcal{I}_G = \{[V_1], \cdots, [V_n]\}$, then $\{\chi_{V_1}, \cdots, \chi_{V_n}\}$ is an orthonormal basis of the vector space $\mathbb{C}_{\text{class}}(G)$ of class functions.

4.2.5. Character Tables. We follow [**44**] [**6**] to develop this subsection. A finite group G has a finite number of conjugacy classes and the same finite number of non-isomorphic irreducible representations. The character of a group representation is

C_{2v}	E	C_2	σ_v	σ_v'	$h = 4$	
A_1	1	1	1	1	z, z^2, x^2, y^2	
A_2	1	1	-1	-1	xy	R_z
B_1	1	-1	1	-1	x, xz	R_y
B_2	1	-1	-1	1	y, yz	R_x

FIGURE 7. The Character Table of C_{2v}

constant on each conjugacy class. Hence, the values of the characters can be written as an array, known as a *character table.* Typically, the rows are given by the irreducible non-isomorphic representations and the columns are given by the conjugacy classes.

A character table often contains enough information to identify a given abstract group and distinguish it from others. However, there exist non-isomorphic groups which nevertheless have the same character table, for example D_8 (the symmetry group of the square) and A_8 (the quaternion group)[**6**].

EXAMPLE. [**44**] Chemists and physicists use a special convention for representing character tables which is applied especially to the so-called point groups, which are the 32 finite symmetry groups possible in a lattice. In the table 7, the numbered regions contain the following contents ([**20**] pp. $90 - 92$).

(1) The symbol used to represent the group in question (in this case C_{2v}).
(2) The conjugacy classes, indicated by number and symbol, where the sum of the coefficients gives the order h of the group.
(3) Mulliken symbols one for each irreducible representation.

D_{3h}	E	σ_h	$2C_3$	$2S_3$	$3C_2'$	$3\sigma_v$	$h = 12$	
A_1'	1	1	1	1	1	1	$z^2, x^2 + y^2$	
A_2'	1	1	1	1	-1	-1		R_z
A_1''	1	-1	1	-1	1	-1		
A_2''	1	-1	1	-1	-1	1	z	
E'	2	2	-1	-1	0	0	(x, y), $(xy, x^2 - y^2)$	
E''	2	-2	-1	1	0	0	(xz, yz)	(R_x, R_y)

FIGURE 8. The Character Table of D_{3h}

(4) An array of the group characters of the non-isomorphic irreducible representations of the group, with one column for each conjugacy class, and one row for each irreducible representation.

(5) and (6) are not going to be used in this thesis, so we do not discuss them.

EXAMPLE. We know by the above theorem that two rows in a character table are orthogonal with respect to the inner product defined above. For instance, in the character table of D_{3h} (see Figure 8), the two rows A_1'' and E'' satisfy:

$$\begin{aligned}(\chi_{A_1''}, \chi_{E''}) &= \frac{1}{|D_{3h}|} \sum_{c \text{ is a conjugacy class}} |c| \chi_{A_1''}(c) \chi_{E''}(c) \\ &= \frac{1}{12}[1 \times 1 \times 2 + 1 \times (-1) \times (-2) + 2 \times 1 \times (-1) + 2 \times (-1) \times 1 \\ &\quad + 3 \times 1 \times 0 + 3 \times (-1) \times 0] \\ &= 0.\end{aligned}$$

4.2.6. Decomposition of Representations.

Definition. Let $\rho_V : G \to GL(V)$, $\rho_W : G \to GL(W)$ be two representations of the same group G, then the direct sum of these two representations is defined as:

$$\rho_V \oplus \rho_W : G \to GL(V \oplus W) : \quad g \mapsto \rho_V(g) \oplus \rho_W(g).$$

Let $V = \mathbb{C}^n, W = \mathbb{C}^m$, then $\rho_V(g) \in \mathbb{C}^{n\times n}, \rho_W(g) \in \mathbb{C}^{m\times m}$, and $(\rho_V \oplus \rho_W)(g) \in \mathbb{C}^{(n+m)\times(n+m)}, (\rho_V \oplus \rho_W)(g) = \left(\begin{smallmatrix} \rho_V(g) & 0 \\ 0 & \rho_W(g) \end{smallmatrix}\right)$.

Lemma. Let $\rho_V : G \to GL(V)$ and $\rho_W : G \to GL(W)$ be representations of G. Then $\chi_{V\oplus W} = \chi_\rho + \chi_W$.

Definition. Let V be a vector space. Define $V^{\oplus a} = V \overset{1}{\oplus} \cdots \overset{a-1}{\oplus} V$, where $a \in \mathbb{N}, a \geq 2$.

Theorem. Suppose $\mathcal{I}_G = \{[V_1], \cdots, [V_n]\}$, where G is a finite group. Suppose V is an arbitrary finite dimensional representation of G. If $\chi_V = \sum_{i=1}^n a_i \chi_{V_i}$ where $a_i = (\chi_{V_i}, \chi_V)$, then $a_i \geq 0$ is an integer and $V \cong V_1^{\oplus a_1} \oplus \cdots \oplus V_n^{\oplus a_n}$.

Example. We continue the the last example in §4.2.4 where D_{3h} for PCl_5 is discussed. For the representation ρ''' for the character is $\chi_{\rho'''} = (3, 1, 0, -2, -1, 1)^T$. Then

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} = \begin{pmatrix} (\chi_{A_1'}, \chi_{\rho'''}) \\ (\chi_{A_2'}, \chi_{\rho'''}) \\ (\chi_{A_1''}, \chi_{\rho'''}) \\ (\chi_{A_2''}, \chi_{\rho'''}) \\ (\chi_{E'}, \chi_{\rho'''}) \\ (\chi_{E''}, \chi_{\rho'''}) \end{pmatrix}$$

$$= \frac{1}{|D_{3h}|} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 \\ 2 & 2 & -1 & -1 & 0 & 0 \\ 2 & -2 & -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 0 \\ -2 \\ -1 \\ 1 \end{pmatrix}$$

$$= \frac{1}{12} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 \\ 2 & 2 & -1 & -1 & 0 & 0 \\ 2 & -2 & -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 0 \\ -4 \\ -3 \\ 3 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Thus $\chi_{\rho'''} = \chi_{E'} + \chi_{A_2''}$, and hence $V_{\rho'''} = V_{E'} \oplus V_{A_2''}$. We show the details as follows.

$$\rho'''(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

$$\rho'''(\sigma_h) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(C_3(+1)) = \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} \\ \frac{-\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

$$\rho'''(C_3(-1)) = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

$$\rho'''(S_3) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(S_3') = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(C_2(3,4)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(C_2(4,5)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(C_2(5,3)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} -1 \end{pmatrix},$$

$$\rho'''(\sigma_v(1,5)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

$$\rho'''(\sigma_v(1,4)) = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

$$\rho'''(\sigma_v(1,3)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix},$$

where the 2×2 matrices are the irreducible representation E' of D_{3h} and the 1×1 matrices are the irreducible representation A_2'' of D_{3h}.

Chapter 5

The Conformation and Configuration of the H_3 System

In this chapter we first discuss the shapes (conformations) of the H_3 system in general coordinates: how to mathematically represent a shape and how to mathematically classify the shapes into the three different categories: the non-collinear, the collinear, and the one-point-coincident, and then we build up three particular internal coordinate systems which treat the three hydrogen nuclei in a symmetrical way.

5.1. The Shape Space of the H_3 System

In this section, we will discuss the mathematical representation of the conformations of the H_3 system. After rigid motion, the conformation (or shape) of a particular molecule does not change. To identify the mathematical representation of the conformation of a molecule, we introduce the concept of orbit. We also investigate the categories within the set of all orbits of configurations of H_3 according to the orbit's dimensionality.

Let $R = (\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3) \in (\mathbb{R}^3)^3$ be the position vectors of the three nuclei of the H_3 system. R is called a *configuration*. Let $\mathbf{b} \in \mathbb{R}^3$ be a translation vector. Let $A \in \mathrm{SO}(3)$ be a rotation matrix. We define $G_a := \mathbb{R}^3 \times \mathrm{SO}(3)$, which is the group of rigid motions of $\mathbb{R}^3$. The group law for G_a is, for any $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^3$ and $A_1, A_2 \in \mathrm{SO}(3)$,

$$(\mathbf{b}_1, A_1)(\mathbf{b}_2, A_2) = (\mathbf{b}_1 + A_1\mathbf{b}_2, A_1A_2)$$

$(\mathbf{b}, A) \in G_a$ acts on the left of $(\mathbb{R}^3)^3$ by the rule

$$(\mathbf{b}, A)(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3) = (\mathbf{b} + A\mathbf{R}_1, \mathbf{b} + A\mathbf{R}_2, \mathbf{b} + A\mathbf{R}_3),$$

which represents the configuration R of the molecular system after the rigid motion $(\mathbf{b}, A)$ is applied.

Let's check that this is a left action:

$$\begin{aligned}
&[(\mathbf{b}_1, A_1)(\mathbf{b}_2, A_2)](\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3) \\
&\quad = (\mathbf{b}_1 + A_1\mathbf{b}_2, A_1A_2)(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3) \\
&\quad = (\mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_1, \mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_2, \mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_3), \\
&(\mathbf{b}_1, A_1)[(\mathbf{b}_2, A_2)(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3)] \\
&\quad = (\mathbf{b}_1, A_1)(\mathbf{b}_2 + A_2\mathbf{R}_1, \mathbf{b}_2 + A_2\mathbf{R}_2, \mathbf{b}_2 + A_2\mathbf{R}_3) \\
&\quad = (\mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_1, \mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_2, \mathbf{b}_1 + A_1\mathbf{b}_2 + A_1A_2\mathbf{R}_3). \quad \square
\end{aligned}$$

To identify the conformation of a molecule, we introduce the concept of orbit as follows.

Definition. The *orbit* of a three-atom molecule configuration $R = (\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3)$ is defined as

$$G_aR := \{(\mathbf{b}, A)R \in (\mathbb{R}^3)^3 | (\mathbf{b}, A) \in G_a\}$$

R is called a *representative of the orbit.*

Definition. The *Shape space* of the three-atom molecule is the set of all its orbits in $(\mathbb{R}^3)^3$, denoted by $G_a\backslash\backslash(\mathbb{R}^3)^3$,

$$G_a\backslash\backslash(\mathbb{R}^3)^3 := \{G_aR | R \in (\mathbb{R}^3)^3\}.$$

With the notations defined above, we now start to investigate the classifications of the orbits of the H_3 molecule. We claim that there are three categories of orbits in shape space:

$$\dim\, G_a R = \begin{cases} 6, & \text{if } R \text{ is a non-collinear configuration;} \\ 3, & \text{if } \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3; \\ 5, & \text{if } R \text{ is a collinear configuration but not } \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3. \end{cases}$$

where $\dim\, G_a R$ denotes the dimension of the obit $G_a R$ thought of as a manifold, i.e. the number of independent variables needed to parameterize $G_a R$.

Definition. Given a configuration R, its *isotropy subgroup* within the group G_a is defined as

$$\mathrm{Iso}_{G_a} R := \{(\mathbf{b}, A) \in G_a \mid (\mathbf{b}, A)R = R\}.$$

Fact. Given a configuration R, the mapping of the group onto the orbit:

$$G_a \to G_a R : (\mathbf{b}, A) \mapsto (\mathbf{b}, A)R$$

is always surjective. It induces a one-to-one and onto mapping:

$$f : G_a // \mathrm{Iso}_{G_a} R \to G_a R : (\mathbf{b}, A)\mathrm{Iso}_{G_a} R \mapsto (\mathbf{b}, A)R,$$

where $G_a // \mathrm{Iso}_{G_a} R$ is the quotient set of $\mathrm{Iso}_{G_a} R$ in G.

This bijective mapping is well-defined since if $(\mathbf{b}', A') = (\mathbf{b}, A)g, g \in \mathrm{Iso}_{G_a} R$, then $(\mathbf{b}', A')\mathrm{Iso}_{G_a} R = (\mathbf{b}, A)\mathrm{Iso}_{G_a} R$ and $(\mathbf{b}', A')R = (\mathbf{b}, A)gR = (\mathbf{b}, A)R$.

This fact tells us that we can identify the members of the class of rigid motions $(\mathbf{b}, A)\mathrm{Iso}_{G_a} R$ because after acting on the configuration R they all produce the same outcome $(\mathbf{b}, A)R$.

Now we consider the orbit classification problem of the H_3 system. There are three situations for the configuration R: non-collinear, collinear and one-point-coincident. We will show in a theorem that the isotropy subgroups corresponding to these three kinds of R are of different dimensions, and consequently so are the orbits. We need the following lemma and corollary as a preparation.

LEMMA. If $A \in \mathrm{SO}(3)$, and $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^3$, then $A(\mathbf{e}_1 \times \mathbf{e}_2) = A\mathbf{e}_1 \times A\mathbf{e}_2$.

PROOF. : For $A \in \mathrm{SO}(3)$, by the lemma in the previous chapter, we assume that $A = R(e^{i\theta}, \mathbf{u}) = \mathbf{u}\mathbf{u}^T + [I - \mathbf{u}\mathbf{u}^T]\cos\theta + [\mathbf{u}\times]\sin\theta$, where $R(e^{i\theta}, \mathbf{u})$ represents the rotation along the axis $\mathbf{u}$ through an angle θ. Then we have

$$A(\mathbf{e}_1 \times \mathbf{e}_2) = \mathbf{u}[\mathbf{u}\cdot(\mathbf{e}_1 \times \mathbf{e}_2)] + \{\mathbf{e}_1 \times \mathbf{e}_2 - \mathbf{u}[\mathbf{u}\cdot(\mathbf{e}_1 \times \mathbf{e}_2)]\}\cos\theta + \mathbf{u} \times (\mathbf{e}_1 \times \mathbf{e}_2)\sin\theta,$$

and

$$\begin{aligned}
&(A\mathbf{e}_1) \times (A\mathbf{e}_2)\\
&=\{\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_1) + [\mathbf{e}_1 - \mathbf{u}(\mathbf{u}\cdot\mathbf{e}_1)]\cos\theta + \mathbf{u}\times\mathbf{e}_1\sin\theta\}\\
&\qquad \times \{\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2) + [\mathbf{e}_2 - \mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2)]\cos\theta + \mathbf{u}\times\mathbf{e}_2\sin\theta\}\\
&=\cos\theta[\mathbf{e}_1 \times \mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2) + \mathbf{u}\times\mathbf{e}_2(\mathbf{u}\cdot\mathbf{e}_1)]\\
&\qquad + \sin\theta[\mathbf{u}\times(\mathbf{u}\times\mathbf{e}_2)(\mathbf{u}\cdot\mathbf{e}_1) + (\mathbf{u}\times\mathbf{e}_1)\times\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2)]\\
&\qquad + (\cos\theta)^2[\mathbf{e}_1\times\mathbf{e}_2 - \mathbf{e}_1\times\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2) - \mathbf{u}\times\mathbf{e}_2(\mathbf{u}\cdot\mathbf{e}_1)]\\
&\qquad + (\sin\theta)^2(\mathbf{u}\times\mathbf{e}_1)\times(\mathbf{u}\times\mathbf{e}_2)\\
&\qquad + \sin\theta\cos\theta[\mathbf{e}_1\times(\mathbf{u}\times\mathbf{e}_2) - \mathbf{u}\times(\mathbf{u}\times\mathbf{e}_2)(\mathbf{u}\cdot\mathbf{e}_1) - (\mathbf{u}\times\mathbf{e}_1)\times\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2)\\
&\qquad + (\mathbf{u}\times\mathbf{e}_1)\times\mathbf{e}_2]\\
&=\cdots \quad \text{(we will show the calculation details after this equation)}\\
&=\mathbf{u}[\mathbf{u}\cdot(\mathbf{e}_1\times\mathbf{e}_2)] + \{\mathbf{e}_1\times\mathbf{e}_2 - \mathbf{u}[\mathbf{u}\cdot(\mathbf{e}_1\times\mathbf{e}_2)]\}\cos\theta + \mathbf{u}\times(\mathbf{e}_1\times\mathbf{e}_2)\sin\theta.
\end{aligned}$$

Here are the calculation details: because of the identity $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = -\mathbf{a}(\mathbf{b}\cdot\mathbf{c}) + \mathbf{b}(\mathbf{a}\cdot\mathbf{c})$, the coefficient of $\cos\theta$ is (denoted as equ.1.)

$$\begin{aligned}
&\mathbf{e}_1\times\mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2) + \mathbf{u}\times\mathbf{e}_2(\mathbf{u}\cdot\mathbf{e}_1)\\
&\quad = [\mathbf{e}_1(\mathbf{e}_2\cdot\mathbf{u}) - \mathbf{e}_2(\mathbf{e}_1\cdot\mathbf{u})]\times\mathbf{u}\\
&\quad = [(\mathbf{e}_2\times\mathbf{e}_1)\times\mathbf{u}]\times\mathbf{u}
\end{aligned}$$

$$= -(\mathbf{e}_2 \times \mathbf{e}_1)(\mathbf{u} \cdot \mathbf{u}) + \mathbf{u}[(\mathbf{e}_2 \times \mathbf{e}_1) \cdot \mathbf{u}]$$

$$= \mathbf{e}_1 \times \mathbf{e}_2 - \mathbf{u}[\mathbf{u} \cdot (\mathbf{e}_1 \times \mathbf{e}_2)].$$

The coefficient of $\sin\theta$ is (denoted as equ.2.)

$$\mathbf{u} \times (\mathbf{u} \times \mathbf{e}_2)(\mathbf{u} \cdot \mathbf{e}_1) + (\mathbf{u} \times \mathbf{e}_1) \times \mathbf{u}(\mathbf{u} \cdot \mathbf{e}_2)$$

$$= -[\mathbf{u} \times \mathbf{e}_2(\mathbf{e}_1 \cdot \mathbf{u})] \times \mathbf{u} + [\mathbf{u} \times \mathbf{e}_1(\mathbf{e}_2 \cdot \mathbf{u})] \times \mathbf{u}$$

$$= \{\mathbf{u} \times [-\mathbf{e}_2(\mathbf{e}_1 \cdot \mathbf{u}) + \mathbf{e}_1(\mathbf{e}_2 \cdot \mathbf{u})]\} \times \mathbf{u}$$

$$= \{\mathbf{u} \times [(\mathbf{e}_2 \times \mathbf{e}_1) \times \mathbf{u}]\} \times \mathbf{u}$$

$$= \{[(\mathbf{e}_1 \times \mathbf{e}_2) \times \mathbf{u}] \times \mathbf{u}\} \times \mathbf{u}$$

$$= \{-\mathbf{e}_1 \times \mathbf{e}_2(\mathbf{u} \cdot \mathbf{u}) + \mathbf{u}[\mathbf{u} \cdot (\mathbf{e}_1 \times \mathbf{e}_2)]\} \times \mathbf{u}$$

$$= -(\mathbf{e}_1 \times \mathbf{e}_2) \times \mathbf{u}$$

$$= \mathbf{u} \times (\mathbf{e}_1 \times \mathbf{e}_2).$$

The coefficient of $\sin^2\theta$ is

$$(\mathbf{u} \times \mathbf{e}_1) \times (\mathbf{u} \times \mathbf{e}_2)$$

$$= \mathbf{e}_1[\mathbf{u} \cdot (\mathbf{u} \times \mathbf{e}_2)] - \mathbf{u}[\mathbf{e}_1 \cdot (\mathbf{u} \times \mathbf{e}_2)]$$

$$= \mathbf{u}[\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{u})]$$

$$= \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{u}].$$

Combining the terms involving $\cos^2\theta$ and $\sin^2\theta$ we get

$$\cos^2\theta\{\mathbf{e}_1 \times \mathbf{e}_2 - (\mathbf{e}_1 \times \mathbf{u})(\mathbf{u} \cdot \mathbf{e}_2) - (\mathbf{u} \times \mathbf{e}_2)(\mathbf{u} \cdot \mathbf{e}_1)\}$$

$$+ (1 - \cos^2\theta)\mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{u}]$$

$$= \cos^2\theta\{\mathbf{e}_1 \times \mathbf{e}_2 - (\mathbf{e}_1 \times \mathbf{u})(\mathbf{u} \cdot \mathbf{e}_2) - (\mathbf{u} \times \mathbf{e}_2)(\mathbf{u} \cdot \mathbf{e}_1)$$

$$- \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{u}]\} + \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{u}]$$

$$
\begin{aligned}
&= \cos^2\theta\{\mathbf{e}_1 \times \mathbf{e}_2 - \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2)\cdot\mathbf{u}] - (\mathbf{e}_1 \times \mathbf{u})(\mathbf{u}\cdot\mathbf{e}_2) \\
&\qquad - (\mathbf{u} \times \mathbf{e}_2)(\mathbf{u}\cdot\mathbf{e}_1)\} + \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2)\cdot\mathbf{u}] \\
&= \mathbf{u}[(\mathbf{e}_1 \times \mathbf{e}_2)\cdot\mathbf{u}]
\end{aligned}
$$

The last step of the above equality is because the coefficient of $\cos^2\theta$ vanishes by equ.1.

The coefficient of $\sin\theta\cos\theta$ is:

$$
\begin{aligned}
&\mathbf{e}_1 \times (\mathbf{u} \times \mathbf{e}_2) - \mathbf{u} \times (\mathbf{u} \times \mathbf{e}_2)(\mathbf{u}\cdot\mathbf{e}_1) - (\mathbf{u} \times \mathbf{e}_1) \times \mathbf{u}(\mathbf{u}\cdot\mathbf{e}_2) + (\mathbf{u} \times \mathbf{e}_1) \times \mathbf{e}_2 \\
&= \mathbf{e}_1 \times (\mathbf{u} \times \mathbf{e}_2) + (\mathbf{e}_1 \times \mathbf{e}_2) \times \mathbf{u} + (\mathbf{u} \times \mathbf{e}_1) \times \mathbf{e}_2 \text{ (by equ.2)} \\
&= -(\mathbf{u} \times \mathbf{e}_2) \times \mathbf{e}_1 - (\mathbf{e}_2 \times \mathbf{e}_1) \times \mathbf{u} - (\mathbf{e}_1 \times \mathbf{u}) \times \mathbf{e}_2 \\
&= \mathbf{0} \text{ (by Jacobi identity).}
\end{aligned}
$$

□

Corollary. If $A \in \mathrm{SO}(3)$, and $A\mathbf{x}_1 = \mathbf{x}_1, A\mathbf{x}_2 = \mathbf{x}_2$, where $\mathbf{x}_1, \mathbf{x}_2$ are orthogonal unit vectors of $\mathbb{R}^3$, then $A = I$.

Proof. We define $\mathbf{x}_3 = \mathbf{x}_1 \times \mathbf{x}_2$, then $X = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is an orthonormal basis of $\mathbb{R}^3$. By the lemma, $AX = X$, then $A = A(XX^T) = (AX)X^T = XX^T = I$. □

Theorem.

$$
\mathrm{Iso}_{G_a} R = \begin{cases} \{(\mathbf{0}, I)\}, & \text{if } R \text{ is a non-collinear configuration.} \\ \{((I-A)\mathbf{R}_1, A) | A \in \mathrm{SO}(3)\}, & \text{if } \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3; \\ \{((I - R(e^{i\theta}, \mathbf{u}))\mathbf{R}_1, R(e^{i\theta}, \mathbf{u})) | \theta \in [0, 2\pi)\}, & \text{otherwise.} \end{cases}
$$

where in the 3rd case $\mathbf{u}$ is a unit vector in $\mathrm{span}\{\mathbf{R}_2 - \mathbf{R}_1, \mathbf{R}_3 - \mathbf{R}_1\}$.

PROOF. : Suppose $(b, A) \in \mathrm{Iso}_{G_a} R$. In the case that R is non-collinear, $\{\mathbf{R}_2 - \mathbf{R}_1, \mathbf{R}_3 - \mathbf{R}_1\}$ is a linearly independent set in $\mathbb{R}^3$. Since

$$\mathbf{b} + A\mathbf{R}_1 = \mathbf{R}_1,$$

$$\mathbf{b} + A\mathbf{R}_2 = \mathbf{R}_2,$$

$$\mathbf{b} + A\mathbf{R}_3 = \mathbf{R}_3.$$

We have that

$$A(\mathbf{R}_2 - \mathbf{R}_1) = \mathbf{R}_2 - \mathbf{R}_1,$$

$$A(\mathbf{R}_3 - \mathbf{R}_1) = \mathbf{R}_3 - \mathbf{R}_1,$$

i.e. A has the eigenvalue 1, and its eigenspace is at least 2 dimensional. Let $\{\mathbf{x}_1, \mathbf{x}_2\}$ be orthonormal basis of $\mathrm{span}\{\mathbf{R}_2 - \mathbf{R}_1, \mathbf{R}_3 - \mathbf{R}_1\}$, then $A\mathbf{x}_1 = \mathbf{x}_1, A\mathbf{x}_2 = \mathbf{x}_2$. By the corollary, we have $A = I$. Together with $\mathbf{b} + A\mathbf{R}_1 = \mathbf{R}_1$, we have $\mathbf{b} = \mathbf{0}$ and $(\mathbf{b}, A) = (\mathbf{0}, I)$. Thus, $\mathrm{Iso}_{G_a} R = \{(\mathbf{0}, I)\}$, which is 0 dimensional.

Let's consider the second case: $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3$. We have $(\mathbf{b}, A)\mathbf{R}_1 = \mathbf{b} + A\mathbf{R}_1 = \mathbf{R}_1 \Rightarrow \mathbf{b} = (I - A)\mathbf{R}_1$. Then $\mathrm{Iso}_{G_a} R \subseteq \{((I - A)\mathbf{R}_1, A), A \in \mathrm{SO}(3)\}$; it is quite clear that the reverse inclusion also holds, thus $\mathrm{Iso}_{G_a} R = \{((I - A)\mathbf{R}_1, A), A \in \mathrm{SO}(3)\}$; it is 3 dimensional since $A = R(e^{i\theta}, \mathbf{u})$ has 3 free variables.

The third case is that R is collinear but it is not the case that $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3$. Let $\mathbf{u}$ be a unit vector in $\mathrm{span}\{\mathbf{R}_2 - \mathbf{R}_1, \mathbf{R}_3 - \mathbf{R}_1\}$ (the space is one dimensional). We have $A\mathbf{u} = \mathbf{u}$. By the argument in the Lemmas developed in the previous section, for the eigenvalue 1 of A, the corresponding eigenspace is either 3 dimensional, in which case $A = I$, or 1 dimensional, in which case $A = R(e^{i\theta}, \mathbf{u})$, for some $\theta \in (0, 2\pi)$. Summing up these two cases, we can represent $A = R(e^{i\theta}, \mathbf{u})$, for some $\theta \in [0, 2\pi)$. So A has one free variable $\theta \in [0, 2\pi)$. $(\mathbf{b}, A)\mathbf{R}_1 = \mathbf{b} + A\mathbf{R}_1 = \mathbf{R}_1 \Rightarrow \mathbf{b} = (I - A)\mathbf{R}_1$. Then we have shown that $\mathrm{Iso}_{G_a}(R) \subseteq \{((I - R(e^{i\theta}, \mathbf{u}))\mathbf{R}_1, R(e^{i\theta}, \mathbf{u})) | \theta \in [0, 2\pi)\}$.

We are going to show the reverse inclusion. Suppose $((I - A)\mathbf{R}_1, A) \in \{((I - R(e^{i\theta}, \mathbf{u}))\mathbf{R}_1, R(e^{i\theta}, \mathbf{u}))|\theta \in [0, 2\pi)\}$, we claim $((I - A)\mathbf{R}_1, A)\mathbf{R}_i = \mathbf{R}_i, i = 1, 2, 3$, and hence $\{((I - R(e^{i\theta}, \mathbf{u}))\mathbf{R}_1, R(e^{i\theta}, \mathbf{u}))|\theta \in [0, 2\pi)\} \subseteq \mathrm{Iso}_{G_a}(R)$.

To see,

$$\begin{aligned}
((I - A)\mathbf{R}_1, A)\mathbf{R}_1 &= (I - A)\mathbf{R}_1 + A\mathbf{R}_1 \\
&= \mathbf{R}_1; \\
((I - A)\mathbf{R}_1, A)\mathbf{R}_2 &= (I - A)\mathbf{R}_1 + A\mathbf{R}_2 \\
&= \mathbf{R}_1 + A(\mathbf{R}_2 - \mathbf{R}_1) \\
&= \mathbf{R}_1 + (\mathbf{R}_2 - \mathbf{R}_1) \\
&= \mathbf{R}_2; \\
((I - A)\mathbf{R}_1, A)\mathbf{R}_3 &= (I - A)\mathbf{R}_1 + A\mathbf{R}_3 \\
&= \mathbf{R}_1 + A(\mathbf{R}_3 - \mathbf{R}_1) \\
&= \mathbf{R}_1 + (\mathbf{R}_3 - \mathbf{R}_1) \\
&= \mathbf{R}_3.
\end{aligned}$$

In sum, we have $\mathrm{Iso}_{G_a}(R) = \{((I - R(e^{i\theta}, \mathbf{u}))\mathbf{R}_1, R(e^{i\theta}, \mathbf{u}))|\theta \in [0, 2\pi)\}$. □

By the theorem above, we have

$$\begin{aligned}
&\dim G_a R \\
&= \dim\,(G_a // Iso_{G_a} R) \\
&= \dim G_a - \dim \mathrm{Iso}_{G_a} R \\
&= \begin{cases} 6 - 0 = 6, & \text{if } R \text{ is non-collinear configuration;} \\ 6 - 3 = 3, & \text{if } \mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3; \\ 6 - 1 = 5, & \text{otherwise.} \end{cases}
\end{aligned}$$

In the next section we will see that the 6 dimensional orbits form the interior of a cone, the 5 dimensional orbits form the surface of the cone, and the 3 dimensional orbit form the tip of the cone.

5.2. Different Coordinates on the Shape Space of the H_3 System.

In this section we discuss the shape space of the H_3 system using three particular coordinate systems. Because the conformation (or shape) of the H_3 molecule only depends on the relative positions of its nuclei, the coordinates need to be internal. It is also desirable to define a coordinate system that treats the three nuclei in a symmetrical way. There are three internuclear distances for three nuclei. We are going to use them and some appropriate functions of them to define internal coordinates.

The first coordinate system on the shape space is (l_{12}, l_{13}, l_{23}), where $l_{ij} = \|\mathbf{R}_j - \mathbf{R}_i\|, i, j = 1, 2, 3, i \neq j$ (see figure 9). These are constrained by the triangle inequalities $l_{23} \leq l_{12} + l_{13}, l_{13} \leq l_{12} + l_{23}, l_{12} \leq l_{23} + l_{13}$.

The conformation of the molecular system can be identified by the 3 internuclear distances (l_{12}, l_{23}, l_{13}). We claim that these distances of the H_3 molecular system do not change after any rigid motion $(\mathbf{b}, A)$ and hence we can identify them with an orbit. To see this, after the rigid motion $(\mathbf{b}, A)$ is applied, the square of the new internuclear distances are:

$$\begin{aligned} l'^2_{ij} &= \|\mathbf{R}'_j - \mathbf{R}'_i\|^2 \\ &= \|(\mathbf{b} + A\mathbf{R}_j) - (\mathbf{b} + A\mathbf{R}_i)\|^2 \\ &= \|A(\mathbf{R}_j - \mathbf{R}_i)\|^2 \\ &= [A(\mathbf{R}_j - \mathbf{R}_i)]^T A(\mathbf{R}_j - \mathbf{R}_i) \\ &= (\mathbf{R}_j - \mathbf{R}_i)^T A^T A(\mathbf{R}_j - \mathbf{R}_i) \end{aligned}$$

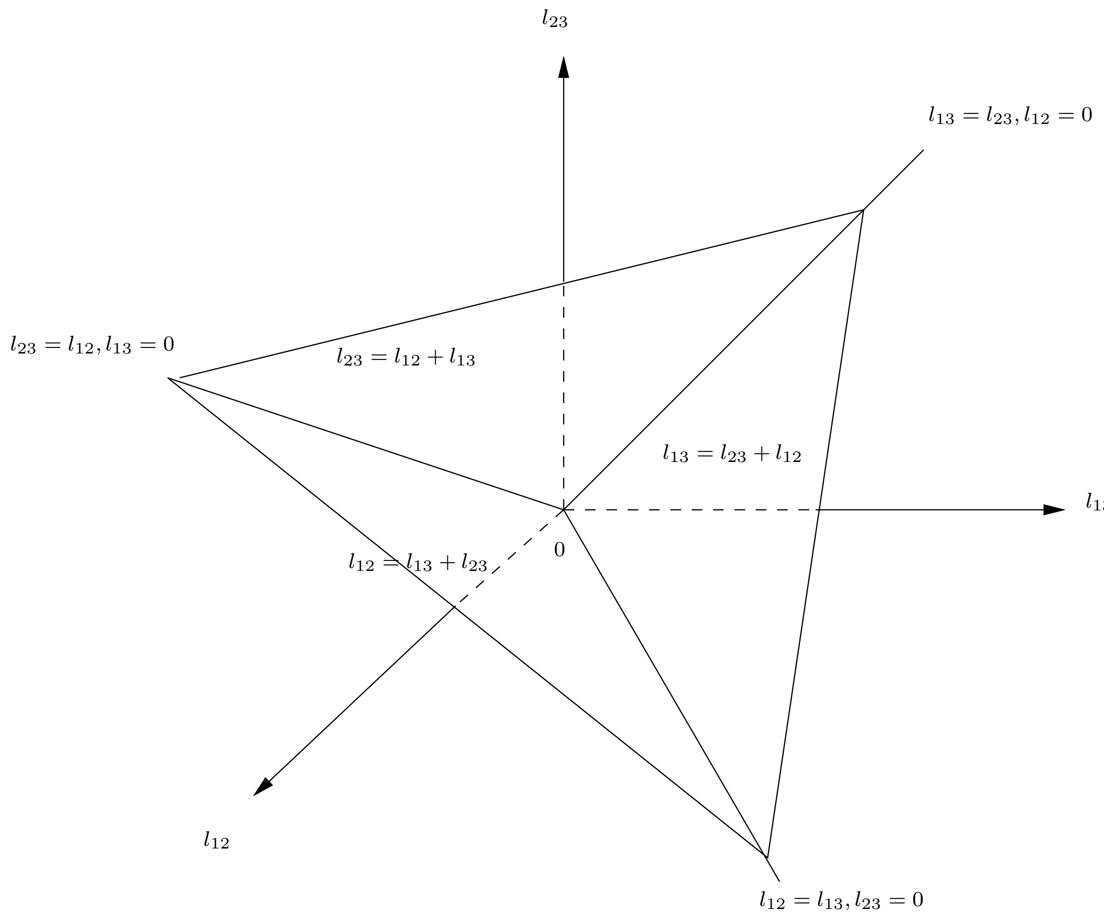

FIGURE 9. (l_{12}, l_{13}, l_{23}) Coordinate.

$$= (\mathbf{R}_j - \mathbf{R}_i)^T(\mathbf{R}_j - \mathbf{R}_i)$$

$$= \|\mathbf{R}_j - \mathbf{R}_i\|^2$$

$$= l_{ij}^2,$$

where $i, j = 1, 2, 3, i \neq j$.

The second coordinate system on the shape space is $(l_{12}^2, l_{13}^2, l_{23}^2)$. (See figure 10.)

In the following lemma, we will show that the collinear configurations are on the surface of a cone in this coordinate system.

LEMMA. We define

$s_1 := l_{12}^2, s_2 := l_{13}^2, s_3 := l_{23}^2,$

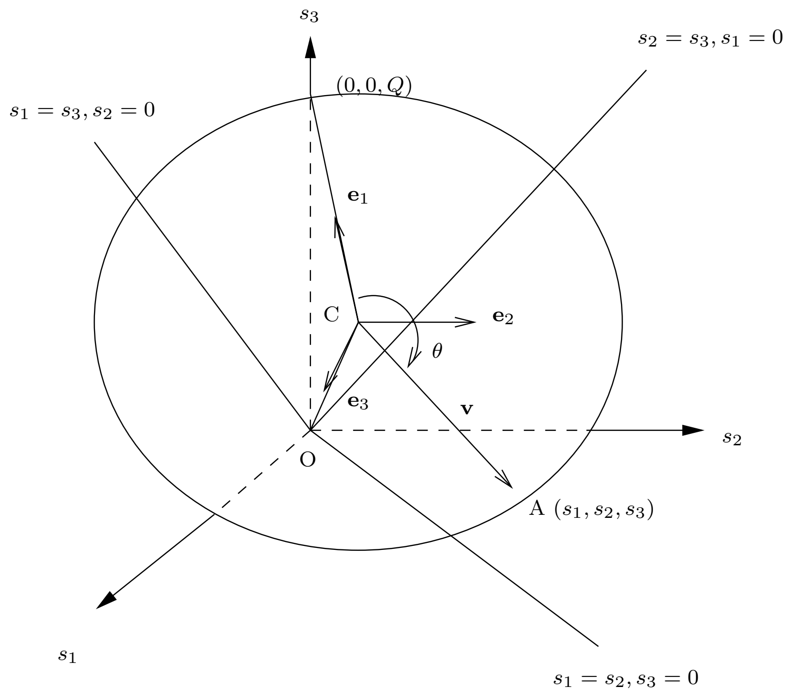

FIGURE 10. The Collinear configurations are on the surface of the cone $\mathcal{C}$. The intersection of the plane $s_1+s_2+s_3 = Q$ with the three surfaces $\sqrt{s_1} + \sqrt{s_2} = \sqrt{s_3}$, $\sqrt{s_2} + \sqrt{s_3} = \sqrt{s_1}$, and $\sqrt{s_1} + \sqrt{s_3} = \sqrt{s_2}$ is a circle centered at $C = (Q/3, Q/3, Q/3)$ with radius $\frac{Q}{\sqrt{6}}$. Hence the cone $\mathcal{C}$ is bounded by the three surfaces $\sqrt{s_1} + \sqrt{s_2} = \sqrt{s_3}, \sqrt{s_2} + \sqrt{s_3} = \sqrt{s_1}$, and $\sqrt{s_1} + \sqrt{s_3} = \sqrt{s_2}$. The axis of the cone is $s_1 = s_2 = s_3$.

$$\mathcal{Q} := \{(l_{12}, l_{13}, l_{23}) \in \mathbb{R}^3 | l_{12} + l_{23} \geq l_{13}, l_{13} + l_{23} \geq l_{12}, l_{12} + l_{13} \geq l_{23}; l_{12}, l_{13}, l_{23} \geq 0\},$$

$$\mathcal{C} := \{(s_1, s_2, s_3) \in \mathbb{R}^3 | s_1, s_2, s_3 \geq 0; (s_1 - \frac{s_1+s_2+s_3}{3})^2 + (s_2 - \frac{s_1+s_2+s_3}{3})^2 + (s_3 - \frac{s_1+s_2+s_3}{3})^2 \leq \frac{(s_1+s_2+s_3)^2}{6}\}.$$

Then the mapping $(l_{12}, l_{13}, l_{23}) \mapsto (s_1, s_2, s_3)$ maps $\mathcal{Q}$ bijectively onto $\mathcal{C}$.

PROOF. Because the mapping $(l_{12}, l_{13}, l_{23}) \mapsto (s_1, s_2, s_3)$ maps $\{(l_{12}, l_{13}, l_{23}) | l_{12}, l_{13}, l_{23} \geq 0\}$ bijectively onto $\{(s_1, s_2, s_3) | s_1, s_2, s_3 \geq 0\}$, it is enough to show, for every point (l_{12}, l_{13}, l_{23}) on the boundary of $\mathcal{Q}$, its image is on the boundary of $\mathcal{C}$. Let C be the point on the intersection of the plane $s_1 + s_2 + s_3 = Q$ with the line $s_1 = s_2 = s_3$; then $C = (Q/3, Q/3, Q/3)$. Let $A = (s_1, s_2, s_3)$ be an arbitrary point on the intersection of the plane $s_1 + s_2 + s_3 = Q$ with the surface $\sqrt{s_1} + \sqrt{s_2} = \sqrt{s_3}$.

Note that

$$l_{12}^4 + l_{13}^4 + (l_{12}+l_{13})^4 = 2(l_{12}^4 + l_{13}^4) + 6l_{12}^2 l_{13}^2 + 4l_{12}^3 l_{13} + 4l_{12} l_{13}^3.$$

For the point (s_1, s_2, s_3) on the surface of $\sqrt{s_1} + \sqrt{s_2} = \sqrt{s_3}$,

$$\begin{aligned}
Q^2 &= (s_1 + s_2 + s_3)^2 \\
&= (l_{12}^2 + l_{13}^2 + l_{23}^2)^2 \\
&= [l_{12}^2 + l_{13}^2 + (l_{12}+l_{13})^2]^2 \\
&= l_{12}^4 + l_{13}^4 + (l_{12}+l_{13})^4 + 2l_{12}^2 l_{13}^2 + 2l_{12}^2(l_{12}+l_{13})^2 + 2l_{13}^2(l_{12}+l_{13})^2 \\
&= l_{12}^4 + l_{13}^4 + (l_{12}+l_{13})^4 + 2(l_{12}^4 + l_{13}^4) + 6l_{12}^2 l_{13}^2 + 4l_{12}^3 l_{13} + 4l_{12} l_{13}^3 \\
&= 2[l_{12}^4 + l_{13}^4 + (l_{12}+l_{13})^4] \text{ (by the previous equality).} \\
\Rightarrow & l_{12}^4 + l_{13}^4 + (l_{12}+l_{13})^4 = \frac{Q^2}{2}. \\
\|AC\|^2 &= (s_1 - \frac{Q}{3})^2 + (s_2 - \frac{Q}{3})^2 + (s_3 - \frac{Q}{3})^2 \\
&= (l_{12}^2 - \frac{Q}{3})^2 + (l_{13}^2 - \frac{Q}{3})^2 + (l_{23}^2 - \frac{Q}{3})^2 \\
&= {l_{12}}^4 + {l_{13}}^4 + {l_{23}}^4 - 2\frac{Q}{3}({l_{12}}^2 + {l_{13}}^2 + {l_{23}}^2) + 3(\frac{Q}{3})^2 \\
&= {l_{12}}^4 + {l_{13}}^4 + (l_{12}+l_{13})^4 - \frac{Q^2}{3}
\end{aligned}$$

(the above equality holds because that A is on the plane $l_{12}^2 + l_{13}^2 + l_{23}^2 = Q$

and A is on the surface $l_{12} + l_{13} = l_{23}$)

$$= \frac{Q^2}{6}.$$

Therefore $\|AC\| = \frac{Q}{\sqrt{6}}$ is a constant, and hence the intersection of the plane $s_1 + s_2 + s_3 = Q$ with each of the three surfaces $l_{12} + l_{13} = l_{23}, l_{13} + l_{23} = l_{12}$ and $l_{23} + l_{12} = l_{13}$ is a circular arc centered at $C = (Q/3, Q/3, Q/3)$ with radius $\frac{Q}{\sqrt{6}}$. So A is on the boundary of $\mathcal{C}$. In the other words, the collinear configurations are on the surface of the cone $\mathcal{C}$. □

The third coordinate system is (Q, s, θ) (see figure 10) where $Q \in [0, +\infty), s \in [0, 1], \theta \in [0, 2\pi]$, are defined as follows:

$$
\begin{aligned}
Q &= l_{12}^2 + l_{13}^2 + l_{23}^2 = s_1 + s_2 + s_3, \\
Qs\cos\theta &= 2l_{23}^2 - l_{12}^2 - l_{13}^2 = 2s_3 - s_1 - s_2, \\
Qs\sin\theta &= \sqrt{3}(l_{13}^2 - l_{12}^2) = \sqrt{3}(s_2 - s_1).
\end{aligned}
$$

We will explain the meaning of this coordinate system as follows. As we showed above, the intersection of the plane $s_1 + s_2 + s_3 = Q$ with $\mathcal{C}$ is a circle together with its interior. The radius of this circle is $\|AC\| = \frac{Q}{\sqrt{6}}$ (see figure 10).

We define $\mathbf{e}_1$ to be the unit vector pointing from $C(\frac{Q}{3}, \frac{Q}{3}, \frac{Q}{3})$ to $(0, 0, Q)$; then

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{(0, 0, Q) - (\frac{Q}{3}, \frac{Q}{3}, \frac{Q}{3})}{\|(0, 0, Q) - (\frac{Q}{3}, \frac{Q}{3}, \frac{Q}{3})\|} \\
&= \frac{(-\frac{Q}{3}, -\frac{Q}{3}, \frac{2Q}{3})}{Q\frac{\sqrt{6}}{3}} \\
&= \frac{1}{\sqrt{6}}(-1, -1, 2).
\end{aligned}
$$

We define $\mathbf{e}_3$ to be the unit vector pointing from C to O. then

$$
\begin{aligned}
\mathbf{e}_3 &= (-\frac{Q}{3}, -\frac{Q}{3}, -\frac{Q}{3})\frac{1}{\frac{Q}{\sqrt{3}}} \\
&= (-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}).
\end{aligned}
$$

We define

$$
\begin{aligned}
\mathbf{e}_2 &= \mathbf{e}_3 \times \mathbf{e}_1 \\
&= (-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}) \times \frac{1}{\sqrt{6}}(-1, -1, 2) \\
&= (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0).
\end{aligned}
$$

Let $\mathbf{v}$ be a vector pointing from C to $A(s_1, s_2, s_3)$, one point in the cone $\mathcal{C}$; then

$$
\mathbf{v} = (s_1, s_2, s_3) - (\frac{Q}{3}, \frac{Q}{3}, \frac{Q}{3})
$$

$$
\begin{aligned}
&= (s_1 - \frac{Q}{3}, s_2 - \frac{Q}{3}, s_3 - \frac{Q}{3}),\\
\|\mathbf{v}\| &= \sqrt{(s_1 - \frac{Q}{3})^2 + (s_2 - \frac{Q}{3})^2 + (s_3 - \frac{Q}{3})^2}.
\end{aligned}
$$

Let θ be the angle between $\mathbf{v}$ and $\mathbf{e}_1$, then

$$
\begin{aligned}
\mathbf{v} &= \mathbf{e}_1(\mathbf{e}_1 \cdot \mathbf{v}) + \mathbf{e}_2(\mathbf{e}_2 \cdot \mathbf{v})\\
&= \mathbf{e}_1\|\mathbf{v}\|\cos\theta + \mathbf{e}_2\|\mathbf{v}\|\sin\theta,\\
\|\mathbf{v}\|\cos\theta &= \mathbf{e}_1 \cdot \mathbf{v}\\
&= \frac{1}{\sqrt{6}}(-1, -1, 2)(s_1 - \frac{Q}{3}, s_2 - \frac{Q}{3}, s_3 - \frac{Q}{3})\\
&= \frac{1}{\sqrt{6}}(2s_3 - s_1 - s_2),\\
Qs\cos\theta &= 2s_3 - s_2 - s_1 = \sqrt{6}\|\mathbf{v}\|\cos\theta\\
\Rightarrow s &= \sqrt{6}\frac{\|\mathbf{v}\|}{Q}.
\end{aligned}
$$

If C is on the bounding circle $\|\mathbf{v}\| = \frac{Q}{\sqrt{6}}$ and $s = 1$. A similar calculation confirms that $\mathbf{e}_1 \times \mathbf{v} = \mathbf{e}_3\|\mathbf{e}_1\|\|\mathbf{v}\|\sin\theta$. In summary, in the third coordinate system (Q, s, θ), $Q = \sqrt{3}\|OC\|, s = \sqrt{6}\frac{\|\mathbf{v}\|}{Q}$, and θ is the angle between $\mathbf{e}_1$ and vector $\mathbf{v}$ measured as a positive rotation around the axis $\mathbf{e}_3$. This coordinate system is used in the papers [**2**], [**15**], [**45**] and [**41**].

Standard Configuration. In §6.3.1, we will find it necessary to introduce the concept of a standard configuration. Given a conformation, we want to define a particular configuration that is a member of that conformation, and we will use that configuration to define the Hamiltonian. Moreover, we will find that when a symmetry operation acts on a three-atom molecule, it is desirable to have the geometric center of the triangle to be the origin.

Definition. Given the conformation $b = (l_{12}, l_{13}, l_{23})$ of the non-collinear H_3 system, and given $\tilde{\mathbf{R}}_1(b) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tilde{\mathbf{R}}_2(b) = \begin{pmatrix} l_{12} \\ 0 \\ 0 \end{pmatrix}, \tilde{\mathbf{R}}_3(b) = \begin{pmatrix} l_{13}\cos\theta \\ l_{13}\sin\theta \\ 0 \end{pmatrix}$, where $\theta =$

$\cos^{-1}(\frac{l_{12}^2+l_{13}^2-l_{23}^2}{2l_{12}l_{13}})$, we define a *standard configuration* $(\mathbf{R}_1(b), \mathbf{R}_2(b), \mathbf{R}_3(b))$ of the H_3 system by giving the three nuclei $(\tilde{\mathbf{R}}_1(b), \tilde{\mathbf{R}}_2(b), \tilde{\mathbf{R}}_3(b))$ specific new coordinates as follows. The origin of the coordinates is at $\tilde{\mathbf{R}}(b) = \frac{\tilde{\mathbf{R}}_1(b)+\tilde{\mathbf{R}}_2(b)+\tilde{\mathbf{R}}_3(b)}{3} = \begin{pmatrix} \frac{1}{3}(l_{12}+l_{13}\cos\theta) \\ \frac{1}{3}l_{13}\sin\theta \\ 0 \end{pmatrix}$. Thus define

$$\mathbf{R}_1(b) = \tilde{\mathbf{R}}_1(b) - \tilde{\mathbf{R}}(b) = \begin{pmatrix} -\frac{1}{3}(l_{12}+l_{13}\cos\theta) \\ -\frac{1}{3}l_{13}\sin\theta \\ 0 \end{pmatrix},$$

$$\mathbf{R}_2(b) = \tilde{\mathbf{R}}_2(b) - \tilde{\mathbf{R}}(b) = \begin{pmatrix} \frac{1}{3}(2l_{12}-l_{13}\cos\theta) \\ -\frac{1}{3}l_{13}\sin\theta \\ 0 \end{pmatrix},$$

$$\mathbf{R}_3(b) = \tilde{\mathbf{R}}_3(b) - \tilde{\mathbf{R}}(b) = \begin{pmatrix} \frac{1}{3}(-l_{12}+2l_{13}\cos\theta) \\ \frac{2}{3}l_{13}\sin\theta \\ 0 \end{pmatrix},$$

where we understand that $\cos\theta = \frac{l_{12}^2+l_{13}^2-l_{23}^2}{2l_{12}l_{13}}$, and $\sin\theta = \sqrt{1-(\frac{l_{12}^2+l_{13}^2-l_{23}^2}{2l_{12}l_{13}})^2}$. This θ is different from the θ in the previous section.

Therefore $(\mathbf{R}_1(b), \mathbf{R}_2(b), \mathbf{R}_3(b))$ is completely determined by the conformation (l_{12}, l_{13}, l_{23}) of the non-collinear H_3 system.

Chapter 6

Fiber Bundles and Connections in H_3 System

In this chapter we introduce the differential geometry concepts: fiber bundles, and the special case of Hermitian vector bundles, on which we define connections, parallel translation, and the covariant derivative. We illustrate these concepts and results in two concrete cases: a surface in $\mathbb{R}^3$ and the H_3 system.

6.1. Fiber Bundles and Connections

Definition. A *fiber* of a map $f : X \to Y$ is the pre-image of a point $y \in Y$: $f^{-1}(y) = \{x \in X | f(x) = y\}$.

Definition. Suppose M is a Hausdorff topological space and $n \geq 1$ is an integer. Suppose $\{\phi_i\}_{i\in\mathcal{I}}$ is a collection of homeomorphisms (a bijective map between two topological spaces which is continuous in both directions) from open sets of M to open sets of $\mathbb{R}^n$, where

(1) $\forall i \in \mathcal{I}, \phi_i : \mathrm{dom}(\phi_i) \to \mathrm{codom}(\phi_i)$, $\mathrm{dom}(\phi_i)$ is open in M, and $\mathrm{codom}(\phi_i)$ is open in $\mathbb{R}^n$.

(2) $M = \cup_{i\in\mathcal{I}} \mathrm{dom}(\phi_i)$.

(3) For all $i, j \in \mathcal{I}$, $\phi_i \circ \phi_j^{-1} : \phi_j(\mathrm{dom}(\phi_i) \cap \mathrm{dom}(\phi_j)) \to \phi_i(\mathrm{dom}(\phi_i) \cap \mathrm{dom}(\phi_j))$ is C^∞ smooth; notice that $\phi_i(\mathrm{dom}(\phi_i) \cap \mathrm{dom}(\phi_j))$ is open in $\mathbb{R}^n, \forall i, j \in \mathcal{I}$.

(4) (Maximality) If U is open in M and V is open in $\mathbb{R}^n$ and $\phi : U \to V$ is a homeomorphism and if for all $i \in \mathcal{I}$, $\phi_i \circ \phi^{-1} : \phi(\mathrm{dom}(\phi_i) \cap U) \to \phi_i(\mathrm{dom}(\phi_i) \cap U)$ is a diffeomorphism (a map between open subsets of $\mathbb{R}^n$

which is infinitely differentiable and has a infinitely differentiable inverse), then there exists $j \in \mathcal{I}$ such that $\phi = \phi_j$.

Then $(M, \{\phi_i\}_{i\in\mathcal{I}})$ is called a n-dimensional *smooth manifold.* If it will not cause confusion, we denote a n-dimensional smooth manifold as M_n or just M when the dimensionality of the manifold is not in focus.

Definition. Suppose m-dimensional $(M, \{\phi_i\}_{i\in\mathcal{I}})$ and n-dimensional $(N, \{\psi_i\}_{j\in\mathcal{J}})$ are smooth manifolds and $f : M_m \to N_n$ satisfies:

(1) f is continuous.

(2) $\forall i \in \mathcal{I}, \forall j \in \mathcal{J}$, $\psi_j \circ f \circ \phi_i^{-1} : \phi_i(\mathrm{dom}(\phi_i) \cap f^{-1}(\mathrm{dom}(\psi_j))) \subset \mathbb{R}^m \to \mathbb{R}^n$ is smooth.

Then f is called a *smooth map* between M_m and N_n.

Definition. If $f : M_m \to N_n$ is a bijective smooth map between smooth manifolds M_m and N_n and $f^{-1} : N_n \to M_m$ is a smooth map also, then f is a *diffeomorphism*. This implies that $n = m$.

Definition. Let E, B and F be three smooth manifolds; let $\pi : E \to B$ be a smooth map; let U be an open subset of B. A *local trivialization* over U is a diffeomorphism $\tau : U \times F \to \pi^{-1}(U)$ s.t. $\forall b \in U, \forall y \in F, \pi(\tau(b, y)) = b$ (see figure 11).

Definition. Let G be a Lie group (a smooth manifold which is also a group and which satisfies the additional condition that the group operations are smooth) and F be a smooth manifold. A *(smooth) left action* of G on F is a (smooth) map $G \times F \to F$ defined as $(g, y) \mapsto g \cdot y$ which satisfies $g \cdot (h \cdot y) = (gh) \cdot y$ and $e \cdot y = y$, for all $y \in F$, where e is the identity of G. A left action of G on F is *faithful* if for any $g_1, g_2 \in G, g_1 \neq g_2$, then $g_1 \cdot y \neq g_2 \cdot y$, for some $y \in F$.

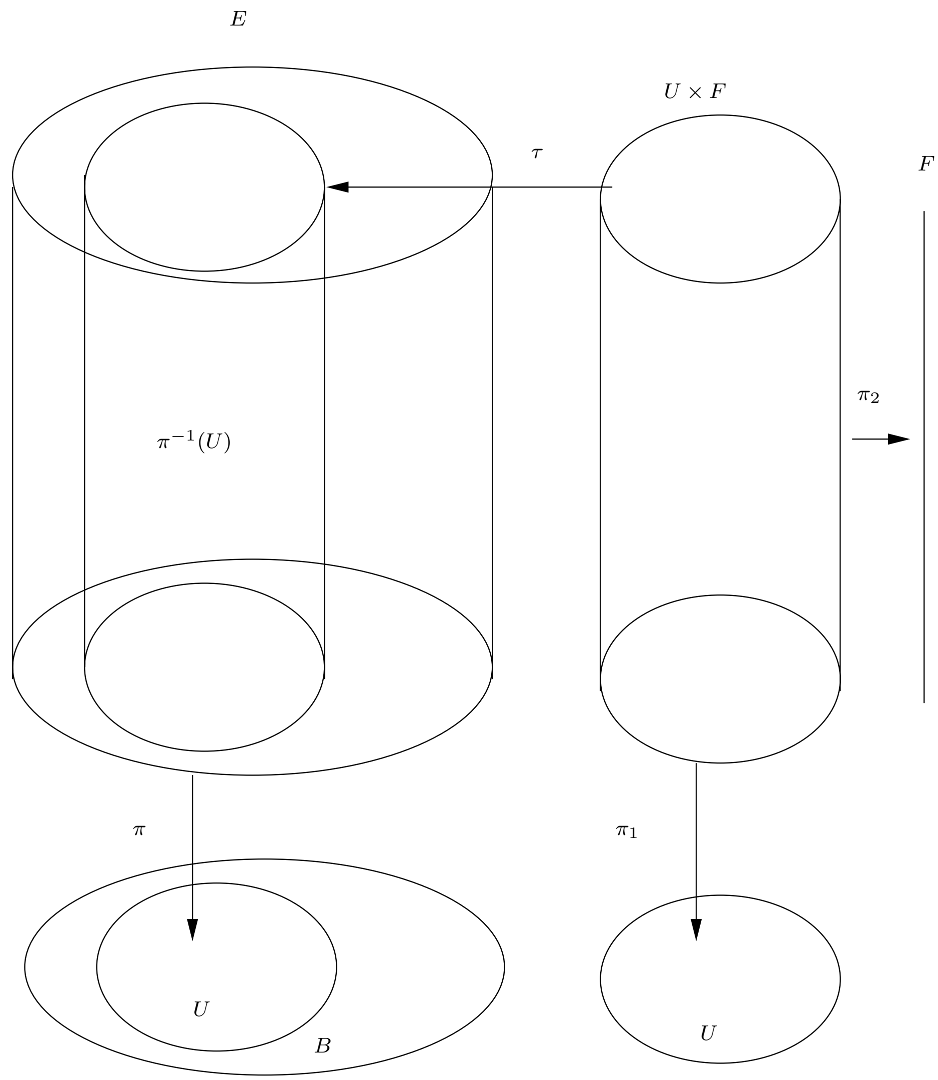

FIGURE 11. Local Trivialization τ Over U .

DEFINITION. Let E be a smooth manifold, called the *total space*; let B be a smooth manifold, called the *base space*; let $\pi : E \to B$ be a smooth map; let F be a smooth manifold, called the *standard fiber*; let G be a Lie group, called the *structure group*; assume that $G \times F \to F$ defines a smooth left action of G on F; let $\{\tau_i : U_i \times F \to \pi^{-1}(U_i)\}_{i \in \mathcal{I}}$ be a family of local trivializations satisfy $\pi(\tau_i(b, y)) = b, \forall i \in \mathcal{I}, b \in U_i$, where $\{U_i\}_{i \in \mathcal{I}}$ is an open covering of B; $(E, B, \pi, F, G, \{\tau_i\}_{i \in \mathcal{I}})$ determines a *fiber bundle* with standard fiber F and structure group G if there is a family $\{g_{ij} : U_i \cap U_j \to G\}_{(i,j) \in \mathcal{I}^2}$ of smooth maps such that $\forall b \in U_i \cap U_j, \forall y \in F, (\tau_i^{-1} \circ$

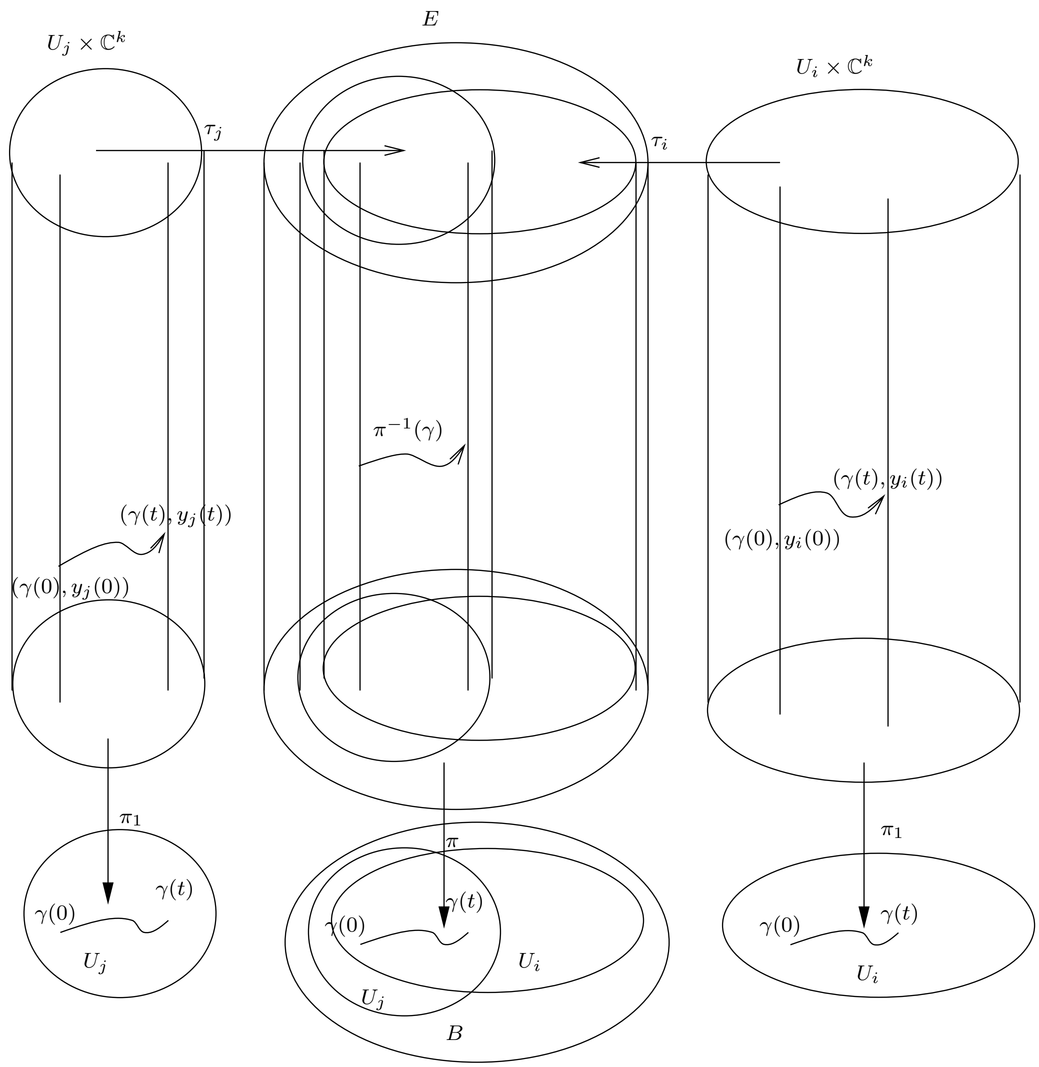

FIGURE 12. Two Overlapping trivializations and Parallel translation

$\tau_j)(b, y) = (b, g_{ij}(b) \cdot y)$. $\{\tau_i\}_{i\in\mathcal{I}}$ is said to be *a atlas of smooth local trivializations* with the smooth *cocycle* $\{g_{ij}\}_{(i,j)\in\mathcal{I}^2}$ (see figure 12). Suppose $\{\tau_i\}_{i\in\mathcal{I}}$ and $\{\tilde{\tau}_j\}_{j\in\mathcal{J}}$ are atlases of local trivializations with G-valued cocycles, then we say $\{\tau_i\}_{i\in\mathcal{I}} \overset{G}{\sim} \{\tau_j\}_{j\in\mathcal{J}}$ iff $\{\tau_i\}_{i\in\mathcal{I}} \cup \{\tilde{\tau}_j\}_{j\in\mathcal{J}}$ has a G-valued cocycle. $\overset{G}{\sim}$ is an equivalence relation on the set of atlases of local trivializations of fiber bundles with structure group G. We say $(E, B, \pi, F, G, \{\tau_i\}_{i\in\mathcal{I}})$ and $(E, B, \pi, F, G, \{\tilde{\tau}_j\}_{j\in\mathcal{J}})$ determine the same fiber bundle with standard fiber F and structure group G iff $\{\tau_i\}_{i\in\mathcal{I}} \overset{G}{\sim} \{\tilde{\tau}_j\}_{j\in\mathcal{J}}$.

In particular, if $F = \mathbb{C}^k$ and $G = \mathfrak{U}(k) := \{U \in \mathbb{C}^{k\times k} | U^\dagger U = I\}$ and $G \times F \to F$ is the usual left action of $\mathfrak{U}(k)$ on $\mathbb{C}^k$ (matrix-vector multiplication), then a fiber bundle $(E, B, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$ with standard fiber $\mathbb{C}^k$ and structure group $\mathcal{U}(k)$ is called a *Hermitian vector bundle*.

In the case of a Hermitian vector bundle, the local trivialization τ_i determines a field of "orthonormal" bases of the fibers in $\pi^{-1}(U_i)$ in that $\forall b \in U_i, \pi^{-1}(b)$ has the basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\} := \{\tau_i(b, \hat{\mathbf{e}}_1), \cdots, \tau_i(b, \hat{\mathbf{e}}_k)\}$, where $\{\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_k\}$ is the standard basis of $\mathbb{C}^k$. $\pi^{-1}(b) = E_b$ is equipped with a unique $\mathbb{C}$-vector space structure and a unique inner product such that $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\}$ is an orthonormal basis. These structures are independent of τ_i such that $b \in U_i$.

Fact. If $g : (-\epsilon, 1 + \epsilon) \to \mathfrak{U}(k)$ is a smooth map with $g(0) = I$, then for all $t_0 \in [0, 1]$, we have $g'(t_0)g(t_0)^{-1} \in T_I\mathfrak{U}(k)$, where $T_I\mathfrak{U}(k)$ is the tangent space of $\mathfrak{U}(k)$ at I.

Proof. Fix $t_0 \in [0, 1]$. We define a map $F : \mathbb{C}^{k\times k} \to \mathbb{C}^{k\times k} : h \mapsto hg(t_0)^{-1}$. Since F is $\mathbb{C}$-linear, $DF(g(t_0)) : T_{g(t_0)}\mathbb{C}^{k\times k} \to T_I\mathbb{C}^{k\times k} : (g(t_0), \tilde{h}) \mapsto (I, \tilde{h}g(t_0)^{-1})$. Let $F_r : \mathfrak{U}(k) \to \mathfrak{U}(k)$ be the restriction of F to $\mathfrak{U}(k)$, then $DF_r(g(t_0)) : T_{g(t_0)}\mathfrak{U}(k) \to T_I\mathfrak{U}(k)$ is a restriction of $DF(g(t_0))$. Since $(g(t_0), g'(t_0)) \in T_{g(t_0)}\mathfrak{U}(k)$, we get $DF_r(g(t_0))(g(t_0), g'(t_0)) = (I, g'(t_0)g(t_0)^{-1}) \in T_I\mathfrak{U}(k)$. □

Remark: Formally $T_g\mathfrak{U}(k)$ consists of pairs (g, h), but we often informally write $h \in T_g\mathfrak{U}(k)$.

Definition. A *Lie algebra* is a vector space V equipped with a bilinear *Lie bracket* $[\cdot, \cdot] : V \times V \to V$ which satisfies $\forall A, B, C \in V$,

(1) antisymmetry: $[A, B] = -[B, A]$.

(2) Jacobi identity: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

FACT. $\mathfrak{u}(k) := T_I\mathfrak{U}(k) = \{A \in \mathbb{C}^{k\times k} | A^\dagger = -A\}$ is a vector subspace of $\mathbb{C}^{k\times k}$, which is also closed under the commutator: $[A,B] = AB - BA \in \mathfrak{u}(k), \forall A, B \in \mathfrak{u}(k)$. So $\mathfrak{u}(k)$ is a Lie algebra.

FACT. If $\omega : (-\epsilon, 1+\epsilon) \to \mathfrak{u}(k)$ is a smooth map and $g : (\epsilon, 1+\epsilon) \to \mathbb{C}^{k\times k}$ is the solution of the ODE initial value problem: $g'(t) = \omega(t)g(t), t \in (-\epsilon, 1+\epsilon), g(0) = I$, then $g(t_0) \in \mathfrak{U}(k), \forall t_0 \in [0,1]$.

PROOF. $\frac{d}{dt}[g(t)^\dagger g(t)] = g'(t)^\dagger g(t) + g(t)^\dagger g'(t) = [\omega(t)g(t)]^\dagger g(t) + g(t)^\dagger \omega(t) g(t) = g(t)^\dagger[\omega(t)^\dagger + \omega(t)]g(t) = 0$. Since $g(0)^\dagger g(0) = I$ we have $g(t_0)^\dagger g(t_0) = I, \forall t_0 \in [0,1]$. i.e. $g(t_0) \in \mathfrak{U}(k), \forall t_0 \in [0,1]$. □

DEFINITION. If U is an open subset of $\mathbb{R}^n$, $TU = U \times \mathbb{R}^n$ is its tangent space, W is a $\mathbb{R}$-vector space, then a *W-valued 1-form* is a smooth map, $\omega : TU \to W$ s.t. $\forall b \in U, T_bU \to W : (b,v) \mapsto \omega(b,v)$ is $\mathbb{R}$-linear.

In the following we will introduce the concepts of parallel translation and connection in the context of Hermitian vector bundles instead of general fiber bundles. The reason is that we will focus on the two systems: the H_3 system and the surface in $\mathbb{R}^3$, in both of which cases Hermitian vector bundles are to be defined. These concepts for the fiber bundle can be developed in a similar way.

DEFINITION. Let $\tau_i : U_i \times \mathbb{C}^k \to \pi^{-1}(U_i)$ be a local trivialization of the Hermitian vector bundle $(E, B, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$ where $U_i \subset B$ is open. Using coordinate charts on B we will assume U_i is open in $\mathbb{R}^n$. Let $\omega_i : TU_i \to \mathfrak{u}(k)$ be a $\mathfrak{u}(k)$-valued 1-form. Let $\gamma : (-\epsilon, 1+\epsilon) \to U_i$ be a smooth path and let $g_i : (-\epsilon, 1+\epsilon) \to \mathfrak{U}(k)$ solve the ODE initial value problem $g_i'(t) = -\omega_i(\gamma(t), \gamma'(t))g_i(t), g_i(0) = I$. Then the *parallel translation* map over γ for the local trivialization τ_i is an isometric $\mathbb{C}$-linear map $h_\gamma^{(i)} : \pi^{-1}(\{\gamma(0)\}) \to \pi^{-1}(\{\gamma(1)\})$ such that $h_\gamma^{(i)}(\tau_i(\gamma(0), y)) = \tau_i(\gamma(1), g_i(1)\cdot y)$ for all $y \in \mathbb{C}^k$.

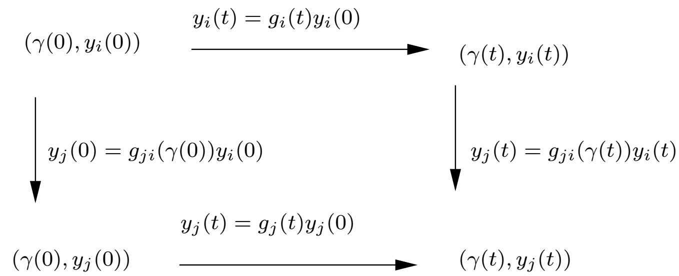

FIGURE 13. Parallel translation in two local trivializations

The parallel translation map above is defined on a curve that is entirely in the open set of a local trivialization. Now we want to extend the definition of the parallel translation map over an arbitrary smooth curve, which is independent of the covering of the curve.

First, we must show that $h_\gamma^{(i)} = h_\gamma^{(j)}$ for a curve γ lying entirely in the open set $U_i \cap U_j$ of two local trivializations τ_i and τ_j. Thus we can write $h_\gamma = h_\gamma^{(i)}$ whenever range$(\gamma) \in U_i$.

FACT. Suppose $i, j \in \mathcal{I}, i \neq j$; and $U_i \cap U_j \neq \emptyset$. Then the connection 1-forms ω_i and ω_j determine the same parallel translation map over each smooth path $\gamma : (-\epsilon, 1+\epsilon) \to U_i \cap U_j$ if and only if ω_i and ω_j satisfy the "transformation rule" i.e. for all $(x, v) \in T(U_i \cap U_j)$,

$$\omega_j(x, v) = -Dg_{ji}(x)(v)g_{ji}(x)^{-1} + g_{ji}(x)\omega_i(x, v)g_{ji}(x)^{-1}.$$

PROOF. ($\Rightarrow$) Let $(x, v) \in T(U_i \cap U_j)$ and $\gamma : (-\epsilon, 1+\epsilon) \to T(U_i \cap U_j)$ be a smooth curve such that $(\gamma(0), \gamma'(0)) = (x, v)$. Suppose the local connection 1-forms ω_i and ω_j determine the same parallel translation map over γ. Then the two ways from $y_i(0)$ to $y_j(t)$ pictured in figure 13 give the same $y_j(t)$.

Way 1: $y_j(t) = g_{ji}(\gamma(t))y_i(t)$, where $y_i(t) = g_i(t) \cdot y_i(0)$. $\Rightarrow y_j(t) = g_{ji}(\gamma(t))g_i(t) \cdot y_i(0)$, where $g_i(t)$ solves $g_i'(t) = -\omega_i(\gamma(t), \gamma'(t))g_i(t), g_i(0) = I$.

Way 2: $y_j(t) = g_j(t) \cdot y_j(0)$, where g_j solves $g_j'(t) = -\omega_j(\gamma(t), \gamma'(t))g_j(t), g_j(0) = I$, $y_j(0) = g_{ji}(\gamma(0))y_i(0)$. $\Rightarrow y_j(t) = g_j(t)g_{ji}(\gamma(0))y_i(0)$.

Since $g_{ji}(\gamma(t))g_i(t)\cdot y_i(0) = g_j(t)g_{ji}(\gamma(0))y_i(0)$ holds for all $y_i(0)\in\mathbb{C}^k$, and that the left action of $\mathfrak{U}(k)$ on $\mathbb{C}^k$ is faithful, we have $g_{ji}(\gamma(t))g_i(t) = g_j(t)g_{ji}(\gamma(0)) \Rightarrow g_j(t)^{-1}g_{ji}(\gamma(t))g_i(t) = g_{ji}(\gamma(0))$. Differentiating both sides with respect to t, we have:

$$
\begin{aligned}
0 &= \frac{d}{dt}g_{ji}(\gamma(0)) = \frac{d}{dt}[g_j(t)^{-1}g_{ji}(\gamma(t))g_i(t)] \\
&= -g_j(t)^{-1}[g_j'(t)g_j(t)^{-1}]g_{ji}(\gamma(t))g_i(t) \\
&\quad + g_j(t)^{-1}[\frac{d}{dt}g_{ji}(\gamma(t))]g_i(t) + g_j(t)^{-1}g_{ji}(\gamma(t))g_i'(t) \\
&= g_j(t)^{-1}\omega_j(\gamma(t),\gamma'(t))g_{ji}(\gamma(t))g_i(t) \\
&\quad + g_j(t)^{-1}\{\frac{d}{dt}g_{ji}(\gamma(t)) - g_{ji}(\gamma(t))\omega_i(\gamma(t),\gamma'(t))\}g_i(t) \\
&= g_j(t)^{-1}\{\omega_j(\gamma(t),\gamma'(t))g_{ji}(\gamma(t)) + \frac{d}{dt}g_{ji}(\gamma(t)) - g_{ji}(\gamma(t))\omega_i(\gamma(t),\gamma'(t))\}g_i(t).
\end{aligned}
$$

Therefore $\omega_j(\gamma(t),\gamma'(t))g_{ji}(\gamma(t)) + \frac{d}{dt}g_{ji}(\gamma(t)) - g_{ji}(\gamma(t))\omega_i(\gamma(t),\gamma'(t)) = 0$.

In particular, if $t=0$, we have
$\omega_j(\gamma(0),\gamma'(0)) = -\frac{d}{dt}g_{ji}(\gamma(t))|_{t=0}g_{ji}(\gamma(0))^{-1} + g_{ji}(\gamma(0))\omega_i(\gamma(0),\gamma'(0))g_{ji}(\gamma(0))^{-1}$.

($\Leftarrow$) Suppose that the local connection 1-forms $\{\omega_i\}_{i\in\mathcal{I}}$, satisfy the *transformation rule*. Since the derivation of the above proof is invertible, it is clear that inverse claim holds. □

Second, we claim that the parallel translation map h_γ satisfies a subdivision property. To see, assume that $\gamma([0,1])\subset U_i, \gamma_1 = \gamma|_{[0,t_0]}, \gamma_2 = \gamma|_{[t_0,1]}$, where t_0 is arbitrary in $(0,1)$. Assume that the $g_1(t_0)$ corresponding to h_{γ_1} solves $g_1'(t) = -\omega_i(\gamma(t),\gamma'(t))g_1(t), t\in[0,t_0]; g_1(0)=I$; and the $g_2(1)$ corresponding to h_{γ_2} solves $g_2'(t) = -\omega_i(\gamma(t),\gamma'(t))g_2(t), t\in[t_0,1]; g_2(t_0)=I$. Let $h_\gamma = h_{\gamma_2}\circ h_{\gamma_1}$, then the corresponding $g(t) = \begin{cases} g_1(t) & \text{if } t\in[0,t_0]; \\ g_2(t)g_1(t_0) & \text{if } t\in[t_0,1] \end{cases}$, it solves $g'(t) = -\omega_i(\gamma(t),\gamma'(t))g(t), t\in[0,1]; g(0)=I$. □

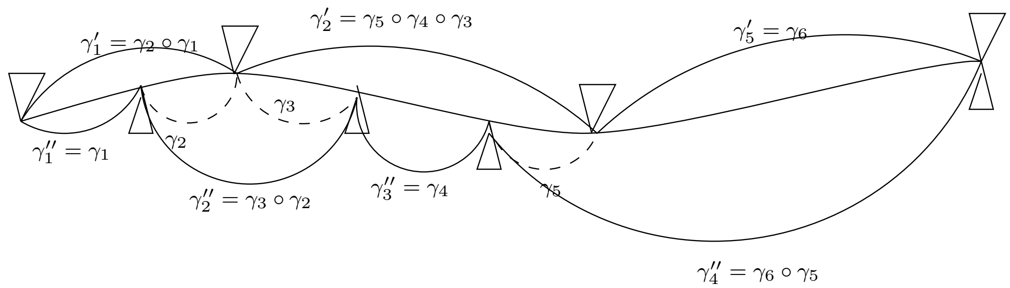

FIGURE 14. Parallel translation independent of local trivialization index i.

Here we showed that the parallel translation map h_γ is independent of the bi-subdivision of the path γ. The claim also holds for any finite subdivision of the path γ by applying the proven result recursively finitely many times.

For the third step, consider an arbitrary path $\gamma : [0,1] \to B$. Consider two coverings $\{U_i\}_{i=1}^m$ and $\{V_j\}_{j=1}^n$ of γ, corresponding to two families of trivializations $\{\tau_i'\}_{i=1}^m$ and $\{\tau_j''\}_{j=1}^n$ respectively taken from the same atlas. Assume that $0 = c_0 < \cdots < c_m = 1$ and $\gamma|[c_{i-1}, c_i] \subset U_i$; $0 = d_0 < \cdots < d_m = 1$ and $\gamma([d_{j-1}, d_j]) \subset V_j$. Assume that the local connection 1-forms $\{\omega_i\}_{i\in\mathcal{I}}$ satisfy the transformation law. For any $i = 1, \cdots, m$, define $h_{\gamma_i'}$ to be the parallel translation map over $\gamma_i' := \gamma|[c_{i-1}, c_i]$. Define $h_\gamma' := h_{\gamma_m'} \circ h_{\gamma_{m-1}'} \circ \cdots \circ h_{\gamma_1'}$. Also define $h_\gamma'' := h_{\gamma_n''} \circ h_{\gamma_{n-1}''} \circ \cdots \circ h_{\gamma_1''}$, where $h_{\gamma_j''}$ is the parallel translation map over $\gamma_j'' := \gamma|[d_{j-1}, d_j]$ (see figure 14).

We claim that $h_\gamma' = h_\gamma''$. To see, let $\{e_k\}_{k=0}^p = \{c_i\}_{i=0}^m \cup \{d_j\}_{j=0}^n$, and assume that $0 = e_0 < \cdots < e_p = 1$, i.e. $\{e_k\}_{k=0}^p$ is a finer subdivision of γ than $\{c_i\}_{i=0}^m$ and $\{d_j\}_{j=0}^n$. By the subdivision property each $h_{\gamma_i'}$ is a composition of $h_{\gamma_k}, k \in I_i$ and $\{1, \cdots, p\} = \cup_{i=1}^m I_i$, $I_i \cap I_j = \emptyset$ if $i \neq j$. Therefore $h_\gamma' = h_{\gamma_p} \circ \cdots \circ h_{\gamma_1}$. Similarly, $h_\gamma'' = h_{\gamma_p} \circ \cdots \circ h_{\gamma_1}$, so $h_\gamma' = h_\gamma''$.

Here we show how to calculate $h_\gamma : \pi^{-1}(\{\gamma(0)\}) \to \pi^{-1}(\{\gamma(1)\})$ along γ. For $0 = t_0 < t_1 < \cdots < t_k = 1$, suppose $\gamma|[t_{i-1}, t_i] \in U_i$ of local trivialization τ_i, $i = 1, \cdots, k$. We denote $h_\gamma[\tau_1(\gamma(0), y)] = \tau_k(\gamma(1), h_{\gamma,1k} \cdot y), \forall y \in F$, where $h_{\gamma,1k}$ is computed as follows: along $\gamma([t_{i-1}, t_i]) \subset U_i$, $g_i(t)$ solves the ODE initial value problem $g_i'(t) = -\omega_i(\gamma(t), \gamma'(t))g_i(t), t \in [t_{i-1}, t_i]; g_i(t_{i-1}) = I$. Then

$$h_{\gamma,1k} = g_k(t_k) \cdots [g_{32}(\gamma(t_2))g_2(t_2)][g_{21}(\gamma(t_1))g_1(t_1)].$$

In particular, if the curve is a smooth loop, $\gamma(0) = \gamma(1)$, we denote $h_{\gamma,1} := g_{1k}(\gamma(0))h_{\gamma,1k}$ corresponding to the parallel translation map h_γ in the trivialization τ_1.

Definition. A *connection* is defined in the Hermitian vector bundle $(E, B, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$ by a family $\{(U_i, \tau_i, \omega_i)\}$ where $\{U_i\}_{i\in\mathcal{I}}$ is an open covering of B, $\{\tau_i : U_i \times \mathbb{C}^k \to \pi^{-1}(U_i)\}_{i\in\mathcal{I}}$ is a family of local trivializations with a smooth $\mathfrak{u}(k)$-valued cocycle $\{g_{ij}\}_{(i,j)\in\mathcal{I}^2}$, and $\{\omega_i : TU_i \to \mathfrak{u}(k)\}_{i\in\mathcal{I}}$ is a family of smooth $\mathfrak{u}(k)$-valued 1-forms, satisfying the transformation laws.

Definition. Let B be the base space, E the total space and $\pi : E \to B$ is the bundle map. A *(global) section* is a smooth map $\psi : B \to E$ s.t. $\forall b \in B, \psi(b) \in \pi^{-1}(\{b\})$. A *local section* is a smooth map $\psi : U \to E$ s.t. $\forall b \in U, \psi(b) \in \pi^{-1}(\{b\})$, where $U \subset B$ is open.

Here we want to show how to take the covariant derivative of a section ψ in a given direction. We use the connection to parallel translate $\psi(\gamma(t))$ from $\pi^{-1}(\gamma(t))$ back to $\pi^{-1}(\gamma(0))$ as follows (see figure 15): for $|t| < \epsilon$, $\gamma(t) \in U_i, \psi(\gamma(t)) = \tau_i(\gamma(t), y_i(t)) \in \pi^{-1}(\{\gamma(t)\})$, where $y_i(t) := \pi_2\tau_i^{-1}(\psi(\gamma(t))) \in \mathbb{C}^k$.

The *covariant derivative* of ψ in the direction $(\gamma(0), \gamma'(0))$ is defined as

$$\begin{aligned}
\nabla_{(\gamma(0),\gamma'(0))}\psi &= \lim_{t\to 0} \frac{[\text{ parallel translation of } \psi(\gamma(t)) \text{ back to } \pi^{-1}(\gamma(0))] - \psi(\gamma(0))}{t} \\
&= \lim_{t\to 0} \frac{\tau_i(\gamma(0), g_i(t)^{-1}y_i(t)) - \tau_i(\gamma(0), y_i(0))}{t} \\
&= \lim_{t\to 0} \tau_i \left(\gamma(0), \frac{g_i(t)^{-1}y_i(t) - y_i(0)}{t}\right) \\
&= \tau_i \left(\gamma(0), \lim_{t\to 0} \frac{g_i(t)^{-1}y_i(t) - y_i(0)}{t}\right) \\
&= \tau_i \left(\gamma(0), \frac{d}{dt}[g_i(t)^{-1}y_i(t)]|_{t=0}\right)
\end{aligned}$$

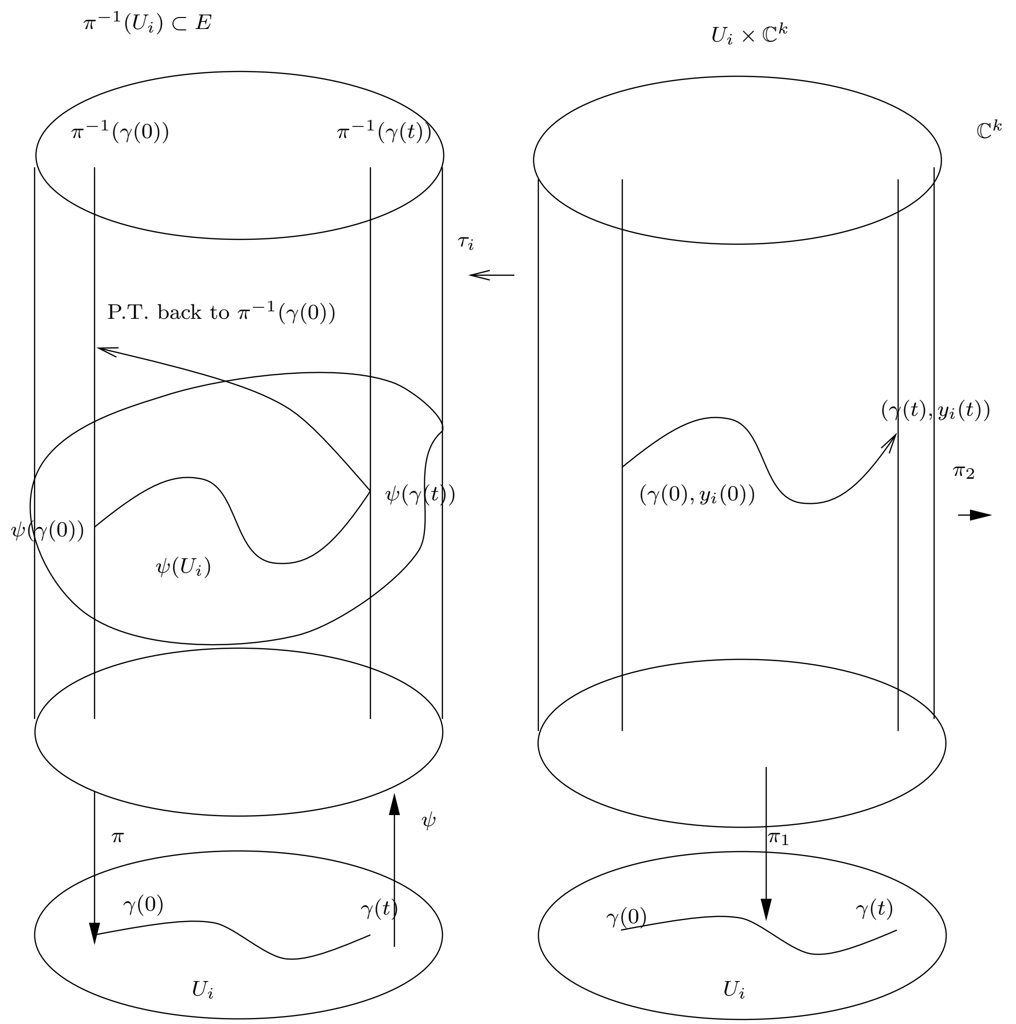

FIGURE 15. Use Connection to Parallel Translate $\psi(\gamma(t))$ from $\pi^{-1}(\gamma(t))$ back to $\pi^{-1}(\gamma(0))$

Because $g_i(t)^{-1}$ can be solved from $g_i'(t) = -\omega_i(\gamma(t), \gamma'(t))g_i(t), g_i(0) = I$, where ω_i is known, we have

$$\begin{aligned}
\frac{d}{dt}(g_i(t)^{-1}y_i(t)) &= \frac{d}{dt}[g_i(t)^{-1}]y_i(t) + g_i(t)^{-1}\frac{d}{dt}y_i(t) \\
&= -g_i(t)^{-1}\frac{d}{dt}[g_i(t)]g_i(t)^{-1}y_i(t) + g_i(t)^{-1}\frac{d}{dt}y_i(t) \\
&= g_i(t)^{-1}[\omega_i(\gamma(t), \gamma'(t))y_i(t) + \frac{d}{dt}y_i(t)] \\
\frac{d}{dt}(g_i(t)^{-1}y_i(t))|_{t=0} &= \omega_i(\gamma(0), \gamma'(0))y_i(0) + \frac{d}{dt}y_i(t)|_{t=0}
\end{aligned}$$

$$\Rightarrow \nabla_{(\gamma(0),\gamma'(0))}\psi = \tau_i\left(\gamma(0), \omega_i(\gamma(0),\gamma'(0))y_i(0) + [\frac{d}{dt}y_i(t)]|_{t=0}\right).$$

The transformation law of connection 1-forms implies this formula gives a result independent of the local trivialization.

FACT.

$$\begin{aligned} &g_{ji}(\gamma(0))[\frac{d}{dt}y_i(t)|_{t=0} + \omega_i(\gamma(0),\gamma'(0))y_i(0)] \\ &= \frac{d}{dt}y_j(t)|_{t=0} + \omega_j(\gamma(0),\gamma'(0))y_j(0). \end{aligned}$$

PROOF. It is known that :

$$y_j(t) = g_{ji}(\gamma(t))y_i(t),$$

$$\omega_j(\gamma(0),\gamma'(0)) = g_{ji}(\gamma(0))\omega_i(\gamma(0),\gamma'(0))g_{ji}(\gamma(0))^{-1} - [\frac{d}{dt}g_{ji}(\gamma(t))]|_{t=0}g_{ji}(\gamma(0))^{-1}$$

Plug them into the right hand side of the identity to be proven, we have

$$\begin{aligned} R.H.S. =&\frac{d}{dt}[g_{ji}(\gamma(t))y_i(t)]|_{t=0} + \{g_{ji}(\gamma(0))\omega_i(\gamma(0),\gamma'(0))g_{ji}(\gamma(0))^{-1} \\ &- [\frac{d}{dt}g_{ji}(\gamma(t))]|_{t=0}g_{ji}(\gamma(0))^{-1}\}y_j(0) \\ =&\frac{d}{dt}[g_{ji}(\gamma(t))]|_{t=0}y_i(0) + g_{ji}(\gamma(0))\frac{d}{dt}[g_i(\gamma(t))]|_{t=0} \\ &+ g_{ji}(\gamma(0))\omega_i(\gamma(0),\gamma'(0))y_i(0) - \frac{d}{dt}[g_{ji}(\gamma(t))]|_{t=0}y_i(0) \\ =&g_{ji}(\gamma(0))[\frac{d}{dt}[g_i(\gamma(t))]|_{t=0} + \omega_i(\gamma(0),\gamma'(0))y_i(0)] \\ =&L.H.S. \quad \square \end{aligned}$$

6.2. HERMITIAN VECTOR BUNDLES AND CONNECTIONS ON A GRAPH IN $\mathbb{R}^3$

Let's look at the concepts introduced above in the case of a surface in $\mathbb{R}^3$.

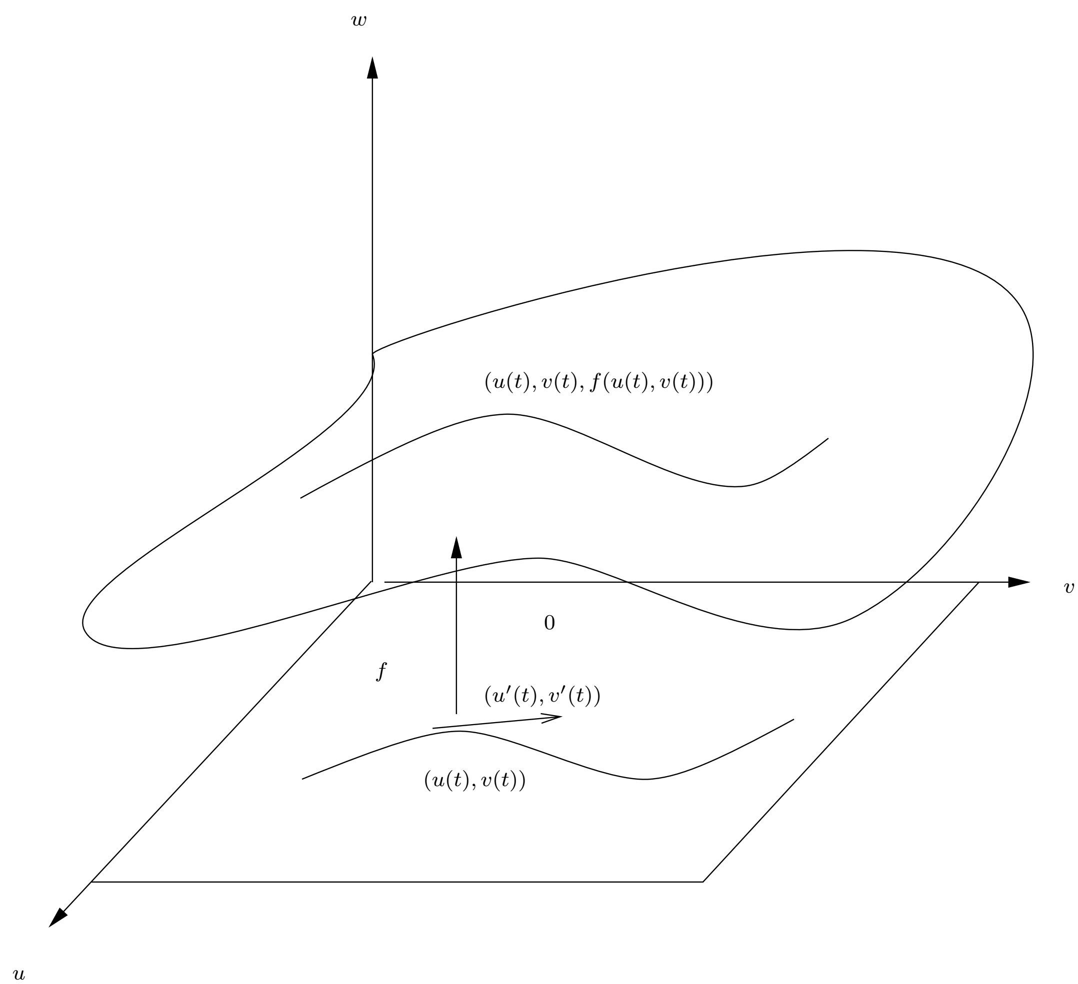

FIGURE 16. Fiber Bundles and Connections in the Case of a Surface in $\mathbb{R}^3$

6.2.1. Three Trivializations and the corresponding Cocycles.

EXAMPLE. Let B be the graph of $f(u,v)$, i.e. $B = \{\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix} \in \mathbb{R}^3 | (u,v)^T \in \mathbb{R}^2\}$. Let $E = TB = \cup_{b \in B} T_b B = \cup_{b \in B} \{(b, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}) | v_3 = v_1 f_u(u,v) + v_2 f_v(u,v), v_1, v_2 \in \mathbb{R}$ where $b = \begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}\}$, the tangent space of B. Let $\pi : E \to B : (\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}) \mapsto \begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}$. Let $F = \mathbb{R}^2$, $G = GL(2, \mathbb{R}^2)$, and the left action $G \times F \to F$ of G on F is matrix-vector multiplication. We define global trivialization $\tau_0 : B \times \mathbb{R}^2 \to \pi^{-1}(B) = E$ as $\tau_0(b, \binom{v_1}{v_2}) = (b, v_1 \begin{pmatrix} 1 \\ 0 \\ f_u(u,v) \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ f_v(u,v) \end{pmatrix})$, then we can define the cocycle $g_{00}(b) = I_{2 \times 2}$ for all $b \in B$.

Example. We use the same E, B, π, F as the ones in the previous example. Let $G = O(2,\mathbb{R})$, the 2×2 orthogonal matrices. Let $U_1 = U_2 = B$. By orthonormalizing the basis of the previous example in two ways, we can define two other global trivializations $\tau_i : U_i \times \mathbb{R}^2 \to E = TB, i = 1, 2$ as follows:

$$\tau_1(\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}) = (\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}, \frac{v_1}{\sqrt{1+f_u^2}}\begin{pmatrix} 1 \\ 0 \\ f_u \end{pmatrix} + \frac{v_2}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\begin{pmatrix} -f_u f_v \\ 1+f_u^2 \\ f_v \end{pmatrix}),$$

$$\tau_2(\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}) = (\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}, \frac{v_1}{\sqrt{1+f_v^2}}\begin{pmatrix} 0 \\ 1 \\ f_v \end{pmatrix} + \frac{v_2}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}}\begin{pmatrix} 1+f_v^2 \\ -f_u f_v \\ f_u \end{pmatrix}).$$

Here we compute the cocycle $g_{01}(b)$, which is defined by $(\tau_0^{-1} \circ \tau_1)(b, \begin{pmatrix} v_1' \\ v_2' \end{pmatrix}) = (b, g_{01}(b)\begin{pmatrix} v_1' \\ v_2' \end{pmatrix})$. i.e.

$$\begin{aligned}
\pi_2(\tau_1(b, \begin{pmatrix} v_1' \\ v_2' \end{pmatrix})) &= \frac{v_1'}{\sqrt{1+f_u^2}}\begin{pmatrix} 1 \\ 0 \\ f_u \end{pmatrix} + \frac{v_2'}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\begin{pmatrix} -f_u f_v \\ 1+f_u^2 \\ f_v \end{pmatrix} \\
&= \frac{v_1'}{\sqrt{1+f_u^2}}\mathbf{x}_u - \frac{v_2' f_u f_v}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\mathbf{x}_u + \frac{v_2'(1+f_u^2)}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\mathbf{x}_v \\
&= (\frac{v_1'}{\sqrt{1+f_u^2}} - \frac{v_2' f_u f_v}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}})\mathbf{x}_u + \frac{v_2'(1+f_u^2)}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\mathbf{x}_v \\
&= v_1\mathbf{x}_u + v_2\mathbf{x}_v \\
&= \pi_2(\tau_0(b, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix})) \\
&= \pi_2(\tau_0(b, g_{01}(b)\begin{pmatrix} v_1' \\ v_2' \end{pmatrix}))
\end{aligned}$$

$\mathbf{x}_u = \begin{pmatrix} 1 \\ 0 \\ f_u \end{pmatrix}$ and $\mathbf{x}_v = \begin{pmatrix} 0 \\ 1 \\ f_v \end{pmatrix}$. Then

$$\begin{aligned}
v_1 &= \frac{v_1'}{\sqrt{1+f_u^2}} - \frac{v_2' f_u f_v}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}} \\
v_2 &= \frac{v_2'(1+f_u^2)}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}} \\
\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= g_{01}(b)\begin{pmatrix} v_1' \\ v_2' \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{1+f_u^2}} & \frac{-f_u f_v}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}} \\ 0 & \frac{\sqrt{1+f_u^2}}{\sqrt{1+f_u^2+f_v^2}} \end{pmatrix} \begin{pmatrix} v_1' \\ v_2' \end{pmatrix}$$

$$\Rightarrow g_{01}(b) = \begin{pmatrix} \frac{1}{\sqrt{1+f_u^2}} & \frac{-f_u f_v}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}} \\ 0 & \frac{\sqrt{1+f_u^2}}{\sqrt{1+f_u^2+f_v^2}} \end{pmatrix}$$

Similarly we can compute the cocycle $g_{02}(b) = \begin{pmatrix} 0 & \frac{\sqrt{1+f_v^2}}{\sqrt{1+f_u^2+f_v^2}} \\ \frac{1}{\sqrt{1+f_v^2}} & \frac{-f_u f_v}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}} \end{pmatrix}$.

We compute g_{12} from g_{01}, g_{02}.

$$\begin{aligned} g_{12}(b) &= g_{10}(b) g_{02}(b) \\ &= g_{01}(b)^{-1} g_{02}(b) \\ &= \begin{pmatrix} \frac{1}{\sqrt{1+f_u^2}} & \frac{-f_u f_v}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}} \\ 0 & \frac{\sqrt{1+f_u^2}}{\sqrt{1+f_u^2+f_v^2}} \end{pmatrix}^{-1} \begin{pmatrix} 0 & \frac{1+f_v^2}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}} \\ \frac{1}{\sqrt{1+f_v^2}} & + \frac{-f_u f_v}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{f_u f_v}{\sqrt{(1+f_u^2)(1+f_v^2)}} & \frac{\sqrt{1+f_u^2+f_v^2}}{\sqrt{(1+f_u^2)(1+f_v^2)}} \\ \frac{\sqrt{1+f_u^2+f_v^2}}{\sqrt{(1+f_u^2)(1+f_v^2)}} & \frac{-f_u f_v}{\sqrt{(1+f_u^2)(1+f_v^2)}} \end{pmatrix} \end{aligned}$$

A simple check shows that $g_{12}(b) \in O(2)$. So $\{\tau_1, \tau_2\}$ determine a fiber bundle with structure group $O(2)$. Also $\{\tau_0\} \overset{GL(2)}{\sim} \{\tau_1, \tau_2\}$, so the two examples describe the same vector bundle.

6.2.2. Covariant derivative of a vector field. Let B be the graph of $f(u,v)$, i.e. $B = \{\begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix} \in \mathbb{R}^3 | (u,v)^T \in \mathbb{R}^2\}$. Let $E = TB = \cup_{b \in B} T_b B = \cup_{b \in B} \{(b, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} | v_3 = v_1 f_u(u,v) + v_2 f_v(u,v), v_1, v_2 \in \mathbb{R}$ where $b = \begin{pmatrix} u \\ v \\ f(u,v) \end{pmatrix}\}$, the tangent space of B. Then a vector field over an open set $U \subset B$ is a local section $\psi : U \to TB$ where $\psi(b) \in T_b B$ for all $b \in U$.

Definition. Let ψ be a differentiable vector field in an open set $U \subset B$. Let $\mathbf{y} \in T_b B, b \in U$. Consider a parametrized curve $\gamma : (-\epsilon, \epsilon) \to U$, with $\gamma(0) = b$ and $\gamma'(0) = \mathbf{y}$, and let $\psi(\gamma(t)), t \in (-\epsilon, \epsilon)$, be the restriction of the vector field ψ to the curve γ. The vector obtained by the orthogonal projection of $(d\psi/dt)(0)$ onto the plane T_bB is called the *covariant derivative* at b of the vector field ψ in the direction $\mathbf{y}$. This covariant derivative is denoted by $(\nabla_{\mathbf{y}}\psi)(b)$.

Using what we already know about the relation between the connection 1-form and the covariant derivative we now seek an expression for the connection 1-form associated to this sense of covariant derivative. To simplify the notation, we denote $\mathbf{e}_i(b)$ as $\mathbf{e}_i$ and $\tilde{\mathbf{e}}_i(b)$ as $\tilde{\mathbf{e}}_i$. Consider a general trivialization $\tau : U \times \mathbb{R}^2 \to E = TB$ as $\tau(b, (\begin{smallmatrix} u \\ v \end{smallmatrix})) = (b, \mathbf{e}_1 u + \mathbf{e}_2 v)$, where $E_b = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}$. Let $\psi : B \to E$ be an arbitrary vector field satisfying $\psi(b) \in E_b, \forall b \in B$ and $\psi(b) = \mathbf{e}_1 a_1 + \mathbf{e}_2 a_2$. A parametrized curved at b is $\gamma : (-\epsilon, \epsilon) \to B$ such that $b = \gamma(0) = \begin{pmatrix} u(0) \\ v(0) \\ f(u(0),v(0)) \end{pmatrix}$.

We consider the orthonormal basis $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2\}$ spanning E, generated by the Gram-Schmidt process starting from $\{\mathbf{e}_1, \mathbf{e}_2\}$:

$$\tilde{\mathbf{e}}_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|},\ \tilde{\mathbf{e}}_2 = \frac{\mathbf{e}_2 - \tilde{\mathbf{e}}_1[\tilde{\mathbf{e}}_1^T\mathbf{e}_2]}{\|\mathbf{e}_2 - \tilde{\mathbf{e}}_1[\tilde{\mathbf{e}}_1^T\mathbf{e}_2]\|}.$$

Then the orthogonal projection operator from $\mathbb{R}^3$ into E_b is

$$\begin{aligned}
P_b &= \tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T + \tilde{\mathbf{e}}_2\tilde{\mathbf{e}}_2^T \\
&= \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1} + \frac{\{\mathbf{e}_2 - \tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T\mathbf{e}_2\}}{\|\mathbf{e}_2 - \tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T\mathbf{e}_2\|}\frac{\{\mathbf{e}_2 - \tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T\mathbf{e}_2\}^T}{\|\mathbf{e}_2 - \tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T\mathbf{e}_2\|} \\
&= \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1} + \frac{\{\mathbf{e}_2 - \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1}\mathbf{e}_2\}}{\{\mathbf{e}_2 - \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1}\mathbf{e}_2\}^T}\frac{\{\mathbf{e}_2 - \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1}\mathbf{e}_2\}^T}{\{\mathbf{e}_2 - \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1}\mathbf{e}_2\}} \\
&= \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1} + \frac{\{\mathbf{e}_2 - \mathbf{e}_1\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1}\}}{\{\mathbf{e}_2^T - \mathbf{e}_1^T\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1}\}}\frac{\{\mathbf{e}_2^T - \mathbf{e}_1^T\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1}\}}{\{\mathbf{e}_2 - \mathbf{e}_1\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1}\}} \\
&= \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1} + \frac{\mathbf{e}_2\mathbf{e}_2^T - \mathbf{e}_2\mathbf{e}_1^T\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1} - \mathbf{e}_1\mathbf{e}_2^T\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1} + \mathbf{e}_1\mathbf{e}_1^T\frac{(\mathbf{e}_1^T\mathbf{e}_2)^2}{(\mathbf{e}_1^T\mathbf{e}_1)^2}}{\mathbf{e}_2\mathbf{e}_2^T - \mathbf{e}_2^T\mathbf{e}_1\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1} - \mathbf{e}_1^T\mathbf{e}_2\frac{\mathbf{e}_1^T\mathbf{e}_2}{\mathbf{e}_1^T\mathbf{e}_1} + \mathbf{e}_1^T\mathbf{e}_1\frac{(\mathbf{e}_1^T\mathbf{e}_2)^2}{(\mathbf{e}_1^T\mathbf{e}_1)^2}} \\
&= \frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1}\frac{(\mathbf{e}_2^T\mathbf{e}_2 - \frac{(\mathbf{e}_1^T\mathbf{e}_2)^2}{\mathbf{e}_1^T\mathbf{e}_1})}{(\mathbf{e}_2^T\mathbf{e}_2 - \frac{(\mathbf{e}_1^T\mathbf{e}_2)^2}{\mathbf{e}_1^T\mathbf{e}_1})} + \frac{\mathbf{e}_2\mathbf{e}_2^T(\mathbf{e}_1^T\mathbf{e}_1) - \mathbf{e}_2\mathbf{e}_1^T(\mathbf{e}_1^T\mathbf{e}_2) - \mathbf{e}_1\mathbf{e}_2^T(\mathbf{e}_1^T\mathbf{e}_2) + (\mathbf{e}_1^T\mathbf{e}_1)\frac{(\mathbf{e}_1^T\mathbf{e}_2)^2}{\mathbf{e}_1^T\mathbf{e}_1}}{(\mathbf{e}_2\mathbf{e}_2^T)(\mathbf{e}_1^T\mathbf{e}_1) - (\mathbf{e}_1^T\mathbf{e}_2)^2}
\end{aligned}$$

$$= \{\mathbf{e}_1\mathbf{e}_1^T(\mathbf{e}_2^T\mathbf{e}_2) + \mathbf{e}_2\mathbf{e}_2^T(\mathbf{e}_1^T\mathbf{e}_1) - \mathbf{e}_2\mathbf{e}_1^T(\mathbf{e}_1^T\mathbf{e}_2) - \mathbf{e}_1\mathbf{e}_2^T(\mathbf{e}_1^T\mathbf{e}_2)\}/\det(h)$$

$$= \frac{1}{\det(h)}(\mathbf{e}_1,\, \mathbf{e}_2)\begin{pmatrix} \mathbf{e}_2^T\mathbf{e}_2\mathbf{e}_1^T - \mathbf{e}_1^T\mathbf{e}_2\mathbf{e}_2^T \\ -\mathbf{e}_1^T\mathbf{e}_2\mathbf{e}_1^T - \mathbf{e}_1^T\mathbf{e}_1\mathbf{e}_2^T \end{pmatrix}$$

$$= (\mathbf{e}_1,\, \mathbf{e}_2)\frac{1}{\det(h)}\begin{pmatrix} \mathbf{e}_2^T\mathbf{e}_2 & -\mathbf{e}_1^T\mathbf{e}_2 \\ -\mathbf{e}_1^T\mathbf{e}_2 & \mathbf{e}_1^T\mathbf{e}_1 \end{pmatrix}\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix}$$

$$= (\mathbf{e}_1,\, \mathbf{e}_2)h^{-1}\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix},$$

where $h := \begin{pmatrix} \mathbf{e}_1^T\mathbf{e}_1 & \mathbf{e}_1^T\mathbf{e}_2 \\ \mathbf{e}_2^T\mathbf{e}_1 & \mathbf{e}_2^T\mathbf{e}_2 \end{pmatrix}$, $\det(h) = \det\begin{pmatrix} \mathbf{e}_1^T\mathbf{e}_1 & \mathbf{e}_1^T\mathbf{e}_2 \\ \mathbf{e}_2^T\mathbf{e}_1 & \mathbf{e}_2^T\mathbf{e}_2 \end{pmatrix} = (\mathbf{e}_1^T\mathbf{e}_1)(\mathbf{e}_2^T\mathbf{e}_2) - (\mathbf{e}_1^T\mathbf{e}_2)^2$, and $h^{-1} = \frac{1}{\det(h)}\begin{pmatrix} \mathbf{e}_2^T\mathbf{e}_2 & -\mathbf{e}_1^T\mathbf{e}_2 \\ -\mathbf{e}_1^T\mathbf{e}_2 & \mathbf{e}_1^T\mathbf{e}_1 \end{pmatrix}$.

On one hand, $P_b\psi = \mathbf{e}_1 a_1 + \mathbf{e}_2 a_2 = (\mathbf{e}_1,\, \mathbf{e}_2)\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right)$; on the other hand $P_b\psi = (\mathbf{e}_1,\, \mathbf{e}_2)h^{-1}\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix}\psi = (\mathbf{e}_1,\, \mathbf{e}_2)\frac{1}{\det(h)}\begin{pmatrix} \mathbf{e}_2^T\mathbf{e}_2 & -\mathbf{e}_1^T\mathbf{e}_2 \\ -\mathbf{e}_1^T\mathbf{e}_2 & \mathbf{e}_1^T\mathbf{e}_1) \end{pmatrix}\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix}\psi$, so we have

$$\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right) = h^{-1}\begin{pmatrix} \mathbf{e}_1^T\psi \\ \mathbf{e}_2^T\psi \end{pmatrix} = \frac{1}{\det(h)}\begin{pmatrix} \mathbf{e}_2^T\mathbf{e}_2 & -\mathbf{e}_1^T\mathbf{e}_2 \\ -\mathbf{e}_1^T\mathbf{e}_2 & \mathbf{e}_1^T\mathbf{e}_1 \end{pmatrix}\begin{pmatrix} \mathbf{e}_1^T\psi \\ \mathbf{e}_2^T\psi \end{pmatrix}.$$

Moreover

$$\begin{aligned} \frac{d}{dt}\psi(\gamma(t)) &= \frac{d}{dt}[\mathbf{e}_1(\gamma(t))a_1(\gamma(t)) + \mathbf{e}_2(\gamma(t))a_2(\gamma(t))] \\ &= \frac{d}{dt}\mathbf{e}_1(\gamma(t)) \cdot a_1(\gamma(t)) + \frac{d}{dt}\mathbf{e}_2(\gamma(t)) \cdot a_2(\gamma(t)) \\ &\quad + \mathbf{e}_1(\gamma(t))\frac{d}{dt}a_1(\gamma(t)) + \mathbf{e}_2(\gamma(t))\frac{d}{dt}a_2(\gamma(t)), \end{aligned}$$

$$\text{where } \frac{d}{dt}\mathbf{e}_j(\gamma(t)) = \frac{\partial \mathbf{e}_j}{\partial u}(\gamma(t)) \cdot u'(t) + \frac{\partial \mathbf{e}_j}{\partial v}(\gamma(t)) \cdot v'(t)$$

Finally we are ready to calculate the covariant derivative:

$$\begin{aligned} P_b\frac{d}{dt}\psi(\gamma(t)) &= (\mathbf{e}_1,\, \mathbf{e}_2)(h^{-1})\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix}\frac{d}{dt}\psi(\gamma(t)) \\ &= (\mathbf{e}_1,\, \mathbf{e}_2)(h^{-1})\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix}\{[\frac{\partial \mathbf{e}_1}{\partial u}u'(t) + \frac{\partial \mathbf{e}_1}{\partial v}v'(t)]a_1 \\ &\quad + [\frac{\partial \mathbf{e}_2}{\partial u}u'(t) + \frac{\partial \mathbf{e}_2}{\partial v}v'(t)]a_2\} + (\mathbf{e}_1,\, \mathbf{e}_2)\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} \\ &= (\mathbf{e}_1,\, \mathbf{e}_2)\{\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} \\ &\quad + (h^{-1})\begin{pmatrix} \mathbf{e}_1^T\frac{\partial \mathbf{e}_1}{\partial u}u'(t) + \mathbf{e}_1^T\frac{\partial \mathbf{e}_1}{\partial v}v'(t) & \mathbf{e}_1^T\frac{\partial \mathbf{e}_2}{\partial u}u'(t) + \mathbf{e}_1^T\frac{\partial \mathbf{e}_2}{\partial v}v'(t) \\ \mathbf{e}_2^T\frac{\partial \mathbf{e}_1}{\partial u}u'(t) + \mathbf{e}_2^T\frac{\partial \mathbf{e}_1}{\partial v}v'(t) & \mathbf{e}_2^T\frac{\partial \mathbf{e}_2}{\partial u}u'(t) + \mathbf{e}_2^T\frac{\partial \mathbf{e}_2}{\partial v}v'(t) \end{pmatrix}\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right)\} \\ &= (\mathbf{e}_1,\, \mathbf{e}_2)\{\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + (h^{-1})u'(t)\begin{pmatrix} \mathbf{e}_1^T\frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_1^T\frac{\partial \mathbf{e}_2}{\partial u} \\ \mathbf{e}_2^T\frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_2^T\frac{\partial \mathbf{e}_2}{\partial u} \end{pmatrix}\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right) \end{aligned}$$

$$+ (h^{-1}) v'(t) \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_1^T \frac{\partial \mathbf{e}_2}{\partial v} \\ \mathbf{e}_2^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_2^T \frac{\partial \mathbf{e}_2}{\partial v} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \}$$

$$= (\mathbf{e}_1, \mathbf{e}_2) \{ \begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + u'(t) \begin{pmatrix} \Gamma^1_{11} & \Gamma^1_{12} \\ \Gamma^2_{11} & \Gamma^2_{12} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + v'(t) \begin{pmatrix} \Gamma^1_{21} & \Gamma^1_{22} \\ \Gamma^2_{21} & \Gamma^2_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \}$$

$$= (\mathbf{e}_1, \mathbf{e}_2) \{ \begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + \omega(\gamma(t), \gamma'(t)) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \},$$

where we define the Christoffel symbols (see [**17**])

$$\begin{pmatrix} \Gamma^1_{11} & \Gamma^1_{12} \\ \Gamma^2_{11} & \Gamma^2_{12} \end{pmatrix} = (h^{-1}) \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_1^T \frac{\partial \mathbf{e}_2}{\partial u} \\ \mathbf{e}_2^T \frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_2^T \frac{\partial \mathbf{e}_2}{\partial u} \end{pmatrix}$$

$$\begin{pmatrix} \Gamma^1_{21} & \Gamma^1_{22} \\ \Gamma^2_{21} & \Gamma^2_{22} \end{pmatrix} = (h^{-1}) \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_1^T \frac{\partial \mathbf{e}_2}{\partial v} \\ \mathbf{e}_2^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_2^T \frac{\partial \mathbf{e}_2}{\partial v} \end{pmatrix}$$

$$\omega(\gamma(t), \gamma'(t)) = u'(t) \begin{pmatrix} \Gamma^1_{11} & \Gamma^1_{12} \\ \Gamma^2_{11} & \Gamma^2_{12} \end{pmatrix} + v'(t) \begin{pmatrix} \Gamma^1_{21} & \Gamma^1_{22} \\ \Gamma^2_{21} & \Gamma^2_{22} \end{pmatrix}$$

EXAMPLE. For the trivialization τ_0, $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ f_u(u,v) \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ f_v(u,v) \end{pmatrix}$. we calculate ω_0 as follows.

$$\begin{aligned} \det(h) &= (\mathbf{e}_1^T \mathbf{e}_1)(\mathbf{e}_2^T \mathbf{e}_2) - (\mathbf{e}_1^T \mathbf{e}_2)^2 \\ &= (1 + f_u^2)(1 + f_v^2) - f_u^2 f_v^2 \\ &= 1 + f_u^2 + f_v^2 \\ h^{-1} &= \frac{1}{\det(h)} \begin{pmatrix} \mathbf{e}_2^T \mathbf{e}_2 & -\mathbf{e}_1^T \mathbf{e}_2 \\ -\mathbf{e}_1^T \mathbf{e}_2 & \mathbf{e}_1^T \mathbf{e}_1 \end{pmatrix} = \frac{1}{1 + f_u^2 + f_v^2} \begin{pmatrix} 1+f_v^2 & -f_u f_v \\ -f_u f_v & 1+f_u^2 \end{pmatrix} \\ \begin{pmatrix} \Gamma^1_{11} & \Gamma^1_{12} \\ \Gamma^2_{11} & \Gamma^2_{12} \end{pmatrix} &= (h^{-1}) \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_1^T \frac{\partial \mathbf{e}_2}{\partial u} \\ \mathbf{e}_2^T \frac{\partial \mathbf{e}_1}{\partial u} & \mathbf{e}_2^T \frac{\partial \mathbf{e}_2}{\partial u} \end{pmatrix} \\ &= \frac{1}{1 + f_u^2 + f_v^2} \begin{pmatrix} 1+f_v^2 & -f_u f_v \\ -f_u f_v & 1+f_u^2 \end{pmatrix} \begin{pmatrix} f_u f_{uu} & f_u f_{uv} \\ f_v f_{uu} & f_v f_{uv} \end{pmatrix} \\ &= \frac{1}{1 + f_u^2 + f_v^2} \begin{pmatrix} f_u f_{uu} & f_u f_{uv} \\ f_v f_{uu} & f_v f_{uv} \end{pmatrix} \\ \begin{pmatrix} \Gamma^1_{21} & \Gamma^1_{22} \\ \Gamma^2_{21} & \Gamma^2_{22} \end{pmatrix} &= (h^{-1}) \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_1^T \frac{\partial \mathbf{e}_2}{\partial v} \\ \mathbf{e}_2^T \frac{\partial \mathbf{e}_1}{\partial v} & \mathbf{e}_2^T \frac{\partial \mathbf{e}_2}{\partial v} \end{pmatrix} \\ &= \frac{1}{1 + f_u^2 + f_v^2} \begin{pmatrix} 1+f_v^2 & -f_u f_v \\ -f_u f_v & 1+f_u^2 \end{pmatrix} \begin{pmatrix} f_u f_{uv} & f_u f_{vv} \\ f_v f_{uv} & f_v f_{vv} \end{pmatrix} \\ &= \frac{1}{1 + f_u^2 + f_v^2} \begin{pmatrix} f_u f_{uv} & f_u f_{vv} \\ f_v f_{uv} & f_v f_{vv} \end{pmatrix} \end{aligned}$$

$$\omega_0 = u'(t)\begin{pmatrix}\Gamma^1_{11} & \Gamma^1_{12}\\ \Gamma^2_{11} & \Gamma^2_{12}\end{pmatrix} + v'(t)\begin{pmatrix}\Gamma^1_{21} & \Gamma^1_{22}\\ \Gamma^2_{21} & \Gamma^2_{22}\end{pmatrix}$$
$$= \frac{u'(t)}{1+f_u^2+f_v^2}\begin{pmatrix}f_uf_{uu} & f_uf_{uv}\\ f_vf_{uu} & f_vf_{uv}\end{pmatrix} + \frac{v'(t)}{1+f_u^2+f_v^2}\begin{pmatrix}f_uf_{uv} & f_uf_{vv}\\ f_vf_{uv} & f_vf_{vv}\end{pmatrix}$$

Example. We are to calculate ω_1 and ω_2 which are corresponding to τ_1 and τ_2 respectively.

In τ_1, $\tilde{\mathbf{e}}_1 = \frac{1}{\sqrt{1+f_u^2}}\begin{pmatrix}1\\0\\f_u\end{pmatrix}, \tilde{\mathbf{e}}_2 = \frac{1}{\sqrt{(1+f_u^2)(1+f_u^2+f_v^2)}}\begin{pmatrix}-f_uf_v\\1+f_u^2\\f_v\end{pmatrix}$. Because $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2\}$ are orthonormal,

$$0 = \tilde{\mathbf{e}}_1^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial u} = \tilde{\mathbf{e}}_1^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial v} = \tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_2}{\partial u} = \tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_2}{\partial v},$$
$$h^{-1} = \frac{1}{\det h}\begin{pmatrix}\tilde{\mathbf{e}}_2^T\tilde{\mathbf{e}}_2 & -\tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_2\\ -\tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_2 & \tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_1\end{pmatrix} = I;$$

moreover, $\tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial u} = -\tilde{\mathbf{e}}_1^T\frac{\partial\tilde{\mathbf{e}}_2}{\partial u}$, and $\tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial v} = -\tilde{\mathbf{e}}_1^T\frac{\partial\tilde{\mathbf{e}}_2}{\partial v}$.

$$\tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial u} = \tilde{\mathbf{e}}_2^T[-(1+f_u^2)^{-3/2}f_uf_{uu}\begin{pmatrix}1\\0\\f_u\end{pmatrix} + \frac{1}{\sqrt{1+f_u^2}}\begin{pmatrix}0\\0\\f_{uu}\end{pmatrix}]$$
$$= \tilde{\mathbf{e}}_2^T[\begin{pmatrix}-f_uf_{uu}\\0\\-f_u^2f_{uu}\end{pmatrix} + \begin{pmatrix}0\\0\\(1+f_u^2)f_{uu}\end{pmatrix}]\frac{1}{(1+f_u^2)^{3/2}}$$
$$= \tilde{\mathbf{e}}_2^T\begin{pmatrix}-f_u\\0\\1\end{pmatrix}\frac{f_{uu}}{\sqrt{1+f_u^2}^3}$$
$$= \frac{f_vf_{uu}}{\sqrt{1+f_u^2+f_v^2}(1+f_u^2)}$$

Similarly $\tilde{\mathbf{e}}_2^T\frac{\partial\tilde{\mathbf{e}}_1}{\partial v} = \frac{f_vf_{uv}}{\sqrt{1+f_u^2+f_v^2}(1+f_u^2)}$.

Therefore

$$\omega_1(\gamma(0), \gamma'(0))$$
$$= u'\begin{pmatrix}0 & \frac{-f_{uu}f_v}{(1+f_u^2)(1+f_u^2+f_v^2)^{1/2}}\\ \frac{f_{uu}f_v}{(1+f_u^2)(1+f_u^2+f_v^2)^{1/2}} & 0\end{pmatrix} + v'\begin{pmatrix}0 & \frac{-f_{uv}f_v}{(1+f_u^2)(1+f_u^2+f_v^2)^{1/2}}\\ \frac{f_{uv}f_v}{(1+f_u^2)(1+f_u^2+f_v^2)^{1/2}} & 0\end{pmatrix}$$

In τ_2, $\tilde{\mathbf{e}}_1 = \frac{1}{\sqrt{1+f_v^2}}\begin{pmatrix}0\\1\\f_v\end{pmatrix}, \tilde{\mathbf{e}}_2 = \frac{1}{\sqrt{(1+f_v^2)(1+f_u^2+f_v^2)}}\begin{pmatrix}1+f_v^2\\-f_uf_v\\f_u\end{pmatrix}$. Because $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2\}$ are orthonormal,

$$0 = \tilde{\mathbf{e}}_1^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial u} = \tilde{\mathbf{e}}_1^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial v} = \tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_2}{\partial u} = \tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_2}{\partial v},$$

$$h^{-1} = \frac{1}{\det(h)}\begin{pmatrix}\tilde{\mathbf{e}}_2^T\tilde{\mathbf{e}}_2 & -\tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_2\\ -\tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_2 & \tilde{\mathbf{e}}_1^T\tilde{\mathbf{e}}_1\end{pmatrix} = I;$$

moreover, $\tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial u} = -\tilde{\mathbf{e}}_1^T \frac{\partial \tilde{\mathbf{e}}_2}{\partial u}$, and $\tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial v} = -\tilde{\mathbf{e}}_1^T \frac{\partial \tilde{\mathbf{e}}_2}{\partial v}$.

$$\begin{aligned}
\tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial u} &= \tilde{\mathbf{e}}_2^T[-(1+f_v^2)^{-3/2} f_v f_{uv}\begin{pmatrix}0\\1\\f_v\end{pmatrix} + \frac{1}{\sqrt{1+f_v^2}}\begin{pmatrix}0\\0\\f_{uv}\end{pmatrix}]\\
&= \tilde{\mathbf{e}}_2^T[\begin{pmatrix}0\\-f_vf_{uv}\\-f_v^2f_{uv}\end{pmatrix} + \begin{pmatrix}0\\0\\(1+f_v^2)f_{uv}\end{pmatrix}]\frac{1}{\sqrt{1+f_v^2}^3}\\
&= \tilde{\mathbf{e}}_2^T\begin{pmatrix}0\\-f_v\\1\end{pmatrix}\frac{f_{uv}}{(1+f_v^2)^{3/2}}\\
&= \frac{f_uf_{uv}}{\sqrt{1+f_u^2+f_v^2}(1+f_v^2)}.\\
\text{Similarly } \tilde{\mathbf{e}}_2^T \frac{\partial \tilde{\mathbf{e}}_1}{\partial v} &= \frac{f_uf_{vv}}{\sqrt{1+f_u^2+f_v^2}(1+f_v^2)}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\omega_2(\gamma(0), \gamma'(0))\\
&= u'\begin{pmatrix}0 & \frac{-f_{uv}f_u}{(1+f_v^2)(1+f_u^2+f_v^2)^{1/2}}\\ \frac{f_{uv}f_u}{(1+f_v^2)(1+f_u^2+f_v^2)^{1/2}} & 0\end{pmatrix} + v'\begin{pmatrix}0 & \frac{-f_{vv}f_u}{(1+f_v^2)(1+f_u^2+f_v^2)^{1/2}}\\ \frac{f_{vv}f_u}{(1+f_v^2)(1+f_u^2+f_v^2)^{1/2}} & 0\end{pmatrix}
\end{aligned}$$

Remark: We find that ω_1 and ω_2 are antisymmetric while ω_0 is not. ω_1 and ω_2 are in the Lie algebra $\mathfrak{o}(2)$, because the corresponding trivializations τ_1 and τ_2 are defined in such a way that they maintain the inner product from the standard fiber $\mathbb{R}^2$ to the total space $E_b = \mathbb{R}^3$ and that their structure group is the Lie group $O(2,\mathbb{R})$. While τ_0 does not have the properties that τ_1 and τ_2 have, and its structure group is $GL(2,\mathbb{R})$, and so ω_0 has values in the Lie algebra $gl(2)$.

6.3. Hermitian Vector Bundles, Connections in the Case of H_3 System

6.3.1. Hermitian Vector Bundles and Structure Group in the Case of H_3 System. For the H_3 system, and given a point $b = (l_{12}^2, l_{13}^2, l_{23}^2)$ inside the cone $\mathcal{C}$, defined at §5.2, let $(\mathbf{R}_1(b), \mathbf{R}_2(b), \mathbf{R}_3(b))$ be the standard configuration of the conformation b. We have the the electronic Hamiltonian for the H_3 system

$$\begin{aligned}\tilde{H}(b) = &-\frac{\hbar^2}{2m}(\triangle_1 + \triangle_2 + \triangle_3)\\ &-\sum_{j=1}^{3}\sum_{k=1}^{3}\frac{Ke^2}{\|\mathbf{r}_j - \mathbf{R}_k(b)\|}\\ &+\sum_{j=1}^{3}\sum_{k=j+1}^{3}\frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|},\end{aligned}$$

where $\mathbf{R}_i(b)$ is the position vector of the ith nucleus in the standard configuration.

Define $\tilde{E}(b) := \min_{\psi\in\mathcal{H},\psi\neq 0}\frac{(\psi,\widetilde{H}(b)\psi)}{(\psi,\psi)}$, which is the electronic ground state energy for the H_3 system. We define

$\mathcal{H} := \mathcal{H}_e \wedge \mathcal{H}_e \wedge \mathcal{H}_e,$ where $\mathcal{H}_e = L^2(\mathbb{R}^3, \mathbb{C}) \otimes \mathbb{C}^2$ is the Hilbert space of a single electron;

$E_b := \{\psi \in \mathcal{H} | \widetilde{H}(b)\psi = \widetilde{E}(b)\psi, \hat{S}_3\psi = \frac{\hbar}{2}\psi\}$

E_b is a $\mathbb{C}$-vector subspace of $\mathcal{H}$ and inherits the inner product of $\mathcal{H}$;

($\hat{S}_3$ is the spin operator defined in §3.1.3.)

$\mathcal{C}_k := \{b \in \mathcal{C} | \dim E_b = k\}$, which is the base space; (how choose k such that $\mathcal{C}_k$ is open in $\mathcal{C}$ is an unsolved issue)

$\mathbb{C}^k$ is the standard fiber;

$E = \{(b, \psi) \in \mathcal{C}_k \times \mathcal{H} | \psi \in E_b\}$, which is the total space;

$\pi : E \to \mathcal{C}_k : (b, \psi) \mapsto b$, which is a smooth mapping;

$\mathfrak{U}(k) := \{A \in \mathbb{C}^{k\times k} | A^\dagger A = I\}$, which is the structure group;

$\tau_i : U_i \times \mathbb{C}^k \to \pi^{-1}(U_i), i \in \mathcal{I}$, which is an atlas of local trivializations,

where U_i is open in $\mathcal{C}_k$. Also we want τ_i to be smooth in the sense that $\forall \tilde{\psi} \in \mathcal{H}$, the map $U_i \times \mathbb{C}^k \to \mathbb{C} : (b, \mathbf{y}) \mapsto (\tilde{\psi}, \tau_i(b, \mathbf{y}))_{\mathcal{H}}$ is smooth with respect to $\frac{d}{dQ}$, $\frac{d}{dS}, \frac{d}{d\theta}, \frac{d}{d\mathbf{y}}$(see §5.2). The existence of of such $\{\tau_i\}_{i\in\mathcal{I}}$, where $\{U_i\}_{i\in\mathcal{I}}$ is an open covering of $\mathcal{C}_k$ and for all $b \in \mathcal{C}_k$ the mapping $y \mapsto \tau_i(b, \mathbf{y}) : \mathbb{C}^k \to E_b$ is an inner product linear isomorphism, is an unsolved issue;

Therefore $(E, \mathcal{C}_k, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$ is a Hermitian vector bundle for the non-collinear H_3 system.

6.3.2. Representation of the Symmetry Group in the Fiber. If b is a scalene triangle, an isosceles triangle or an equilateral triangle, then let $G(b) = \{A \in O(3) | \{A\mathbf{R}_1(b), A\mathbf{R}_2(b), A\mathbf{R}_3(b)\} = \{\mathbf{R}_1(b), \mathbf{R}_2(b), \mathbf{R}_3(b)\}\}$ be the symmetry group isomorphic to C_s, C_{2v} or D_{3h} respectively. If $A \in G(b)$ and $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \mapsto \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$ is in E_b, where $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \mathbb{R}^3$ are the electron position vectors and $s_1, s_2, s_3 \in \{0, 1\}$ are the spin variables, then A induces a mapping $l_A : E_b \to E_b : \psi \mapsto ((\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \mapsto \psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3))$ i.e.

$$(l_A\psi)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = \psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3).$$

We need to show that if $\psi \in E_b$ then $l_A\psi \in E_b$, i.e. $\tilde{H}(b)l_A\psi = \tilde{E}(b)l_A\psi$ and $\hat{S}_3 l_A\psi = \frac{\hbar}{2} l_A\psi$.

(1) To show $\tilde{H}(b)l_A\psi = \tilde{E}(b)l_A\psi$, we apply l_A to the both sides of $\tilde{H}(b))\psi = \tilde{E}(b)\psi$ attaining $l_A\tilde{H}(b))\psi = l_A\tilde{E}(b)\psi$. Because R.H.S= $l_A\tilde{E}(b)\psi = \tilde{E}(b)l_A\psi$, we want to prove that the L.H.S.=$l_A\tilde{H}(b))\psi = \tilde{H}(b))l_A\psi$.

(1.1) We need to prove that

$$[(l_A \circ \triangle_i)\psi](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = [(\triangle_i \circ l_A)\psi](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3), i = 1, 2, 3.$$

By Taylor's theorem, $(l_A\psi)(\mathbf{x}+\mathbf{h}) = \psi(A^T(\mathbf{x}+\mathbf{h})) = \psi(A^T\mathbf{x}+A^T\mathbf{h}) = \psi(A^T\mathbf{x}) + \nabla\psi(A^T\mathbf{x}) \cdot (A^T\mathbf{h}) + o(\|\mathbf{h}\|) = \psi(A^T\mathbf{x}) + [A(\nabla\psi)(A^T\mathbf{x})] \cdot \mathbf{h} + o(\|\mathbf{h}\|)$, so we have

$$\nabla(l_A\psi)(\mathbf{x}) = A(\nabla\psi)(A^T\mathbf{x}) = A(l_A\nabla\psi)(\mathbf{x}).$$

We will use the Green's identity: $\int_\Omega \triangle(\psi)\phi\, d\mathbf{x} + \int_\Omega \nabla\psi \cdot \nabla\phi\, d\mathbf{x} = \int_{\partial\Omega} \phi\frac{\partial\psi}{\partial\nu}ds$, where ψ and ϕ are two smooth functions over a compact region $\Omega \subset \mathbb{R}^3$ with smooth boundary $\partial\Omega$, and where $\frac{\partial\psi}{\partial\nu}$ is the directional derivative of ψ along the outward normal to the boundary of $\partial\Omega$. In our case, we assume ψ to be a smooth function over $\mathbb{R}^3$ instead of in E_b, and we will come back to consider the case $\psi \in E_b$ later. We let ϕ be any smooth function on $\mathbb{R}^3$ with compact support, then we can always find a closed ball Ω centered at the origin enclosing the support of ϕ and such that $\phi(\mathbf{x})|_{\mathbf{x}\in\partial\Omega} = 0$. Therefore $\int_{\partial\Omega} \phi\frac{\partial\psi}{\partial\nu}ds = 0$ and hence $\int_\Omega \triangle(\psi)\phi\, d\mathbf{x} = -\int_\Omega \nabla\psi \cdot \nabla\phi\, d\mathbf{x}$.

By Green's identity and Taylor's theroem, $\int_\Omega \triangle(l_A\psi)\phi\, d\mathbf{x} = -\int_\Omega \nabla(l_A\psi)\cdot\nabla\phi\, d\mathbf{x} = -\int_\Omega Al_A(\nabla\psi) \cdot \nabla\phi\, d\mathbf{x}$.

Moreover,

$$\begin{aligned}
\int_\Omega l_A(\triangle\psi)\phi\, d\mathbf{x} &= \int_\Omega (\triangle\psi)(A^T\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x} \\
&= \int_\Omega (\triangle\psi)(\mathbf{x}')\phi(A\mathbf{x}')\, d\mathbf{x}' \\
&\quad (\text{ by changing variables } \mathbf{x}' = A^T\mathbf{x}, \mathbf{x} = A\mathbf{x}'; \\
&\quad \text{since } |\det A| = 1, d(A\mathbf{x}') = |\det A|\, d\mathbf{x}' = d\mathbf{x}'; A\Omega = \Omega) \\
&= -\int_\Omega \nabla\psi \cdot \nabla(l_{A^T}\phi)\, d\mathbf{x}' \\
&= -\int_\Omega \nabla\psi \cdot (A^T l_{A^T}\nabla\phi)\, d\mathbf{x}' \\
&= -\int_\Omega (A\nabla\psi) \cdot (l_{A^T}\nabla\phi)\, d\mathbf{x}'
\end{aligned}$$

$$= -\int_\Omega (A\nabla\psi)(\mathbf{x}') \cdot \nabla\phi(A\mathbf{x}')\, d\mathbf{x}'$$

$$= -\int_\Omega A l_A(\nabla\psi) \cdot \nabla\phi\, d\mathbf{x}$$

(by changing variables $\mathbf{x}' = A^T\mathbf{x}, \mathbf{x} = A\mathbf{x}'; |\text{det}A^T| = 1$)

It follows $\int_\Omega \triangle(l_A\psi)\phi\, d\mathbf{x} = \int_\Omega (l_A\triangle\psi)\phi\, d\mathbf{x}$ for any $\phi \in \mathbb{R}$ with compact support $\Rightarrow \triangle \circ l_A = l_A \circ \triangle$. Because $\triangle_i$ acts on $\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \in E_b$ in a partial derivative way (variables $\mathbf{r}_j, j \neq i$ and s_1, s_2, s_3 are fixed as constants), the result we derived above also holds for $\psi \in E_b$: $[(-\frac{\hbar^2}{2m}(\triangle_1+\triangle_2+\triangle_3)\circ l_A)\psi](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = [(l_A \circ (-\frac{\hbar^2}{2m}(\triangle_1 + \triangle_2 + \triangle_3)))\psi](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$.

(1.2) We need to prove that

$$l_A \circ \sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|} = \sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|} \circ l_A.$$

To see

$$l_A(\sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|}\psi)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$$

$$= \sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|A^T\mathbf{r}_j - A^T\mathbf{r}_k\|}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

$$= \sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

$$= \sum_{j=1}^{2}\sum_{k=j+1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{r}_k\|}(l_A\psi)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \text{ as desired.}$$

(1.3) We need to prove that

$$l_A \circ (-\sum_{j=1}^{3}\sum_{k=1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{R}_k\|}) = (-\sum_{j=1}^{3}\sum_{k=1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{R}_k\|}) \circ l_A.$$

To see

$$l_A(-\sum_{j=1}^{3}\sum_{k=1}^{3} \frac{Ke^2}{\|\mathbf{r}_j - \mathbf{R}_k\|}\psi)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$$

$$= -\sum_{j=1}^{3}\sum_{k=1}^{3}\frac{Ke^2}{\|A^T\mathbf{r}_j - \mathbf{R}_k\|}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

$$= -\sum_{j=1}^{3}\sum_{k=1}^{3}\frac{Ke^2}{\|A^T\mathbf{r}_j - A^T\mathbf{R}_k\|}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

(because A^T permutes $(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3)$)

$$= -\sum_{j=1}^{3}\sum_{k=1}^{3}\frac{Ke^2}{\|\mathbf{r}_j - \mathbf{R}_k\|}l_A\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \text{ as desired.}$$

In summary, we have $l_A \circ \tilde{H} = \tilde{H} \circ l_A$ and hence $[\tilde{H}(b)l_A\psi](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = \tilde{E}(b)l_A\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$.

(2) To show $\hat{S}_3 l_A\psi = \frac{\hbar}{2}l_A\psi$, we apply l_A to the both sides of $\hat{S}_3\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = \frac{\hbar}{2}\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$. Because $\hat{S}_3$ acts only on the s_1, s_2, s_3 variables, we get

$$\begin{aligned}
L.H.S = [l_A(\hat{S}_3\psi)](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) = [\hat{S}_3(l_A\psi)](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)& \\
=\hat{S}_3\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)& \\
R.H.S. = l_A\frac{\hbar}{2}\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) =&\frac{\hbar}{2}(l_A\psi)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) \\
=&\frac{\hbar}{2}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3) \\
\Rightarrow \hat{S}_3\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3) =&\frac{\hbar}{2}\psi(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)
\end{aligned}$$

Finally, we have $\tilde{H}(b)l_A\psi = \tilde{E}(b)l_A\psi$ and $\hat{S}_3 l_A\psi = \frac{\hbar}{2}l_A\psi$ and hence the following statement holds: if $\psi \in E_b$ then $l_A\psi \in E_b$. □

Moreover, we claim that l_A is linear. To see, if $\psi_1, \psi_2 \in E_b, \alpha \in \mathbb{C}$, then we have

$$[l_A(\psi_1 + \alpha\psi_2)](\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3)$$

$$= (\psi_1 + \alpha\psi_2)(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

$$= \psi_1(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3) + \alpha\psi_2(A^T\mathbf{r}_1, A^T\mathbf{r}_2, A^T\mathbf{r}_3; s_1, s_2, s_3)$$

$$= (l_A\psi_1)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3) + \alpha(l_A\psi_2)(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; s_1, s_2, s_3). \quad \square$$

If $A, B \in G(b)$, then $l_{AB} = l_A \circ l_B$. So we get a group homomorphism $l : G(b) \to GL(E_b)$. Moreover, by the theory developed at §6, we have $l \cong l_1 \oplus \cdots \oplus l_k$, where $l_1, \cdots, l_k$ are irreducible representations. If $k > 1$, the decomposition might change with the base point b. Since k is unknown to us, this decomposition remains an unresolved issue.

6.3.3. A Natural Connection in Certain Hermitian Vector Bundles. Let $(E, B, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$ be a Hermitian vector bundle with standard fiber $\mathbb{C}^k$ and structure group $\mathfrak{U}(k)$. Under certain conditions, there is a natural connection on this Hermitian vector bundle, as studied, e.g., by Bott and Chern[**16**].

Assume that base space B is an n-dimensional manifold. Assume that for all $b \in B$, $E_b = \pi^{-1}(\{b\})$ a subspace of V, where V is a fixed inner product space. Suppose $\{\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_k\}$ is the standard basis of $\mathbb{C}^k$. Assume that $\forall i \in \mathcal{I}, \forall b \in U_i$ the trivialization τ_i determines the basis $\{\mathbf{e}_1(b), \cdots, \mathbf{e}_k(b)\} := \{\tau_i(b, \hat{\mathbf{e}}_1), \cdots, \tau_i(b, \hat{\mathbf{e}}_k)\}$ of $E_b = \pi^{-1}(\{b\})$ which is orthonormal in the inner product of V. To simplify the notation, here we denote $\tau_i(b, \hat{\mathbf{e}}_j)$ as $\mathbf{e}_j^{(i)}(b)$ or just $\mathbf{e}_j$ if the base point $b \in U_i$ and the local trivialization is understood, for $j = 1, \cdots, k$.

Suppose $\tau : U \times \mathbb{C}^k \to \pi^{-1}(U)$ is a smooth local trivialization of E such that $\{\tau_i\}_{i\in\mathcal{I}} \cup \{\tau\}$ has a $GL(k)-$valued cocycle. We have that $\tau(b, \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}) = (b, \mathbf{e}_1 u_1 + \cdots + \mathbf{e}_k u_k)$, where $\mathbf{e}_j = \tau(b, \hat{\mathbf{e}}_j)$. Let $\psi : B \to E$ be an arbitrary vector field satisfying $\psi(b) \in E_b, \forall b \in B$ and $\psi(b) = \mathbf{e}_1 a_1(b) + \cdots + \mathbf{e}_k a_k(b)$. A smooth parametrized curve at b is $\gamma : (-\epsilon, \epsilon) \to B$ such that $b = \gamma(0)$.

Lemma. The orthogonal projection operator from the ambient inner product space V into the $k-$dimensional subspace E_b is $P_b = (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix}$, where $h := \begin{pmatrix} (\mathbf{e}_1, \mathbf{e}_1) & \cdots & (\mathbf{e}_1, \mathbf{e}_k) \\ \vdots & \vdots & \vdots \\ (\mathbf{e}_k, \mathbf{e}_1) & \cdots & (\mathbf{e}_k, \mathbf{e}_k) \end{pmatrix}$. (See the definition of $\mathbf{e}_i^\dagger$ at §3.4.)

Proof. It is enough to show that P_b is self adjoint and $P_b^2 = P_b$. To see P_b is self adjoint:

$$\begin{aligned} P_b^\dagger &= \{(\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \}^\dagger \\ &= (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1\dagger} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \\ &= (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \\ &= P_b. \end{aligned}$$

To see P_b is a projection operator:

$$\begin{aligned} P_b^2 &= (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \\ &= (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} h h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \\ &= (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \\ &= P_b. \end{aligned}$$

□

We define the covariant derivative of ψ in the direction $\gamma'(t)$ as the orthogonal projection of $\psi'(t) \in V$ into $\pi^{-1}(\{\gamma(t)\})$. Suppose $x_1, \cdots, x_n$ are local coordinate on B near $b \in B$, $\gamma(t) = (x_1(t), \cdots, x_n(t))$. On one hand, $P_b\psi = (\mathbf{e}_1, \cdots, \mathbf{e}_k) \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$; on the other hand $P_b\psi = (\mathbf{e}_1, \cdots, \mathbf{e}_k) \cdot h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \psi$, so we have $\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \psi \\ \vdots \\ \mathbf{e}_k^\dagger \psi \end{pmatrix}$.

Moreover

$$\frac{d}{dt}\psi(\gamma(t)) = \frac{d}{dt}[\mathbf{e}_1(\gamma(t))a_1(\gamma(t)) + \cdots + \mathbf{e}_k(\gamma(t))a_k(\gamma(t))]$$

$$= \frac{d}{dt}\mathbf{e}_1(\gamma(t)) \cdot a_1(\gamma(t)) + \mathbf{e}_1(\gamma(t))\frac{d}{dt}a_1(\gamma(t)) + \cdots$$

$$+ \frac{d}{dt}\mathbf{e}_k(\gamma(t)) \cdot a_k(\gamma(t)) + \mathbf{e}_k(\gamma(t))\frac{d}{dt}a_k(\gamma(t)),$$

where $\frac{d}{dt}\mathbf{e}_j(\gamma(t)) = \frac{\partial \mathbf{e}_j}{\partial x_1}(\gamma(t)) \cdot x_1'(t) + \cdots + \frac{\partial \mathbf{e}_j}{\partial x_n}(\gamma(t)) \cdot x_n'(t),\ j = 1, \cdots, k$

$$P_b \frac{d}{dt}\psi(\gamma(t)) = (\mathbf{e}_1, \cdots, \mathbf{e}_k)(h^{-1}) \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \frac{d}{dt}\psi(\gamma(t))$$

$$= (\mathbf{e}_1, \cdots, \mathbf{e}_k)(h^{-1}) \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} \{(\frac{\partial \mathbf{e}_1}{\partial x_1}x_1' + \cdots + \frac{\partial \mathbf{e}_1}{\partial x_n}x_n')a_1 + \cdots$$

$$+ (\frac{\partial \mathbf{e}_k}{\partial x_1}x_1' + \cdots + \frac{\partial \mathbf{e}_k}{\partial x_n}\mathbf{x}_k')a_k\} + (\mathbf{e}_1, \cdots, \mathbf{e}_k) \begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix}$$

$$= (\mathbf{e}_1, \cdots, \mathbf{e}_k)\{\begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix} + (h^{-1})\cdot$$

$$\begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial x_1}x_1' + \cdots + \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial x_n}x_n' & \cdots & \mathbf{e}_1^T \frac{\partial \mathbf{e}_k}{\partial x_1}x_1' + \cdots + \mathbf{e}_1^T \frac{\partial \mathbf{e}_k}{\partial x_n}x_n' \\ \vdots & \vdots & \vdots \\ \mathbf{e}_k^T \frac{\partial \mathbf{e}_1}{\partial x_1}x_1' + \cdots + \mathbf{e}_k^T \frac{\partial \mathbf{e}_1}{\partial x_n}x_n' & \cdots & \mathbf{e}_k^T \frac{\partial \mathbf{e}_k}{\partial x_1}x_1' + \cdots + \mathbf{e}_k^T \frac{\partial \mathbf{e}_k}{\partial x_n}x_n' \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}\}$$

$$= (\mathbf{e}_1, \cdots, \mathbf{e}_k)\{\begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix} + (h^{-1})x_1' \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial x_1} & \cdots & \mathbf{e}_1^T \frac{\partial \mathbf{e}_k}{\partial x_1} \\ \vdots & \vdots & \vdots \\ \mathbf{e}_k^T \frac{\partial \mathbf{e}_1}{\partial x_1} & \cdots & \mathbf{e}_k^T \frac{\partial \mathbf{e}_k}{\partial x_1} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$$

$$+ \cdots + (h^{-1})x_n' \begin{pmatrix} \mathbf{e}_1^T \frac{\partial \mathbf{e}_1}{\partial x_n} & \cdots & \mathbf{e}_1^T \frac{\partial \mathbf{e}_k}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \mathbf{e}_k^T \frac{\partial \mathbf{e}_1}{\partial x_n} & \cdots & \mathbf{e}_k^T \frac{\partial \mathbf{e}_k}{\partial x_n} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}\}$$

$$= (\mathbf{e}_1, \cdots, \mathbf{e}_k)\{\begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix} + x_1' \begin{pmatrix} \Gamma_{11}^1 & \cdots & \Gamma_{1k}^1 \\ \vdots & \vdots & \vdots \\ \Gamma_{11}^k & \cdots & \Gamma_{1k}^k \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$$

$$+ \cdots + x_n' \begin{pmatrix} \Gamma_{n1}^1 & \cdots & \Gamma_{nk}^1 \\ \vdots & \vdots & \vdots \\ \Gamma_{n1}^k & \cdots & \Gamma_{nk}^k \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}\}$$

$$= (\mathbf{e}_1, \cdots, \mathbf{e}_k)\{\begin{pmatrix} a_1' \\ \vdots \\ a_k' \end{pmatrix} + \omega_i(\gamma(t), \gamma'(t)) \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}\}$$

where we define the Christoffel symbols

$$\begin{pmatrix} \Gamma^1_{i1} & \cdots & \Gamma^1_{ik} \\ \vdots & \vdots & \vdots \\ \Gamma^k_{i1} & \cdots & \Gamma^k_{ik} \end{pmatrix} = (h^{-1}) \begin{pmatrix} (\mathbf{e}_1, \frac{\partial \mathbf{e}_1}{\partial \mathbf{x}_i}) & \cdots & (\mathbf{e}_1, \frac{\partial \mathbf{e}_k}{\partial \mathbf{x}_i}) \\ \vdots & \vdots & \vdots \\ (\mathbf{e}_k, \frac{\partial \mathbf{e}_1}{\partial \mathbf{x}_i}) & \cdots & (\mathbf{e}_k, \frac{\partial \mathbf{e}_k}{\partial \mathbf{x}_i}) \end{pmatrix}$$

$$\omega(\gamma(t), \gamma'(t)) = x_1'(t) \begin{pmatrix} \Gamma^1_{11} & \cdots & \Gamma^1_{1k} \\ \vdots & \vdots & \vdots \\ \Gamma^k_{11} & \cdots & \Gamma^k_{1k} \end{pmatrix} + \cdots + x_n'(t) \begin{pmatrix} \Gamma^1_{n1} & \cdots & \Gamma^1_{nk} \\ \vdots & \vdots & \vdots \\ \Gamma^k_{n1} & \cdots & \Gamma^k_{nk} \end{pmatrix}$$

A more readable formula for ω is

$$\omega(\gamma(t), \gamma'(t)) = h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \frac{d\mathbf{e}_1}{dt} & \cdots & \mathbf{e}_1^\dagger \frac{d\mathbf{e}_k}{dt} \\ \vdots & \vdots & \vdots \\ \mathbf{e}_k^\dagger \frac{d\mathbf{e}_1}{dt} & \cdots & \mathbf{e}_k^\dagger \frac{d\mathbf{e}_k}{dt} \end{pmatrix}$$

If $\tau = \tau_j$ for some $j \in \mathcal{I}$ then $h = I_{k\times k}$ and the above formula for $\omega = \omega_j$ simplifies. In that case it is clear that $\omega_j(\gamma(t), \gamma'(t)) \in \mathfrak{u}(k)$.

6.3.4. Berry-Simon Connection in H_3. Like in the previous section, now we define a natural connection on $(E, \mathcal{C}_k, \pi, \mathbb{C}^k, \mathfrak{U}(k), \{\tau_i\}_{i\in\mathcal{I}})$, the Hermitian vector bundle for the non-collinear H_3 system. This connection is called the *Berry-Simon connection.*

To see in more details, we have an ambient inner product space $\mathcal{H} = \mathcal{H}_e \wedge \mathcal{H}_e \wedge \mathcal{H}_e$ such that $\forall b \in U_i$ we have $E_b = \pi^{-1}(\{b\}) \subset \mathcal{H}$. Suppose $\gamma : (-\epsilon, \epsilon) \to B$ is a smooth curve such that $\gamma(0) = b \in U_i$. $\psi : (-\epsilon, \epsilon) \to E$ a smooth vector field over γ i.e. $t \mapsto \psi(t) \in E_{\gamma(t)} = \pi^{-1}(\{\gamma(t)\})$ is smooth. We define the covariant derivative of ψ at the direction $\gamma'(t)$ as the orthogonal projection of $\psi'(t) \in V$ into $\pi^{-1}(\{\gamma(t)\})$.

Suppose $(\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_k)$ is the standard basis of $\mathbb{C}^k$. Then $\forall i \in \mathcal{I}, \forall b \in U_i$ the trivialization τ_i determines an orthonormal set in the inner product space $\mathcal{H}$: $\{\mathbf{e}_1(b), \cdots, \mathbf{e}_k(b)\} := \{\tau_i(x, \hat{\mathbf{e}}_1), \cdots, \tau_i(x, \hat{\mathbf{e}}_k)\}$. Then $\{\mathbf{e}_1(\gamma(t)), \cdots, \mathbf{e}_k(\gamma(t))\}$ is an orthonormal basis of $\pi^{-1}(\{\gamma(t)\})$.

The orthogonal projection operator from the ambient inner product space $\mathcal{H}$ into the $k-$dimensional subspace E_b is $P_b = (\mathbf{e}_1, \cdots, \mathbf{e}_k) h^{-1} \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix} = (\mathbf{e}_1, \cdots, \mathbf{e}_k) \begin{pmatrix} \mathbf{e}_1^\dagger \\ \vdots \\ \mathbf{e}_k^\dagger \end{pmatrix}$,

because $h := \begin{pmatrix} (\mathbf{e}_1, \mathbf{e}_1) & \cdots & (\mathbf{e}_1, \mathbf{e}_k) \\ \vdots & \vdots & \vdots \\ (\mathbf{e}_k, \mathbf{e}_1) & \cdots & (\mathbf{e}_k, \mathbf{e}_k) \end{pmatrix} = I$.

Then the covariant derivative of ψ is

$$P_b\frac{d}{dt}\psi(\gamma(t)) = (\mathbf{e}_1,\cdots,\mathbf{e}_k)\{\frac{d}{dt}\begin{pmatrix}(\mathbf{e}_1(\gamma(t)),\psi(t))\\ \vdots \\ (\mathbf{e}_k(\gamma(t)),\psi(t))\end{pmatrix} + \omega_i(\gamma(t),\gamma'(t))\begin{pmatrix}(\mathbf{e}_1(\gamma(t)),\psi(t))\\ \vdots \\ (\mathbf{e}_k(\gamma(t)),\psi(t))\end{pmatrix}\},$$

where we define

$$\begin{pmatrix}\Gamma^1_{i1} & \cdots & \Gamma^1_{ik}\\ \vdots & \vdots & \vdots \\ \Gamma^k_{i1} & \cdots & \Gamma^k_{ik}\end{pmatrix} = \begin{pmatrix}(\mathbf{e}_1,\frac{\partial\mathbf{e}_1}{\partial\mathbf{x}_i}) & \cdots & (\mathbf{e}_1,\frac{\partial\mathbf{e}_k}{\partial\mathbf{x}_i})\\ \vdots & \vdots & \vdots \\ (\mathbf{e}_k,\frac{\partial\mathbf{e}_1}{\partial\mathbf{x}_i}) & \cdots & (\mathbf{e}_k,\frac{\partial\mathbf{e}_k}{\partial\mathbf{x}_i})\end{pmatrix}$$

$$\omega_i(\gamma(t),\gamma'(t)) = x_1'\begin{pmatrix}\Gamma^1_{11} & \cdots & \Gamma^1_{1k}\\ \vdots & \vdots & \vdots \\ \Gamma^k_{11} & \cdots & \Gamma^k_{1k}\end{pmatrix} + \cdots + x_n'\begin{pmatrix}\Gamma^1_{n1} & \cdots & \Gamma^1_{nk}\\ \vdots & \vdots & \vdots \\ \Gamma^k_{n1} & \cdots & \Gamma^k_{nk}\end{pmatrix}$$

Then the family $\{(U_i,\tau_i,\omega_i)\}$ is the Berry-Simon connection defined on the Hermitian vector bundle of the non-collinear H_3 system. Generally we do not have an explicit formula for the trivialization τ_i, so we cannot give an explicit formula to calculate the Berry-Simon connection.

6.3.5. Remark(I): potential energy surfaces of H_3 at the center of the cone. The potential energy of the H_3 system is determined by the conformation, so we can define the potential energy function over the cone $\mathcal{C}$ which represents all the conformations of the system. Numerical evidence [**41**][**2**][**15**][**45**] shows that the potential energy function has a singularity at the central axis of the cone. Moreover, the dimension of the fiber will jump when the conformation changes from the non-equilateral to the equilateral. As a consequence, the Berry-Simon connection appears to have a singularity at the central axis of the cone and its curvature would not exist there. We will continue our discussion at §7.3 after we introduce the concept of holonomy and Berry phase.

Chapter 7

Holonomy Groups and Berry Phase

In this chapter holonomy groups and Berry phase are considered. Suppose a quantum system undergoes an evolution so that after some time it comes back to its original state. Such an evolution traces out a cycle in shape space. The result of the evolution will be reflected in the phase of the wave function in the form of a geometric phase factor, usually called *Berry phase*. This phase factor can be measured by interfering the initial and the final states. Such geometric phase factors only depend on the loop in the shape space; while they are independent of parameterization of the path in the shape space, and therefore of the speed at which the system moves along the cyclic path. We explain the Berry phase of the H_3 system using the geometric language i.e. in terms of holonomy in a Hermitian vector bundle.

7.1. Holonomy Groups

In this section we first define holonomy group at a global view point and then go to a local one.

Let $(E, B, \pi, F, G, \{\tau_i\}_{i\in\mathcal{I}})$ be a fiber bundle with standard fiber F and structure group G. $\{\tau_i\}_{i\in\mathcal{I}}$ is a atlas of smooth local trivializations with the smooth cocycle $\{g_{ij}\}_{(i,j)\in\mathcal{I}^2}$, i.e. $\{g_{ij} : U_i \cap U_j \to G\}_{(i,j)\in\mathcal{I}^2}$ is a family of smooth maps such that $\forall b \in U_i \cap U_j, \forall y \in F, (\tau_i^{-1} \circ \tau_j)(b, y) = (b, g_{ij}(b) \cdot y)$, where $g \cdot y$ is the left action of $g \in G$ on $y \in F$. Let $b \in B, E_b = \pi^{-1}(\{b\})$. Let $\gamma : [c_1, c_2] \to B$ be a smooth loop such that $\gamma(c_1) = \gamma(c_2) = b$. Given $z \in E_b$, for each $i \in \mathcal{I}$ s.t. $b \in U_i$, there is a $y_i \in F$

s.t. $\tau_i(b, y_i) = z$. In the other words, for $i, j \in \mathcal{I}, i \neq j$, $z = \tau_i(b, y_i) = \tau_j(b, y_j) \Leftrightarrow (b, g_{ji}(b) \cdot y_i) = (\tau_j^{-1} \circ \tau_i)(b, y_i) = (b, y_j) \Leftrightarrow y_j = g_{ji}(b) \cdot y_i$. The parallel translate of y_i along γ is $h_{\gamma,i} y_i \in F$ (Refer to §6.1 for detailed definition).

So the parallel translation along γ gives a well-defined mapping $h_\gamma : E_b \to E_b : z = \tau_i(b, y_i) \mapsto \tau_i(b, h_{\gamma,i} y_i)$ which is independent of i s.t. $b \in U_i$. So we can define the global holonomy group as follows.

Definition. The *global holonomy group* at b is

$$\begin{aligned} \text{holonomy}(b) :=& \{h_\gamma : E_b \to E_b | h_\gamma(\tau_i(b, y_i)) = \tau_i(b, h_{\gamma,i} y_i) \text{ for all } y_i \in F, \\ & \text{for all } i \in \mathcal{I} \text{ s.t. } b \in U_i\} \end{aligned}$$

This set of mappings is a group under composition.

Now we consider the holonomy transformations from the local view point: Let $h_{\gamma,i}$ be the parallel translation along the loop $\gamma(t), t \in [0, 1]$, where $\gamma(0) = \gamma(1) \in U_i$ of local trivialization τ_i (see §6.1 for details definition). Here $h_{\gamma,i}$ is not necessary independent of i.

Definition. The *local holonomy group* at b is

$$\text{holonomy}(b, i) := \{h_{\gamma,i} | \gamma \text{ is a smooth path in } B \text{ from } b \text{ to itself}, b \in U_i\}.$$

It is easy to see that holonomy(b, i) is a subgroup of G. Also holonomy(b, i) has the property that for all $i, j \in I_b$,

$$\text{holonomy}(b, j) = g_{ij}(b)^{-1} \text{holonomy}(b, i) g_{ij}(b)$$

To see, $\tau_i(b, h_{\gamma,i} y_i)$ is independent of i, where $b \in U_i$. Thus for $i, j \in \mathcal{I}, i \neq j, b \in U_i \cap U_j$,

$$\begin{aligned} & \tau_i(b, h_{\gamma,i} y_i) = \tau_j(b, h_{\gamma,j} y_j) \\ \Leftrightarrow & (b, h_{\gamma,j} y_j) = (\tau_j^{-1} \circ \tau_i)(b, h_{\gamma,i} y_i) = (b, g_{ji}(b) h_{\gamma,i} y_i) \end{aligned}$$

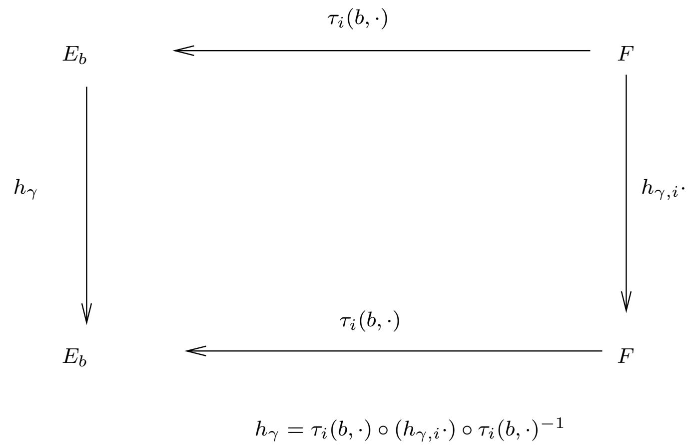

FIGURE 17. The relation between local holonomy and global holonomy.

$$\Leftrightarrow h_{\gamma,j}y_j = g_{ji}(b)h_{\gamma,i}y_i$$

$$\Leftrightarrow h_{\gamma,j}g_{ji}(b)y_i = g_{ji}(b)h_{\gamma,i}y_i$$

Since this will be true for all y_i and the action of G on F is faithful, we see the above is equivalent to $h_{\gamma,j}g_{ji}(b) = g_{ji}(b)h_{\gamma,i}$ i.e. $h_{\gamma,j} = g_{ji}(b)h_{\gamma,i}g_{ji}(b)^{-1}$.

Moreover, the local holonomy is related to the global holonomy as follows. The mapping J : holonomy$(b,i) \to$ holonomy(b) : $h_{\gamma,i} \mapsto h_\gamma$ defines a surjective group homomorphism. If the structure group acts on the standard fiber faithfully, J defines a group isomorphism. To see (refer to figure 18), assume that $h_{\gamma,i}, h_{\tilde{\gamma},i} \in G$ and $h_\gamma = h_{\tilde{\gamma}}$, then for all $y_i \in F$, we have $h_{\gamma,i}\cdot y_i = h_{\tilde{\gamma},i}\cdot y_i$. So $h_{\gamma,i}$ and $h_{\tilde{\gamma},i}$ induce the same mapping on F. For a faithful action $G\times F \to F$, we conclude that $h_{\gamma,i} = h_{\tilde{\gamma},i}$, therefore holonomy$(b,i) \to$ holonomy(b) is one-to-one, and by the definition of the holonomy(b), the mapping is always onto. In summary, J is an isomorphism.

7.2. Introduction to Berry's Phase

Under the same context as §6.3.3, except that we assume the standard fiber to be one dimensional, the Berry-Simon connection 1-form is:

$$\omega(\gamma(t),\gamma'(t)) = \mathbf{e}(\gamma(t))^\dagger \frac{d}{dt}\mathbf{e}(\gamma(t)) = (\mathbf{e}(\gamma(t)), \frac{d}{dt}\mathbf{e}(\gamma(t))),$$

where $\mathbf{e}(b) \in E_b$ and $\|\mathbf{e}(b)\| = 1$, i.e. $\{\mathbf{e}(b)\}$ is a normalized basis of E_b.

We claim that ω is a purely imaginary number. To see,

$$\begin{aligned}
&\mathbf{e}(\gamma(t))^\dagger \mathbf{e}(\gamma(t)) = 1 \\
\Rightarrow&[\frac{d}{dt}\mathbf{e}(\gamma(t))]^\dagger \mathbf{e}(\gamma(t)) + \mathbf{e}(\gamma(t))^\dagger \frac{d}{dt}\mathbf{e}(\gamma(t)) = 0 \\
\Rightarrow&[\frac{d}{dt}\mathbf{e}(\gamma(t))]^\dagger \mathbf{e}(\gamma(t)) = -\mathbf{e}(\gamma(t))^\dagger \frac{d}{dt}\mathbf{e}(\gamma(t)) \\
\Rightarrow&\overline{\omega(\gamma(t), \gamma'(t))} = (\frac{d}{dt}\mathbf{e}(\gamma(t)), \mathbf{e}(\gamma(t))) = -\omega(\gamma(t), \gamma'(t)) \\
\Rightarrow&\omega(\gamma(t), \gamma'(t)) = i(-i\omega(\gamma(t), \gamma'(t))), \text{ where } -i\omega(\gamma(t), \gamma'(t)) \in \mathbb{R}. \quad \square
\end{aligned}$$

With the above result, we are able to find a simple solution for the ODE initial value problem $g'(t) = -\omega(\gamma(t), \gamma'(t))g(t); g(0) = 1$.

$$\begin{aligned}
\Rightarrow& g(t) = e^{i\theta(t)} \\
\Rightarrow& g'(t) = ie^{i\theta(t)}\theta'(t) \\
\Rightarrow& ie^{i\theta(t)}\theta'(t) = -\omega(\gamma(t), \gamma'(t))e^{i\theta(t)} \\
\Rightarrow& \theta'(t) = i\omega(\gamma(t), \gamma'(t)) = i\mathbf{e}(\gamma(t))^\dagger \frac{d}{dt}\mathbf{e}(\gamma(t)) \\
\Rightarrow& \theta(t_1) - \theta(t_0) = i\int_{t_0}^{t_1} \mathbf{e}(\gamma(t))^\dagger \frac{d}{dt}\mathbf{e}(\gamma(t))dt,
\end{aligned}$$

(The following material is adapted from [**4**].) In 1984, M. V. Berry published his very influential findings on the quantum phase factors arising in a cyclic adiabatic quantum evolution [**11**]. Since then this phase got the names 'geometric', 'topological', 'non-integrable' and 'Berry's phase'. Berry investigated a quantum system governed by the Hamiltonian which depends on time through the slowly varying parameters. Then according to the adiabatic theorem, the system evolves in one of its instantaneous eigenstates as predicted by the Schrődinger equation. After an adiabatic evolution ends and the system completes a closed path in the parameter space, the instantaneous eigenstate acquires a phase factor, dependent only on the path traced out in the parametric space. This served as a reason to name the phase geometric. The geometric phase, unlike the rest of the total phase, is independent of the rate at which the system state moves along the cyclic path. The difference between the total and geometric phases received the name of dynamical phase. It was Barry Simon [**37**] who first recognized the geometrical meaning of Berry's phase to be the holonomy in a fiber bundle over the parameter space.

Soon after Berry's discovery of the adiabatic phase, Y. Aharonov and J. Anandan released geometric phase from the adiabatic constraint [**3**]. They defined geometric phase for cyclic evolutions of the system state being an eigenstate of the time evolution operator. This phase reduces to the Berry's phase in the adiabatic limit. But contrary to the Berry's phase, the Aharonov-Anandan phase is defined in the projective Hilbert space, not the parameter space. It generalizes the geometric concept of the geometric phase.

Although there are no widely recognized practical applications of the geometric phase, its experimental observations have been reported in many fields of science. The largest group of experiments have been carried out on polarized light [**12, 13, 18**] and polarized neutrons [**7, 14, 22, 42, 43**]. The geometric phase has also been observed in magnetic resonance experiments [**38**], mesoscopic structures [**24**] and molecular systems [**27**]. Analogues of GP – the Hannay angles have been shown to exist in classical mechanical systems [**21**], the most famous example of which is the Foucault pendulum. For a more complete account on the geometric phase manifestations the reader is referred to the Resource Letter [**5**].

7.3. Berry Phase in H_3 System

Define two local trivializations on the Hermitian vector bundle on the H_3 system as follows. $\tau_0 : U_0 \times \mathbb{C}^k \to \pi^{-1}(U_0)$ and $\tau_1 : U_1 \times \mathbb{C}^k \to \pi^{-1}(U_1)$, where U_0 and U_1 are open in $B = \mathcal{C}_k$. Let $\gamma_0 : (-\epsilon, 1+\epsilon) \to U_0$ be a smooth path such that $\gamma_0(0) = q, \gamma_0(1) = p$;

let $\gamma_1 : (-\epsilon, 1+\epsilon) \to U_1$ be a smooth path such that $\gamma_1(0) = p, \gamma_1(1) = q$. (See figure 18.) Then $\tau_1 \circ \tau_0^{-1} : (U_0 \cap U_1) \times \mathbb{C}^k \to (U_0 \cap U_1) \times \mathbb{C}^k : (b, y) \mapsto (b, g_{01}(b)y)$, where $g_{01} : U_0 \cap U_1 \to \mathfrak{U}(k)$ is a smooth cocycle. $g(t)$ and $\tilde{g}(t)$ are the solutions of the initial value problems: $g'(t) = -\omega_1(\gamma_1(t), \gamma_1'(t))g(t)$, $g(0) = I$; and $\tilde{g}'(t) = -\omega_0(\gamma_0(t), \gamma_0'(t))\tilde{g}(t)$, $\tilde{g}(0) = I$ respectively.

As shown in figure 18, γ_1 followed by γ_0 gives a closed path called γ at $p \in B$. The holonomy transformation in the trivialization τ_1 is given through the following steps:

(1) $y \in \mathbb{C}^k$ (in τ_1);

(2) $g(1)y \in \mathbb{C}^k$ (in τ_1);

(3) $g_{01}(q)g(1)y \in \mathbb{C}^k$ (in τ_0);

(4) $\tilde{g}(1)g_{01}(q)g(1)y \in \mathbb{C}^k$ (in τ_0);

(5) $g_{10}(p)\tilde{g}(1)g_{01}(q)g(1)y \in \mathbb{C}^k$ (in τ_1).

$h_{\gamma,1} = g_{10}(p)\tilde{g}(1)g_{01}(q)g(1)$ is the holonomy transformation in τ_1, i.e. the Berry phase on the loop γ in τ_1. Generally, it is not easy to know what the dimensionality k of the standard fiber $\mathbb{C}^k$ is, neither are the trivializations τ_0 nor τ_1 explicitly known. Thus we do not have the explicit formula to calculate the Berry phase of the H_3 system, but what we do above explains the computational procedure once τ_0, τ_1 and k are known to us.

Remark(II): potential energy surfaces of H_3 at the center of the cone We continue now our discussion from §6.3.5. Mead et.al's paper [**41**] claims that the curvature is zero almost everywhere on the cone except at some points where the potential energy singularities exist. By the relation between curvature, connection and holonomy in differential geometry [**25**], the holonomy and hence the Berry phase is likely to be more complex if the H_3 molecule travels around a loop enclosing the central axis of the cone.

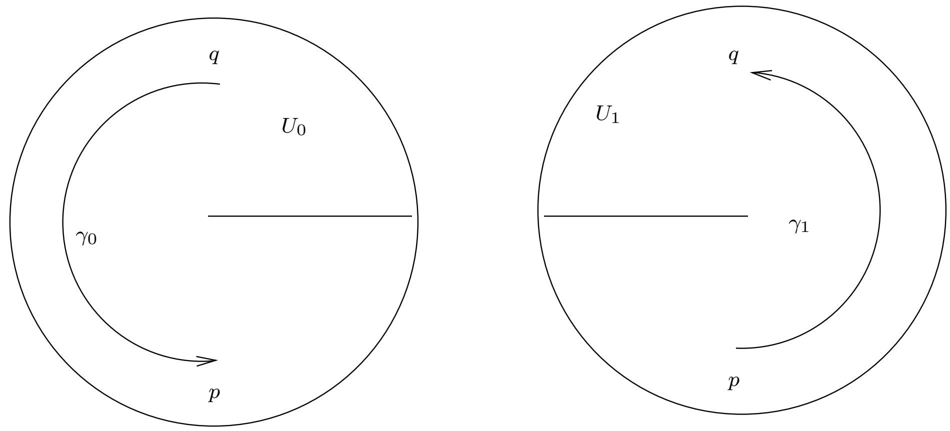

FIGURE 18. Holonomy Transformation in H_3 in trivialization τ_1.

Bibliography

[1] M.Abramowitz and I.A. Stegun *Handbook of mathematical functions* Dover Publication Inc. NY. 1972.

[2] Ravinder Abrol and Aron Kuppermann, *An optimal adiabatic to diabatic transformation of the 1A' and 2A' states of H_3* Thesis preprint, Sept.2001.

[3] Y. Aharonov and J. Anandan, *Phase change during a cyclic quantum evolution, Phys. Rev. Lett. 58 (1987), 1593-1596.*

[4] R. Aleksiejunas and V. Ivaska, *"Geometric phase estimation using time-frequency analysis", J. Phys. A: Math. Gen., V. 34, p. 8835-8850 (2001).* Also on http://signalogram.free-hosting.lt/Software/AnalysisHelp/GPDecHelp/introduction.htm *Introduction To the Geometric Phase Decomposition*

[5] J. Anandan, J. Christian, and K. Wanelik, Resource letter GPP-1: *Geometric phases in physics, Am. J. Phys. 65 (1997), 180-185.*

[6] P.W. Atkins *Physical Chemistry*, Fourth Edition.

[7] G. Badurek, H. Rauch, and D. Tuppinger, *Neutron interferometric double-resonance experiment, Phys. Rev. A. 34 (1986), 2600-2608.*

[8] Baldi P., and Brunak S., *Bioinformatics: The Machine Learning Approach.(1998)*

[9] Roger Balian, *From Microphysics to Macrophysics*

[10] D.R.Bates, K.ledsham, A.L. Stewwart *Wave functions of the hydrogen molecular ion* Philosophical Transaction of Loyal Society of London. Series A. Vol246. I911(1953) Pages 215-240.

[11] M. V. Berry, *Quantum phase factors accompanying adiabatic changes, Proc. R. Soc. A. 392 (1984), 45-57.*

[12] R.Bhandari, *Polarization of light and topological phases, Phys. Reports 281 (1997), 1-64.*

[13] R. Bhandari and J. Samuel, *Observation of topological phase by use of a laser interferometer, Phys. Rev. Lett. 60 (1988), 1211-1213.*

[14] T. Bitter and D. Dubbers, *Manifestation of Berry topological phase in neutron spin rotation, Phys. Rev. Lett. 59 (1987), 251-254.*

[15] Arnold I. Boothroyd, William J. Keogh, Peter G. Martin, and MIchael R. Peterson *A refined H_3 potential energy surface* J. Chem. Phys. 104(18), 8 May 1996.

[16] R. Bott and S. Chern, *Acta Math. 114, 71(1965)*

[17] Manfredo P. Carmo *Differential Geometry of Curves and Surfaces*

[18] R. Y. Chiao, A. Antaramian, K. M. Ganga, H. Jiao, S. R. Wilkinson, and H. Nathel, *Observation of a topological phase by means of a nonplanar Mach-Zehnder interferometer, Phys. Rev. Lett. 60 (1988), 1214-1217.*

[19] John B. Conway, *A Course in Functional Analysis*

[20] Cotton, F. A. *Chemical Applications of Group Theory, 3rd ed. 1990.*

[21] J. H. Hannay, *Angle variable holonomy in adiabatic exursion of an integrable Hamiltonian, J. Phys. A: Math. Gen. 18 (1985), 221-230.*

[22] Y. Hasegawa, M. Zawisky, H. Rauch, and A. I. Ioe, *Geometric phase in coupled neutron interference loops, Phys. Rev. A 53 (1996), 2486-2492.*

[23] N.P.Landsman *Mathematical Topics between Classical and Quantum Mechanics* Springer, 1998.

[24] D. Loss and P. M. Goldbart, *Persistent currents from Berry's phase in mesoscopic systems, Phys. Rev. B 45 (1992), 13544-13561.*

[25] Shoshichi Kobayashi and Katsumi Nomizu *Foundations of Differential Geometry I.*

[26] Donald Allan McQuarrie, *Statistical Mechanics*

[27] C. A. Mead, *The geometric phase in molecular systems, Rev. Mod. Phys. 64 (1992), 51-85.*

[28] Mikio Nakahara *Geometry, Topology and Physics*

[29] von Neumann *Mathematical Foundation of Quantum Mechanics* Princeton Press, NJ, 1955.

[30] M. Ozawa *Canonical approximation quantum measurement* J.Math.Phys. 34(12) Dec, 1993.

[31] K.R. Parthasarathy *Introduction to Probability and Measure* MacMillian Press LTD. 1977, London.

[32] R.Pauncz, *Spin Eigenfunctions, Construction and Use* Plenum Press, NY, 1979.

[33] Bent E. Peterson, *Introduction to the Fourier Transform and Pseudo-differential Operators*

[34] Mike Reed and Barry Simon *Methods of Morden Mathematical Physics Vol.I*

[35] D. Robert *Semi-classical approximation in quantum mechanics: A survey of recent and old mathematical results* Helv.Phys.Acta 71. 44-116(1998).

[36] H.L.Royden, *Real analysis*

[37] Barry Simon *Holonomy, the Quantum Adiabatic Theorem, and Berry's Phase* Physical Review Letters Vol.51,NO.24. 1983.

[38] D. Suter, K. T. Mueller, and A. Pines, *Study of the Aharonov-Anandan quantum phase by NMR interferometry, Phys. Rev. Lett. 60 (1988), 1218-1220.*

[39] F.Treves, *Topological Vector Spaces, Distribution and kernels, 1967*

[40] V.S.Varadarajan, *Geometry of Quantum Theory*

[41] Antonio J. C. Varandas, Franklin B. Brown, C. Alden Mead, Donald G. Truhlar, and Normand C. Blais, *A double many-body expansion of the two lowest-energy potential surfaces and nonadiabatic coupling for* H_3 J. Chem. Phys. 1 June, 1987.

[42] A. G. Wagh et al., *Experimental separation of geometric and dynamical phases using neutron interferometry, Phys. Rev. Lett. 78 (1997), 755-759.*

[43] A. G. Wagh, V. C. Rakhecha, P. Fischer, and A. Ioe, *Neutron interferometric observation of noncyclic phase, Phys. Rev. Lett. 81 (1998), 1992-1995.*

[44] Eric W. Weisstein, *www.mathworld.wolfram.com*

[45] Y.-S.Mark Wu, Aron Kuppermann and James B. Anderson, *A very high accuracy potential energy surface for* H_3 Phys. Chem. 1999, 1, 929-937.

[46] *http://hyperphysics.phy-astr.gsu.edu/hbase/spin.html*