# Non-commutative Probability: From Classical to Quantum Probability

David Paul Dimech

Supervised by Prof Emanuel Chetcuti

DEPARTMENT OF Mathematics
FACULTY OF SCIENCE
UNIVERSITY OF MALTA

September 2025

A Dissertation submitted in partial fulfilment of the requirements for the Degree of Master of Science in Mathematics

## Abstract

In classical probability theory, it is implicitly assumed that random variables commute. However, this assumption does not necessarily hold in all mathematical frameworks, such as those involving matrices. In this thesis, we will explore quantum probability, a non-commutative extension of classical probability. The main aim of this thesis shall be to examine how key concepts from classical probability can be generalized in the quantum setting.

## Acknowledgments

I would like to thank God for giving me the strength to finish this thesis. In addition, I would like to thank Prof Chetcuti for his guidance and fruitful discussions which helped hone my skill in the subject.

# Contents


Abstract i

Acknowledgments ii

Glossary 1

Chapter 1. Introduction 4

Chapter 2. Prerequisites 6

2.1. Stochastic Processes 6

2.2. Stochastic Integration 24

Chapter 3. Foundations of Quantum Probability 31

3.1. Elementary Notions of Quantum Probability 31

3.2. Gleason's Theorem 33

Chapter 4. Quantum Probability Spaces 50

4.1. Proving the Spectral Theorem 50

4.2. Finite Probability Spaces 54

Chapter 5. Multiple Probability Spaces 63

5.1. Tensor Product of Hilbert Spaces 63

5.2. Tensor Product of Operators 69

5.3. Symmetric and Antisymmetric Tensor products 73

5.4. Fock Space 75

Chapter 6. Quantum Stochastic Integration 78

6.1. Creation and Annihilation Operators 78

6.2. Probability Gauge Spaces and Non Commutative $L^p$ Spaces 82

6.3. The Itô-Clifford Integral 88

Chapter 7. Conclusion 95

7.1. Summary of Results 95
7.2. Further Research 95
Bibliography 97
Appendix A. Conditional Expectation 100

# Glossary

| Symbol | Description |
|---|---|
| $Y$ | Random Variable |
| $\Omega$ | Domain of a Random Variable or of an Observable |
| $\mathcal{F}$ | Sigma Algebra |
| $f_Y$ | Probability Density Function |
| $\mathcal{B}(\mathbb{R}^k)$ | Borel sets in $\mathbb{R}^k$ |
| $\mathbb{E}[\cdot]$ | Expectation |
| $\mathrm{Var}(\cdot)$ | Variance |
| $\mathrm{Cov}(\cdot,\cdot)$ | Covariance |
| $T$ | Indexing Set |
| $\mathbb{R}^T$ | Set of Functions on $\mathbb{R}$ |
| $(Y_t)_{t\in T}$ | Stochastic Process |
| $\mathbb{P}$ | Probability Measure |
| $u^*$ | Classical Measure |
| $\nu_{t_1,\cdots,t_k}$ | Finite Dimensional Distributions |
| $\mathcal{R}_0^T$ | Family of Cylinder Sets |
| $D_{\mathbb{Q}}$ | Dyadic Rationals |
| $\Sigma$ | Variance-Covariance Matrix |
| $m_Y(\cdot)$ | Mean Function |
| $k_Y(\cdot,\cdot)$ | Covariance Function |
| $\nu_{\mathbb{W}}$ | Weiner Measure |
| $(B_t)_{t\in[0,T]}$ | Brownian Motion |
| $\{\mathcal{F}_t : t \in T\}$ | Filtration |
| $L^2([a,b],\Omega)$ | Space of Adapted Stochastic Processes |
| $\mathcal{H}$ | Separable Hilbert Space |
| $\mathcal{A}$ | Lattice of Closed Subspaces |
| $\mathcal{P}(\mathcal{H})$ | Set of Orthogonal Projections over $\mathcal{H}$ |

*Continued on next page*

| Symbol | Description |
|---|---|
| $\pi(\cdot)$ | Spectral Measure |
| $\mu$ | Quantum Probability Measure |
| $\rho$ | State |
| $\vee$ | Join |
| $W$ | Weight |
| $Y_{lm}$ | Spherical Harmonics |
| $Q_l$ | Space of Spherical Harmonics of degree $l$ |
| $P_l^m$ | Associated Legendre Polynomial of index $l, m$ |
| $\mathrm{osc}(\cdot, \cdot)$ | Oscillation |
| $\mathrm{Tr}(\cdot)$ | Trace |
| $P_A$ | Orthogonal Projection onto Subspace $A$ |
| $G$ | Hermitian Matrix |
| $X$ | Observable |
| $\mathfrak{F}$ | Family of Matrices |
| $K$ | Kernel |
| $K(H)$ | Space of Kernels on the set $H$ |
| $\sigma$ | Permutation of $\{1, 2, \cdots, n\}$ |
| $\mathcal{S}_n$ | Symmetrizer |
| $\mathcal{S}_n^a$ | Antisymmetrizer |
| $\mathcal{H}^{\otimes^n}$ | $n$-fold Tensor Product of $\mathcal{H}$ |
| $\mathcal{H}^{\otimes_s^n}$ | $n$-fold Symmetric Tensor Product of $\mathcal{H}$ |
| $\mathcal{H}^{\otimes_a^n}$ | $n$-fold Antisymmetric Tensor Product of $\mathcal{H}$ |
| $\Gamma_{fr}(\mathcal{H})$ | Free Fock Space |
| $\Gamma_s(\mathcal{H})$ | Boson Fock Space |
| $\Gamma_a(\mathcal{H})$ | Fermion Fock Space |
| $\phi_0$ | Vacuum Vector |
| $a(\cdot)$ | Annihilation Operator |
| $a^\dagger(\cdot)$ | Creation Operator |
| $\Gamma_a^0(\mathcal{H})$ | Finite Particle Vectors in $\Gamma_a(\mathcal{H})$ |
| $J$ | Conjugation |
| $\mathcal{B}(\mathcal{H})$ | Bounded Operators on $\mathcal{H}$ |
| $\mathfrak{U}$ | Von Neumann Algebra |

*Continued on next page*

| Symbol | Description |
|---|---|
| $m(\cdot)$ | Gauge |
| $m(\cdot\|\mathcal{B})$ | Conditional Expectation with respect to $\mathcal{B}$ |
| $\mathfrak{h}$ | Space of Non-Commutative Adapted Processes |

CHAPTER 1

# Introduction

Classical probability theory was initially formalized in a measure theoretic manner by Kolmogorov in 1933. Until this day, it remains a highly active area of research, offering a rich interplay between deep theoretical developments and a wide range of practical applications. Despite this however, it is not without its limitations. Our main focus in this thesis shall be relaxing the underlying assumption of commutativity present in Kolmogorov's probability. This was prompted by fields such as quantum mechanics where the mathematical objects used do not satisfy this assumption.

The field would come to be known as Non-Commutative Probability. Today there exist several frameworks of Non-Commutative Probability. Two such frameworks are Free Probability spurred mainly by Dan Voiculescu [26] which he initially developed to study free group factors and Quantum Probability which was formalized by John Von Neumann [27] and later expanded on by others such as Luigi Accardi [1]. Our dissertation shall focus on the latter.

In particular we shall be seeing how aspects from classical probability generalize to the quantum framework. In Chapter 2 we shall be discussing some fundamental prerequisites from classical probability ranging from the Kolmogorov Extension and Continuity Theorems to the construction of the Itô integral. In Chapter 3 we give some elementary definitions related to quantum probability and also prove Gleason's Theorem, a very important theorem which shall give us the general form of a measure in this framework.

Following this, we start our discussion in Chapter 4 with the proof of the Spectral Theorem, another important theorem which we shall utilize when calculating probabilities. We shall also see how to construct a quantum probability space and show some examples of finite dimensional quantum probability spaces. As in classical probability, we may combine multiple quantum probability spaces into one and this is precisely what is discussed in Chapter 5. One particular important space is the Fock space, which we discuss in this chapter as well.

We then end our discussion on Quantum Probability by constructing the quantum stochastic integral which, in turn, requires discussion of non-commutative $L^p$ spaces and so-called probability gauge spaces.

CHAPTER 2

# Prerequisites

## 2.1. Stochastic Processes

One of the extensions of classical probability to the quantum framework we shall be discussing is the construction of a quantum stochastic integral. This chapter shall be dedicated to giving a brief background of the classical stochastic integral. We start with an elementary definition. All results for this section, unless otherwise stated, are taken from [17]. As a matter of notation, we fix $\mathcal{B}(\mathbb{R}^k)$ to be the Borel subsets of $\mathbb{R}^k$.

DEFINITION 2.1 (Random Variable). *Let $\Omega$ be a set and $\mathcal{F}$ a $\sigma$-algebra on $\Omega$. A random variable $Y$ is a map $Y : \Omega \to \mathbb{R}$ which is measurable with respect to $\mathcal{F}$.*

When studying random variables, it might be of interest to analyse the behaviour of a particular quantity across time. Mathematically this would mean that we are analysing random variables $Y_1, Y_2, ...$ that is a set of random variables indexed by time. The indexing of random variables across time is formalized by the notion of a stochastic process which we discuss later. To each random variable $Y$ we may associate a probability density function $f_Y(y)$ which satisfies the following

$$\mathbb{P}(\{\omega : Y(\omega) \in A\}) = \int_A f_Y(y)dy \quad A \in \mathcal{B}(\mathbb{R}).$$

As to abide by the axioms of probability, we require that $f_Y(y)$ is non-negative and that $\int_{\mathbb{R}} f_Y(y)dy = 1$. Two other very important quantities are the expectation as well as the variance.

Definition 2.2 (Expectation and Variance of a Random Variable). *The expectation* $\mathbb{E}[Y]$ *of a random variable* $Y$ *is given by*

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy.$$

*The variance is given by*

$$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y]))^2.$$

The notion of the variance may be generalized to covariance which expresses how two variables change with respect to one another.

Definition 2.3. *Let* $Y_1$ *and* $Y_2$ *be two random variables then the covariance is given by*

$$\mathrm{Cov}(Y_1, Y_2) = \mathbb{E}[Y_1 Y_2] - \mathbb{E}[Y_1]\mathbb{E}[Y_2].$$

We may now proceed with the notion of a stochastic process.

Definition 2.4 (Stochastic Process). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. A stochastic process is a collection of random variables* $(Y_t)_{t \in T}$ *where each* $Y_t$ *is defined on* $(\Omega, \mathcal{F}, \mathbb{P})$*. The set* $T$ *is said to be the indexing set of the process.*

There are various types of stochastic processes, which find applications in many fields. For example, branching processes are widely used in epidemiology to model the spread of diseases. The interested reader is referred to [25] for further details.

In the coming theorems, we mention regularity of the measure. As the definition is not always the same across sources, we elect to provide it as well as the definition of a measure for the sake of completeness and clarity.

Definition 2.5 (Classical Measure). *Let* $\Omega$ *be a set and* $\mathcal{F}$ *be a* $\sigma$*-algebra over* $\Omega$*. Let* $\mu^* : \Omega \to [0, 1]$ *be a set function. Then* $\mu^*$ *is said to be a measure if the following hold:*

(i) *For all* $E \in \mathcal{F}, \mu^*(E) \geq 0$.

(ii) $\mu^*(\emptyset) = 0$.

(iii) *For all countable pairwise disjoint collections* $\{E_i\}_{i=1}^{\infty}$ *satisfying* $E_i \in \mathcal{F}$ *for all* $i$,

$$\mu^*\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu^*(E_i).$$

Definition 2.6 (Regularity). *Let* $\Omega$ *be a set and* $\mathcal{F}$ *be a* $\sigma$*-algebra over* $\Omega$. *Then a measure* $\mu^*$ *is said to be regular if for every measurable set* $A$ *the following hold*

$$\mu^*(A) = \sup_{K \subseteq A} \{\mu^*(K) : K \textit{ compact and measurable}\},$$

$$\mu^*(A) = \inf_{A \subseteq U} \{\mu^*(U) : U \textit{ open and measurable}\}.$$

**2.1.1. Kolmogorov's Extension Theorem.** The existence of a stochastic process is a non-trivial matter, indeed one often cannot prove directly that a stochastic process exists. Hence, it was necessary to derive results with which we may show existence and in this section we shall discuss the Kolmogorov Extension Theorem which is an example of such a result. In this section we shall be going through all the theory required to establish this result. The proofs related to this theorem are found in [6].

The intuition behind this result lies in the following observation. We may easily see from Definition 2.4 that a stochastic process may be equivalently described as a collection of random variables $\{Y_t : t \in T\}$ where $T$ is some ordered indexing set. Take a $k$-tuple $(t_1, ..., t_k)$ of distinct elements of $T$, then the vector of random variables $(Y_{t_1}, ..., Y_{t_k})$ has a distribution $\nu_{t_1,...,t_k}$ over $\mathbb{R}^k$

$$\nu_{t_1,\dots,t_k}(H) = \mathbb{P}[\{\omega : (Y_{t_1}(\omega), \dots, Y_{t_k}(\omega)) \in H\}], \; H \in \mathcal{B}(\mathbb{R}^k). \tag{1}$$

The measures $\nu_{t_1,\dots,t_k}$ are said to be the finite dimensional distributions of the stochastic process $\{Y_t : t \in T\}$.

Now (1) implies two properties of $\nu_{t_1,\dots,t_k}$ which shall be of great importance for the theorem. Let $H$ in (1) have the form $H = H_1 \times \dots \times H_k$ for $H_i \in \mathcal{B}(\mathbb{R})$. Consider a permutation $\sigma$ of $(1, 2, \dots, k)$. Since $[(Y_{t_1}(\omega), \dots, Y_{t_k}(\omega)) \in H_1 \times \dots \times H_k]$ and $[(Y_{t_{\sigma(1)}}(\omega), \dots, Y_{t_{\sigma(k)}}(\omega)) \in H_{\sigma(1)} \times \dots \times H_{\sigma(k)}]$ are the same event, we are justified by (1) in concluding that

(2) $$\nu_{t_1,\ldots,t_k}(H_1 \times \cdots \times H_k) = \nu_{t_{\sigma(1)},\ldots,t_{\sigma(k)}}(H_{\sigma(1)} \times \cdots \times H_{\sigma(k)}).$$

For example, if $\nu_{s,t} = \nu_1 \times \nu_2$, then necessarily $\nu_{t,s} = \nu_2 \times \nu_1$.
The second property is

(3) $$\nu_{t_1,\ldots,t_{k-1}}(H_1 \times \cdots \times H_{k-1}) = \nu_{t_1,\ldots,t_k}(H_1 \times \cdots \times H_{k-1} \times \mathbb{R}^1).$$

This is clear because $(Y_{t_1}(\omega), \ldots, Y_{t_{k-1}}(\omega))$ lies in $H_1 \times \cdots \times H_{k-1}$ if and only if $(Y_{t_1}(\omega), \ldots, Y_{t_{k-1}}(\omega), Y_{t_k}(\omega))$ lies in $H_1 \times \cdots \times H_{k-1} \times \mathbb{R}$. We shall call (2) and (3) the consistency conditions of the measure $\nu_{t_1,\ldots,t_{k-1}}$. We may combine these into one condition. Indeed, let $(u_1, \cdots, u_m)$ where $k \leq m$ be an $m$-tuple of distinct elements of $T$ such that each of $t_1, t_2, \cdots, t_k$ corresponds to some distinct $u_i$ for $i \in \{1, \cdots, m\}$. Then $(t_1, \cdots, t_k)$ must be the initial segment of some permutation of $(u_1, \cdots, u_m)$ that is, for some permutation $\sigma$ of $\{1, \cdots, m\}$, we have

$$(u_{\sigma^{-1}(1)}, \cdots, u_{\sigma^{-1}(m)}) = \{t_1, \cdots, t_k, t_{k+1}, \cdots, t_m\}$$

where $t_{k+1}, \cdots, t_m$ are elements of $(u_1, \cdots, u_m)$ which are not present in $(t_1, \cdots, t_k)$. We define $\psi : \mathbb{R}^m \to \mathbb{R}^k$ by

(4) $$\psi(x_1, \cdots, x_m) = (x_{\sigma^{-1}(1)}, \cdots, x_{\sigma^{-1}(k)}).$$

That is $\psi$ applies $\sigma$ to the coordinates and then projects onto the first $k$ of them. Now, as $\psi(Y_{u_1}, \cdots, Y_{u_m}) = (Y_{t_1}, \cdots, Y_{t_k})$ and so

(5) $$\nu_{t_1,\cdots,t_k} = \nu_{u_1,\cdots,u_m}\psi^{-1}.$$

It may easily be verified that (5) contains (2) and (3) as special cases. Furthermore $\psi$ is a coordinate permutation followed by a sequence of projections of the type $(x_1, \cdots, x_l) \to (x_1, \cdots, x_{l-1})$ and so equivalence follows.

While it is trivial that the finite dimensional distributions of any stochastic process satisfy these conditions, it may be shown that the converse also holds. That is, if a system of measures satisfies the consistency conditions, then there exists a stochastic process having the measures in this system as its finite dimensional distributions.

Prior to proceeding with the statement of this theorem however, we need to first discuss the notion of cylinder sets.

To start off we let $T$ be any indexing set equipped with some ordering and we denote the collection of all real functions on $T$ by $\mathbb{R}^T$. An element of $\mathbb{R}^T$ shall be written as $z$ and the value of $z$ at $t$ shall be given by $z_t$. The following shall be fundamental later on:

Definition 2.7 (Coordinate Function). *For $t \in T$ define $Y_t : \mathbb{R}^T \to \mathbb{R}$ by*

$$Y_t(z) = z_t.$$

*Then $Y_t$ is said to be a coordinate function.*

Remark 2.1. *We note the conflicting notation with the definition of a random variable given in Definition 2.1 however later when we define a probability measure on $\mathbb{R}^T$ these shall be random variables and so we elect to ignore this conflict.*

Now we let $\mathcal{R}^T$ be the $\sigma$-field generated by all the coordinate functions $Y_t$. Then $\mathcal{R}^T$ is generated by sets of the form

$$\{z \in \mathbb{R}^T : Y_t(z) \in H\}$$

for $t \in T$ and $H \in \mathcal{B}(\mathbb{R})$. Consider the class $\mathcal{R}_0^T$ consisting of cylinders which we define below.

Definition 2.8. *Let*

$$A = \{z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_k}(z)) \in H\} \tag{6}$$

*whereby $H \subseteq \mathbb{R}^k$ for some $k \in \mathbb{N}$ and $t_1, t_2 \cdots, t_k$ are distinct elements of $T$. Then $A$ is said to be a cylinder.*

Trivially we may see that $\mathcal{R}_0^T$ generates $\mathcal{R}^T$. It may be shown that $\mathcal{R}_0^T$ is a field.

Theorem 2.1. *$\mathcal{R}_0^T$ is a field.*

Proof. The complementation of (6) is precisely the set $\mathbb{R}^T \setminus A = \{z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_k}(z)) \in \mathbb{R}^k \setminus H\}$ showing $\mathcal{R}_0^T$ is closed when taking complements. Let

$$A_1 = \{z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_k}(z)) \in H\}$$

and

$$A_2 = \{z \in \mathbb{R}^T : (Y_{s_1}(z), \cdots, Y_{s_j}(z)) \in I\}$$

be two cylinder sets where $H \in \mathcal{B}(\mathbb{R}^k), I \in \mathcal{B}(\mathbb{R}^j)$. Let $(u_1, \cdots, u_m)$ be an $m$-tuple containing all the $t_\alpha$ and $s_\beta$. Then once again we argue that $(t_1, \cdots, t_k)$ must be the initial segment of some permutation of $(u_1, \cdots, u_m)$ and taking $\psi$ as defined in (4) we obtain that $H' = \psi^{-1}H$ for some $H' \in \mathcal{B}(\mathbb{R}^m)$. Thus $A_1$ may be written as

$$A_1 = \{z \in \mathbb{R}^T : (Y_{u_1}(z), \cdots, Y_{u_m}(z)) \in H'\}. \tag{7}$$

Similarly we have that $A_2$ is equivalently given by

$$A_2 = \{x \in \mathbb{R}^T : (Y_{u_1}(z), \cdots, Y_{u_m}(z)) \in I'\}. \tag{8}$$

Then we have that

$$A_1 \cup A_2 = \{z \in \mathbb{R}^T : (Y_{u_1}(z), \cdots, Y_{u_m}(z)) \in H' \cup I'\}.$$

As $H' \cup I' \in \mathcal{B}(\mathbb{R}^m)$, $A_1 \cup A_2$ is a cylinder showing $\mathcal{R}_0^T$ is a field. □

With these results established, we proceed to prove the main result of this section.

Theorem 2.2 (Kolmogorov Extension Theorem). *For every $\{t_1, \cdots, t_k\} \subseteq T$ where $k \in \mathbb{N}$, let $\nu_{t_1,\cdots,t_k}$ be probability measures which satisfy the consistency conditions in (2) and (3). Then there is a probability measure $\mathbb{P}$ on $\mathcal{R}^T$ such that the coordinate value process $(Y_t : t \in T)$ on $(\mathbb{R}^T, \mathcal{R}^T, \mathbb{P})$ has each of the $\nu_{t_1,\cdots,t_k}$ as its finite dimensional distributions.*

PROOF. Let $A_1$ be defined as in the proof of Theorem 2.1 and define

$$\mathbb{P}(A_1) = \nu_{t_1,\cdots,t_k}(H). \tag{9}$$

As the cylinder $A_1$ may have multiple representations, it is important that we verify consistency of the definition. Let $A_1$ coincide with the cylinder $A_2$ where $A_2$ is defined as in Theorem 2.1. As argued previously, if $(u_1, \cdots, u_m)$ contains each $t_\alpha$ and $s_\beta$ then $A_1$ is also given by (7) where $H' = \psi^{-1}H$. As the consistency conditions imply the condition (5) we have that $\mathbb{P}(A_1) = \nu_{t_1,\cdots,t_k}(H) = \nu_{u_1,\cdots,u_m}(H')$ and similarly $\mathbb{P}(A_2) = \nu_{s_1,\cdots,s_j}(I) = \nu_{u_1,\cdots,u_m}(I')$. As the $u_\gamma s$ are distinct then for any real numbers $w_1, \cdots, w_m$ there are points $z$ of $\mathbb{R}^T$ for which $(Y_{u_1}(z), \cdots, Y_{u_m}(z)) = (w_1, \cdots, w_m)$. Due to this, we may conclude that the cylinders in (7) and (8) coincide so that $H' = I'$. Thus $A_1 = A_2$ gives us that

$$\mathbb{P}(A_1) = \nu_{u_1,\cdots,u_m}(H') = \nu_{u_1,\cdots,u_m}(I') = \mathbb{P}(A_2).$$

Thus we have verified that (9) is indeed consistent.

Now we consider two disjoint cylinders $A_1$ and $A_2$ and without loss of generality we may assume that their index sets are identical so that $A_1$ and $A_2$ are given by (7) and (8) respectively. If $H' \cap I' \neq \emptyset$ then $A_1 \cap A_2 \neq \emptyset$ and thus by the contrapositive we must have that since $A_1$ and $A_2$ are disjoint then so are $H'$ and $I'$. The following is then evident

$$\begin{aligned}\mathbb{P}(A_1 \cup A_2) &= \nu_{u_1,\cdots,u_m}(H' \cup I') \\ &= \nu_{u_1,\cdots,u_m}(H') + \nu_{u_1,\cdots,u_m}(I') \\ &= \mathbb{P}(A_1) + \mathbb{P}(A_2).\end{aligned}$$

Showing $\mathbb{P}$ is finitely additive on $\mathcal{R}_0^T$ and clearly $\mathbb{P}(\mathcal{R}_0^T) = 1$. We shall now show that $\mathbb{P}$ is countably additive on $\mathcal{R}_0^T$ for if this holds we may use Carathéodory's Extension Theorem to extend it to a probability measure on $\mathcal{R}^T$. Due to the definition of $\mathbb{P}$ on $\mathcal{R}_0^T$, we will have that

$$\mathbb{P}[z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_k}(z)) \in H] = \nu_{t_1,\cdots,t_k}(H)$$

meaning the coordinate process $(Y_t : t \in T)$ will have the required finite-dimensional distributions.

Hence we proceed to show countable additivity on $\mathcal{R}_0^T$ and this will follow if $A_n \in \mathcal{R}_0^T$ and $A_n \downarrow \emptyset$ imply that $\mathbb{P}(A_n) \downarrow 0$. We show this by the contrapositive. Let $A_1 \supseteq A_2 \cdots$ such that $\mathbb{P}(A_n) \geq \epsilon$ for all $n$. As the index set of a cylinder may be permuted and expanded as previously discussed, there exists a sequence $t_1, t_2, \cdots$ of points in $T$ for which

$$A_n = \{z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_n}(z)) \in H_n\},$$

where $H_n \in \mathbb{R}^n$.

Clearly, $\mathbb{P}(A_n) = \nu_{t_1,\cdots,t_n}(H_n)$. Due to regularity there exists a compact set $K_n$ in $H_n$ satisfying $\nu_{t_1,\cdots,t_n}(H_n \setminus K_n) < \frac{\epsilon}{2^{n+1}}$. Let $A_n' = \{z \in \mathbb{R}^T : (Y_{t_1}(z), \cdots, Y_{t_n}(z)) \in K_n\}$ then we have $\mathbb{P}(A_n \setminus A_n') < \frac{\epsilon}{2^{n+1}}$. Also $\bigcap_{k=1}^n A_k' \subseteq A_n' \subseteq A_n$ and $\mathbb{P}(A_n \setminus \bigcap_{k=1}^n A_k') < \frac{\epsilon}{2}$ so that $\mathbb{P}(\bigcap_{k=1}^n A_k') > \frac{\epsilon}{2} > 0$. Thus, $\bigcap_{k=1}^n A_k' \subseteq \bigcap_{k=1}^{n-1} A_k'$ and $\bigcap_{k=1}^n A_k'$ is non-empty.

Choose a point $z^{(n)}$ of $\mathbb{R}^T$ in $\bigcap_{k=1}^n A_k'$. If $n \geq k$, then $z^{(n)} \in \bigcap_{m=1}^n A_m' \in \bigcap_{m=1}^k A_m'$ and hence $(z_{t_1}^{(n)}, \cdots, z_{t_k}^{(n)}) \in K_k$. As $K_k$ is bounded, the sequence $\{z_{t_k}^{(1)}, z_{t_k}^{(2)}, \cdots\}$ is bounded for each $k$. By the diagonal method, we may select an increasing sequence of integers $n_1, n_2, \cdots$ of integers such that $\lim_{i\to\infty} z_{t_k}^{(n_i)}$ exists for each $k$. There is in $\mathbb{R}^T$ some point $z$ whose $t_k$th coordinate is this limit for each $k$. Hence, for each $k$, $(z_{t_1}, \cdots, z_{t_k})$ is the limit of $(z_{t_1}^{(n_i)}, \cdots, z_{t_k}^{(n_i)})$ and hence lies in $K_k$. This however gives us that $z$ itself lies in $A_k'$ and hence in $A_k$. With this, we have obtained that $z \in \bigcap_{k=1}^{\infty} A_k$, completing the proof. □

**2.1.2. Kolmogorov Continuity Theorem.** While the Kolmogorov Extension Theorem gives us the existence of a particular measure, it does not guarantee that the associated process will have almost surely continuous paths that is, the set of discontinuous paths need not form a set of measure zero. This is undesirable as continuity is an important aspect of many stochastic processes.

For example, when modelling Lagrangian acceleration, it is essential that the stochastic models produce continuous velocity paths since velocities cannot jump discontinuously in time. For this reason, all established Lagrangian acceleration models such as those found in [3] are constructed to ensure continuity.

This guarantee of continuity is precisely what is given by the Kolmogorov Continuity Theorem which we shall discuss in this section. The proofs presented may be found in [5].

Prior to proceeding to the main result, we need to first define the notion of a modification of a process.

Definition 2.9 (Modification). *A stochastic process* $(Z_t)_{t\in\mathbb{R}}$ *is said to be a modification of another process* $(Y_t)_{t\in\mathbb{R}}$ *if the following holds for every* $t \in \mathbb{R}^+$

$$\mathbb{P}[\{\omega \in \Omega : Z_t(\omega) = Y_t(\omega)\}] = 1.$$

*Furthermore if* $(Z_t)_{t\in\mathbb{R}}$ *has* $\mathbb{P}$ *almost surely continuous paths, then it is said to be a continuous modification.*

As previously mentioned, continuous modifications help ensure continuity of the paths of the process. As an example of a continuous modification, we may refer to [21] which gives a method to construct continuous modifications on stochastic flows. Stochastic flows, in turn, are used to model phenomena such as the mean velocity profile of wall-bounded flows. The interested reader is directed to [20] for further detail on the subject. We now proceed to an important lemma related to modifications.

Lemma 2.1. *If* $(Z_t)_{t\in\mathbb{R}^+}$ *is a modification of* $(Y_t)_{t\in\mathbb{R}^+}$ *then* $(Y_t)_{t\in\mathbb{R}^+}$ *and* $(Z_t)_{t\in\mathbb{R}^+}$ *have the same finite dimensional distributions.*

Proof. For $t_1, t_2, \cdots, t_n \in \mathbb{R}^+$, let $A_i \in \mathcal{B}(\mathbb{R})$ for each $1 \leq i \leq n$. Furthermore define

$$A = \bigcap_{i=1}^{n} Y_{t_i}^{-1}(A_i) \in \mathcal{F},\ A' = \bigcap_{i=1}^{n} Z_{t_i}^{-1}(A_i) \in \mathcal{F}.$$

If $\omega \in A \setminus A'$ then for some $i$, $\omega \notin Z_{t_i}^{-1}(A_i)$, and $\omega \in Z_{t_i}^{-1}(A_i)$ so $Y_{t_i}(\omega) \neq Z_{t_i}(\omega)$. Therefore

$$A \triangle A' \subset \bigcup_{i=1}^{n} \{\omega \in \Omega : Y_{t_i}(\omega) \neq Z_{t_i}(\omega)\}.$$

Since $(Z_t)_{t\in\mathbb{R}^+}$ is a modification of $(Y_t)_{t\in\mathbb{R}^+}$, the right-hand side is a $\mathbb{P}$-null set as a union of finitely many $\mathbb{P}$-null sets. $A$ and $A'$ each belong to $\mathcal{F}$, so $P(A\triangle A') = 0$. Because $\mathbb{P}(A\triangle A') = 0$, $\mathbb{P}(A) = \mathbb{P}(A')$, meaning that

$$\mathbb{P}(\{\omega : Y_{t_1}(\omega) \in A_1, \ldots, Y_{t_n}(\omega) \in A_n\}) = \mathbb{P}(\{\omega : Z_{t_1}(\omega) \in A_1, \ldots, Z_{t_n}(\omega) \in A_n\})$$

hence establishing the result. □

THEOREM 2.3 (Kolmogorov Continuity Theorem). *Let $(Y_t)_{t\in\mathbb{R}^+}$ be a stochastic process in the measure space $(\Omega, \mathcal{F}, \mathbb{P})$. If there exists positive constants $\alpha, \beta$ and $c$ such that*

$$\mathbb{E}[|Y_t - Y_s|^\alpha] \leq c|t-s|^{1+\beta} \quad \textit{for } s,t \in \mathbb{R}^+ \tag{6}$$

*then $(Y_t)_{t\in\mathbb{R}^+}$ has a continuous modification that also satisfies (6).*

PROOF. Let $0 < \gamma < \frac{\beta}{\alpha}$. and we define $\delta = \beta - \alpha\gamma > 0$. For $m \geq 1$ we take $T_m$ to be the set of all pairs $(s,t)$ where $s,t \in \{j2^{-m} : 0 \leq j \leq 2^m\}$ and $|s-t| = 2^{-m}$. Clearly, $|T_m| = 2^{m+1}$. Let

$$A_m = \bigcup_{(s,t)\in T_m} \{\omega : |Y_s(\omega) - Y_t(\omega)| \geq 2^{-\gamma m}\} \in \mathcal{F}. \tag{7}$$

For $(s,t) \in T_m$ we may obtain the following through Markov's inequality as well as (6)

$$\begin{aligned}
\mathbb{P}(|Y_t - Y_s| \geq 2^{-\gamma m}) &\leq (2^{\gamma m})^\alpha \mathbb{E}[|Y_t - Y_s|^\alpha] \\
&\leq c2^{\alpha\gamma m}|t-s|^{1+\beta} \\
&= c2^{\alpha\gamma m}2^{-m(1+\beta)} \\
&< c2^{-m-\delta m}.
\end{aligned}$$

From this we have that

$$\mathbb{P}(A_m) \leq \sum_{(s,t)\in T_m} \mathbb{P}\big(\{\omega : |Y_s(\omega) - Y_t(\omega)| \geq 2^{-\gamma m}\}\big) < \sum_{(s,t)\in T_m} c2^{-m-\delta m} = c2^{1-\delta m}.$$

It is easily verified that $\sum_m \mathbb{P}(A_m) < \infty$ and thus, by the Borel-Cantelli Lemma, we have that

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m\right) = \mathbb{P}(N_0) = 0,$$

where for each $\omega \in \Omega \setminus N_0$ there is some $m_0(\omega)$ such that $\omega \notin A_m$ when $m \geq m_0(\omega)$. That is, for $\omega \in \Omega \setminus N_0$ there is some $m_0(\omega)$ such that

$$|Y_t(\omega) - Y_s(\omega)| < 2^{-\gamma m}, \ m \geq m_0(\omega), \ (s,t) \in T_m. \tag{8}$$

For the next part of the proof we shall utilize the dyadic rationals which are given by the set

$$D_{\mathbb{Q}} = \bigcup_{i=0}^{\infty}\{2^{-i}j : j \in \mathbb{Z}\}.$$

Now let $\omega \in \Omega \setminus N_0$ and $s,t \in D_{\mathbb{Q}} \cap [0,1]$ satisfy $0 < |s-t| \leq 2^{-m_0(\omega)}$. From here on, for a pair $(s,t)$ we fix $m = m(s,t)$ to be the greatest integer such that $|s-t| \leq 2^{-m}$.

$$2^{-m-1} < |s-t| \leq 2^{-m}, \tag{9}$$

which implies that $m \geq m_0(\omega)$. As $s,t$ are in $D_{\mathbb{Q}} \cap [0,1]$, there are some $i_0, j_0 \in \{0,1,2,\cdots,2^m\}$ such that

$$s_0 = i_0 2^{-m} \leq s < (i_0+1)2^{-m}, \quad t_0 = j_0 2^{-m} \leq t < (j_0+1)2^{-m}$$

As $0 \leq s - s_0 < 2^{-m}$ and $0 \leq t - t_0 < 2^{-m}$, there are sequences $(\varphi_n)_{n\in\mathbb{N}}$ and $(\tau_n)_{n\in\mathbb{N}}$ where $\varphi_j, \tau_j \in \{0,1\}$ for $i,j \in \mathbb{N}$ each of which having cofinitely many zero entries such that

$$s = s_0 + \sum_{j>m}\varphi_j 2^{-j}, \ t = t_0 + \sum_{j>m}\tau_j 2^{-j}.$$

Since $0 \leq s - s_0 < 2^{-m}$ and $0 \leq t - t_0 < 2^{-m}$ the following is evident

$$2^{-m} > |(s - s_0 - (t - t_0))| = |(s-t) - (s_0 - t_0)| \geq |s_0 - t_0| - |s-t|.$$

By (9) we get

$$|s_0 - t_0| < 2^{-m} + |s-t| \leq 2^{-m} + 2^{-m} = 2^{-m+1}.$$

Thus meaning $|i_0 - j_0| < 2$ which implies $|i_0 - j_0|$ may only assume a value of 0 or 1. Hence, we either have that $s_0 = t_0$ or $(s_0, t_0) \in T_m$. If the former holds we have $|Y_{t_0}(\omega) - Y_{s_0}(\omega)| = 0$. On the other hand, if the latter holds, as $m \geq m_0(\omega)$ we have, by (8), that

$$|Y_{t_0}(\omega) - Y_{s_0}(\omega)| < 2^{-\gamma m}. \tag{10}$$

We define inductively

$$s_n = s_{n+1} + \varphi_{m+n} 2^{-(m+n)} = s_0 + \sum_{m<j\leq m+n} \varphi_j 2^{-j}, \ n \geq 1.$$

For each $n \geq 1$, $s_n - s_{n-1} \in \{0, 2^{-(m+n)}\}$ implying that either $s_n = s_{n-1}$ or that $(s_{n-1}, s_n) \in T_{m+n}$ and since $m + n \geq m + 1 > m_0(\omega)$, applying (8) gives

$$|Y_{s_n}(\omega) - Y_{s_{n-1}}(\omega)| < 2^{-\gamma(m+n)}.$$

Since the sequence $s_n$ is eventually equal to $s$ we have that

$$\sum_{n=1}^{\infty} (Y_{s_n}(\omega) - (Y_{s_{n-1}}(\omega)) = Y_s(\omega) - Y_{s_0}(\omega).$$

Consider,

$$|Y_s(\omega) - Y_{s_0}(\omega)| \leq \sum_{n=1}^{\infty} |Y_{s_n}(\omega) - Y_{s_{n-1}}(\omega)| < \sum_{n=1}^{\infty} 2^{-\gamma(m+n)} = \frac{2^{-\gamma(m+1)}}{1 - 2^{-\gamma}} \tag{11}$$

and by similar reasoning we obtain that

$$|Y_t(\omega) - Y_{t_0}(\omega)| < \frac{2^{-\gamma(m+1)}}{1 - 2^{-\gamma}}.$$

From these inequalities and by (10) we obtain the following

$$\begin{aligned} |Y_t(\omega) - Y_s(\omega)| &\leq |Y_t(\omega) - Y_{t_0}(\omega)| + |Y_{t_0}(\omega) - Y_{s_0}(\omega)| + |Y_s(\omega) - Y_{s_0}(\omega)| \\ &< \frac{2^{-\gamma(m+1)}}{1 - 2^{-\gamma}} + 2^{-\gamma m} + \frac{2^{-\gamma(m+1)}}{1 - 2^{-\gamma}} \\ &= 2^{-\gamma(m+1)} C, \end{aligned} \tag{12}$$

where $C = 2^{\gamma} + \frac{2}{1-2^{-\gamma}}$. By (9), $2^{-(m+1)} < |t - s|$ and hence

$$|Y_t(\omega) - Y_s(\omega)| \leq C|t - s|^{\gamma}. \tag{13}$$

This is true for all $s, t \in D_{\mathbb{Q}} \cap [0, 1]$ with $|s - t| \leq 2^{-m_0(\omega)}$, when $|s - t| = 0$ it is immediate. Applying what we have just argued to $Y_{k+1}$, we may infer the existence of a $\mathbb{P}$-null set $N_1' \in \mathcal{F}$ such that for each $\omega \in \Omega \setminus N_1'$ there is some $m_1'(\omega)$ such that $m \geq m_j^i(\omega)$ and $(s,t) \in S_m$ give that $|Y_{t+i}(\omega) - Y_{s+i}(\omega)| < 2^{-\gamma m}$. We define $N_1 = N_0 \cup N_1'$, which is $\mathbb{P}$-null as a union of two $\mathbb{P}$-null sets. In addition, for $\omega \in \Omega \setminus N_1$, we let $m_1(\omega) = \max\{m_0(\omega), m_1'(\omega)\}$. For $s,t \in D_{\mathbb{Q}} \cap [0,1]$ satisfying $|s-t| \leq 2^{-m_1(\omega)}$, we have by previous arguments that

$$|Y_t(\omega) - Y_s(\omega)| \leq C|t-s|^\gamma, \qquad |Y_{t+1}(\omega) - Y_{s+1}(\omega)| \leq C|t-s|^\gamma.$$

Proceeding inductively, we get that for each $k \geq 1$ there are $\mathbb{P}$-null sets $N_0 \subset N_1 \subset \cdots \subset N_k$ whereby for each $\omega \in \Omega \setminus N_k$ there exists some $m_k(\omega)$ such that for $s,t \in D_{\mathbb{Q}} \cap [0,1]$ with $|s-t| \leq 2^{-m_k(\omega)}$,

$$\begin{aligned}
|Y_t(\omega) - Y_s(\omega)| &\leq C|t-s|^\gamma \\
|Y_{t+1}(\omega) - Y_{s+1}(\omega)| &\leq C|t-s|^\gamma \\
&\vdots \\
|Y_{t+k}(\omega) - Y_{s+k}(\omega)| &\leq C|t-s|^\gamma.
\end{aligned}$$

Let

$$N_\gamma = \bigcup_{k \geq 1} N_k,$$

which is an increasing sequence of sets whose union is $\mathbb{P}$-null. For $\omega \in \Omega \setminus N_\gamma$, we are guaranteed the existence of a nondecreasing sequence $(m_k(\omega))_{k \in \mathbb{N}}$ such that when $0 \leq j \leq k$ and $s,t \in D_{\mathbb{Q}} \cap [j, j+1]$ with $|s-t| \leq 2^{-m_k(\omega)}$, then $|Y_t(\omega) - Y_s(\omega)| \leq C|t-s|^\gamma$.

For $s,t \in D \cap [0, k+1]$ with $|s-t| \leq 2^{-m_k(\omega)}$, since $|s-t| \leq \frac{1}{2}$, then one of two cases may occur. Either there is some $0 \leq j \leq k$ for which $s,t \in [j, j+1]$ or there is some $1 \leq j \leq k$ for which, say, $s < j < t$. In the first case, $|Y_t(\omega) - Y_s(\omega)| \leq C|t-s|^\gamma$. In the second case, because $|j-s| \leq |t-s| \leq 2^{-m_k(\omega)}$ and $|t-j| \leq |t-s| \leq 2^{-m_k(\omega)}$, we may conclude that since $s,j \in D_{\mathbb{Q}} \cap [0,1]$ and $j,t \in D_{\mathbb{Q}} \cap [j,j+1]$, then

$$\begin{aligned}
|Y_t(\omega) - Y_s(\omega)| &\leq |Y_t(\omega) - Y_j(\omega)| + |Y_j(\omega) - Y_s(\omega)| \\
&\leq C|t-j|^\gamma + C|j-s|^\gamma \\
&\leq 2C|t-s|^\gamma.
\end{aligned}$$

Thus for

$$C_\gamma = 2C = 2^{\gamma+1} + \frac{4}{1-2^{-\gamma}},$$

we have established that for $\omega \in \Omega \setminus N_\gamma$, $k \geq 1$, and $s, t \in D_{\mathbb{Q}} \cap [0, k+1]$ satisfying $|t-s| \leq 2^{-m_k(\omega)}$, the following holds

$$|Y_t(\omega) - Y_s(\omega)| \leq C_\gamma |t-s|^\gamma \leq C_\gamma 2^{-m_k(\omega)}. \tag{14}$$

This gives us that for each $\omega \in \Omega \setminus N_\gamma$, and for $k \geq 1$, the mapping $t \mapsto Y_t(\omega)$ is uniformly continuous on $D \cap [0, k+1]$. For $t \in \mathbb{R}^+$ and $\omega \in \Omega \setminus N_\gamma$, define

$$Z_t(\omega) = \lim_{D \ni s \to t} Y_s(\omega). \tag{15}$$

We shall now establish that $(Z_t)_{t \in \mathbb{R}^+}$ is a continuous stochastic process. First we observe that, by (14) we have that the map

$$\begin{aligned} t &\mapsto Y_t(\omega) \\ D \cap [0, k+1] &\mapsto \mathbb{R}^d \end{aligned}$$

is uniformly continuous for every $\omega \in \Omega$ and trivially we have that $\mathbb{R}^d$ is a complete metric space. As $D_{\mathbb{Q}} \cap [0, k+1]$ is dense in $[0, k+1]$ we obtain that the map

$$\begin{aligned} f : t &\mapsto Z_t(\omega) \\ [0, k+1] &\mapsto \mathbb{R}^d \end{aligned}$$

is also uniformly continuous and thus we have that the map

$$\begin{aligned} t &\mapsto Z_t(\omega) \\ \mathbb{R}^+ &\mapsto \mathbb{R}^d \end{aligned}$$

is continuous. For $\omega \in N_\gamma$ we define $Z_t(\omega) = 0$ for every $t \in \mathbb{R}^+$. Thus we have obtained that

$$\begin{aligned} t &\mapsto Z_t(\omega) \\ \mathbb{R}^+ &\mapsto \mathbb{R}^d \end{aligned}$$

is continuous for every $\omega \in \Omega$. Next we shall verify the measurability of the $Z_t s$. For $t \in \mathbb{R}^+$, $Z_t$ is measurable as the pointwise limit of measurable functions and hence $(Z_t)_{t \in \mathbb{R}^+}$ is a continuous stochastic process.

For the final part of the proof we must verify that $(Z_t)_{t\in\mathbb{R}^+}$ is a modification of $(Y_t)_{t\in\mathbb{R}^+}$. For $s \in D_{\mathbb{Q}}$ and for all $\omega \in \Omega \setminus N_\gamma$ we have that $Z_s(\omega) = Y_s(\omega)$. Since $D_{\mathbb{Q}}$ is dense then, for any $t \in \mathbb{R}^+$, there is a sequence $s_n \in D_{\mathbb{Q}}$ which converges to $t$. Then for all $\omega \in \Omega \setminus N_\gamma$ we have as a consequence of (15) that $Y_{s_n}(\omega) \to Z_t(\omega)$. Consider, $\mathbb{P}(N_\gamma) = 0$ so that $Y_{s_n}$ converges to $Z_t$ almost surely. Since $Y_{s_n} \to Z_t$ almost surely and $\mathbb{P}$ is a probability measure it follows that $Y_{s_n}$ converges in measure to $Z_t$. On the other hand, for $\eta > 0$, by Markov's inequality and (6),

$$\mathbb{P}(\{\omega : \left|Y_{s_n}(\omega) - Y_t(\omega)\right| \geq \eta\}) \;\leq\; \eta^{-\alpha}\, E\big(|Y_{s_n} - Y_t|^\alpha\big) \;\leq\; \eta^{-\alpha} c\, |s_n - t|^{1+\beta},$$

and as this holds for each $\eta > 0$, we obtain that $Y_{s_n}$ converges in measure to $Y_t$. Hence, the limits $Z_t$ and $Y_t$ are equal as equivalence classes of measurable functions $\Omega \to \mathbb{R}^d$. That is,

$$\mathbb{P}(\{\omega : Z_t(\omega) = Y_t(\omega)\}) = 1.$$

As this holds for every $t \in \mathbb{R}^+$, this gives us that $(Z_t)_{t\in\mathbb{R}^+}$ is a modification of $(Y_t)_{t\in\mathbb{R}^+}$, completing the proof. $\square$

**2.1.3. Gaussian Process.** Being one of the most widely used distributions, the Gaussian distribution is central to probability and in this short section we shall build up to a process which is deeply connected with this distribution, the Gaussian Process.

Definition 2.10 (Gaussian Random Variable). *A random variable $Y$ is said to be a Gaussian random variable if its probability density function is given by*

$$h_Y(y) = \frac{1}{\sqrt{2\pi\,\mathrm{Var}(Y)}} \exp\left(-\frac{(y - \mathbb{E}[Y])^2}{2\,\mathrm{Var}(Y)^2}\right), \quad y \in \mathbb{R}.$$

Remark 2.2. *In the above definition, $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$ are parameters of the distributions. These are generally denoted by $\mu$ and $\sigma$ however we elected for a slight abuse of notation with $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$ so as to avoid conflict of notation.*

We state without proof the following elementary result related to Gaussian random variables.

Theorem 2.4. *If $Y_1$ and $Y_2$ are two Gaussian random variables then for any $\alpha \in \mathbb{R}$, $Y_1 + \alpha Y_2$ is also Gaussian.*

We may then generalize this notion by considering Gaussian random vectors.

DEFINITION 2.11. *A vector of random variables* $\underline{Y} = (Y_1, Y_2, \cdots, Y_n)$ *is said to be an* $n$ *dimensional Gaussian random vector if its probability density function is given by*

$$h_{\underline{Y}}(\underline{y}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(\underline{y} - \mathbb{E}[\underline{Y}])^\top \Sigma^{-1}(\underline{y} - \mathbb{E}[\underline{Y}])\right), \quad \underline{y} \in \mathbb{R}^n. \tag{16}$$

*where* $\mathbb{E}[\underline{Y}] = (\mathbb{E}[Y_1], \mathbb{E}[Y_2], \cdots, \mathbb{E}[Y_n])$ *is the mean vector and* $\Sigma_{ij} = \mathrm{Cov}(Y_i, Y_j)$ *where* $\Sigma_{ij}$ *represents the* $ij^{th}$ *entry of* $\Sigma$.

$\Sigma$ in the above is said to be the variance-covariance matrix. We prove the following results which shall be on use in a later section. This result is due [8]

THEOREM 2.5. *Any variance-covariance matrix is symmetric and positive semi-definite.*

PROOF. Let $\Sigma$ be a $d \times d$ matrix and $\underline{v}$ be a vector in $\mathbb{R}^d$. Then

$$\begin{aligned}
\underline{v}^T \Sigma \underline{v} &= \underline{v}^T \mathbb{E}[(\underline{Y} - \mathbb{E}[\underline{Y}])(\underline{Y} - \mathbb{E}[\underline{Y}])^T]\underline{v} \\
&= \mathbb{E}[\underline{a}^T \underline{a}] \\
&= \mathbb{E}[\|\underline{a}\|^2] \geq 0.
\end{aligned}$$

As symmetry follows trivially from the definition, the proof is complete. $\square$

Next we wish to extend the idea above to the case when we have an infinite collection of Gaussian random variables indexed across time. Clearly, we cannot give a distribution in a closed form as in (16). Instead, due to Kolmogorov's Extension Theorem, we may define the finite dimensional distributions and we are guaranteed that a stochastic process with such finite dimensional distributions exists. Applying this reasoning for the specific case when the finite dimensional distributions are Gaussian, we obtain the following definition.

DEFINITION 2.12 (Gaussian Process). *A stochastic process* $\{Y_t : t \in T\}$ *with mean function* $m_Y(y)$ *and covariance function* $k_Y(x, y)$ *is said to be a Gaussian Process if any finite linear combination of its random variables has a multivariate Gaussian Distribution. The variables of the process must satisfy* $\mathbb{E}[Y_t] = m_Y(t)$ *and* $\mathrm{Cov}(Y_t, Y_s) = k_Y(t, s)$ *for* $t, s \in T$.

REMARK 2.3. *So as to have a valid covariance function, we require the function $k_Y(x, y)$ in Definition 2.12 to be positive semi-definite.*

Gaussian processes are highly flexible models, making them well suited for modelling complex phenomena such as non-linear dynamic systems. For example, they have been used in ionospheric physics to model total electron content over South Africa. We refer the interested reader to [2] for further details.

**2.1.4. Brownian Motion.** One of the most fundamental processes in probability theory, Brownian Motion shall be central when we define the stochastic integral as it will serve as the integrator. In this brief section we shall define Brownian Motion as well as the closely related Weiner measure. For this section, we shall use without proof the easily verifiable fact that $\min(s, t)$ is a valid covariance function for a Gaussian Process.

For any finite collection of times $0 \leq t_1 \leq \cdots, t_n \leq T$, let the Gaussian random variable vector $(Y_{t_1}, \cdots, Y_{t_n})$ satisfy $\mathbb{E}[Y_{t_i}] = 0$ and $\text{Cov}(Y_{t_i}, Y_{t_j}) = \min(i, j)$ for $i, j \in \{1, 2, \cdots, n\}$. For every finite collection of times, define the measures $\nu_{t_1,\cdots,t_n}$ as in (1). It is easily verified that the consistency conditions hold and so we are guaranteed the existence of a measure $\nu_{\mathbb{W}}$ on $\mathbb{R}^T$ which satisfies the conditions of the Kolmogorov Extension Theorem.

DEFINITION 2.13 (The Weiner Measure). *The measure $\nu_{\mathbb{W}}$ is said to be the Weiner measure.*

DEFINITION 2.14. *Let $B = (B_t)_{t\in[0,T]}$ be a random variable $B : \Omega \to \mathbb{R}^T$. Then $B$ is said to be a Brownian Motion if for all Borel sets $A \subseteq \mathbb{R}^T$ the probability of the set $\{\omega : B(\omega) \in A\}$ is given by $\nu_{\mathbb{W}}(A)$.*

The above however, does not guarantee that sample paths of Brownian Motion are almost surely continuous. This is undesirable both from a mathematical standpoint seeing as continuous functions give us much more results that we may work with. At this juncture, we may however utilize the Kolmogorov Continuity Theorem to ensure this. Prior to proving these however, we must show a few elementary results.

THEOREM 2.6. *Let $(B_t)_{t\in\mathbb{R}^+}$ be a Brownian Motion Process. Then $\mathbb{E}[B_t - B_s] = 0$ and $\mathrm{Var}(B_t - B_s) = |s-t|$.*

PROOF. We assume without loss of generality that $s \leq t$.That $\mathbb{E}[B_t - B_s] = 0$ follows trivially by linearity of the Expectation.

$$\begin{aligned}
\mathrm{Var}(B_t - B_s) &= \mathbb{E}[(B_t - B_s)^2] \\
&= \mathbb{E}[B_t^2] + \mathbb{E}[B_s^2] - 2\mathbb{E}[B_t B_s] \\
&= t + s - 2s \\
&= t - s
\end{aligned}$$

A similar case holds for $t < s$. □

We also state the following without proof.

THEOREM 2.7. *Let $Y_1$ be a Gaussian variable with mean 0 and variance $\mathrm{Var}(Y_1)$ and let $Y_2$ be a Gaussian variable with mean 0 and variance 1 then $Y_1$ and $\mathrm{Var}(Y_1)^{\frac{1}{2}} Y_2$ have the same distribution (and hence the same expectation).*

THEOREM 2.8. *Brownian Motion satisfies the conditions of Kolmogorov Continuity Theorem.*

PROOF. Let $\alpha > 2$ and $Y'$ denote a Gaussian random variable with mean 0 and variance 1. By Theorem 2.7 the following holds.

$$\begin{aligned}
\mathbb{E}[|Y_t - Y_s|^\alpha] &= |s-t|^{\frac{\alpha}{2}} \mathbb{E}[|Y'|^\alpha] \\
&= C_\alpha |s-t|^{\frac{\alpha}{2}}
\end{aligned}$$

where

$$C_\alpha = \frac{2^{\frac{\alpha}{2}}}{\sqrt{\pi}} \Gamma(\frac{\alpha+1}{2}) = \int_{-\infty}^{\infty} |y|^\alpha \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) = \mathbb{E}[|Y'|^\alpha].$$

Hence picking $\beta = 1 - \frac{\alpha}{2}$ and $C = C_\alpha$ in (6), we are guaranteed to find a continuous modification of Brownian Motion which with slight abuse of notation we shall denote by $(B_t)_{t\in\mathbb{R}^+}$. □

To end this section, we shall state a lemma which may be easily verified by calculations similar to those of the proof of Theorem 2.6.

LEMMA 2.2. *Brownian Motion increment have independent increments that is* $\mathrm{Cov}(B_{t_1} - B_{t_2}, B_{t_3} - B_{t_4})$ $=\mathbb{E}[(B_{t_1} - B_{t_2})(B_{t_3} - B_{t_4})] = 0$.

### 2.2. Stochastic Integration

While the integration of deterministic functions is known, it was desired to extend this notion to integrals whereby both the integrand and the integrator are random. We shall show how such an integral is constructed. For our purposes, we shall focus on integrals of the following form

$$\int_a^b Y_t \, dB_t \tag{17}$$

whereby $B_t$ is a Brownian motion process and $(Y_t)_{t\in\mathbb{R}^+}$ is a stochastic process. It should be noted that, even when $Y_t$ is deterministic function, the problem of defining the integral in (17) is non trivial. Indeed, for the case when $f$ is deterministic, one may attempt to define (17) through the use of Riemann-Stieltjes integration however this approach cannot be implemented here seeing as $B_t$ has infinite finite variation.

We start the discussion by defining a property which the process $(Y_t)_{t\in\mathbb{R}^+}$ must satisfy for it to be stochastically integrable.

DEFINITION 2.15 (Filtration). *A filtration* $\{\mathcal{F}_t : t \in T\}$ *is an increasing sequence of sigma algebras having some indexing set* $T$.

DEFINITION 2.16 (Adapted Process). *A stochastic process* $\{Y_t : t \in T\}$ *is said to be adapted to the filtration* $\{\mathcal{F}_t : t \in T\}$ *if* $Y_t$ *is* $\mathcal{F}_t$ *measurable for every* $t \in T$.

We are now in a position to define the stochastic integral. We shall fix a Brownian motion $B_t$ and a filtration $\{\mathcal{F}_t : a \leq t \leq b\}$ in such a way that the following conditions are satisfied.

(i) For each $t$, $B_t$ is $\mathcal{F}_t$ measurable.
(ii) For any $s \leq t$ the random variable $B_t - B_s$ is independent of the $\sigma$-algebra $\mathcal{F}_s$.

When dealing with the stochastic integral, we shall only be working with adapted processes. Due to this we shall define the space $L^2_{ad}([a,b]\times\Omega)$ as the space of all stochastic processes $\{Y_t : a \leq t \leq b\}$ which satisfy the following conditions

(i) $Y_t$ is adapted to the filtration $\{\mathcal{F}_t : a \leq t \leq b\}$.

(ii) $\int_a^b \mathbb{E}[|Y_t|^2]\ dt < \infty$.

REMARK 2.4. *For ease of notion. We shall write $Y_t$ as $Y(t)$ for this section.*

**2.2.1. Integral for Step Stochastic Processes.** Just as in the case of the Lebesgue integral, we first define the stochastic integral for step stochastic processes in $L^2_{ad}([a,b]\times\Omega)$. To this effect we let

$$Y(t,\omega) = \sum_{i=1}^{n} \xi_{i-1}(\omega)\mathbb{1}_{[t_{i-1},t_i)}(t)$$

where $\xi_{i-1}$ is $\mathcal{F}_{t_{i-1}}$ measurable and $\mathbb{E}[\xi^2_{i-1}] < \infty$. For these simple processes, we define the integral to be

$$I(Y) = \sum_{i=1}^{n} \xi_{i-1}(B_{t_i} - B_{t_{i-1}}). \tag{18}$$

Linearity of the integral is easily verified. The next result shall require some elementary results related to conditional expectation. We direct readers unfamiliar with this concept to Appendix A.

LEMMA 2.3. *Let $I(Y)$ be as defined in (18). Then $\mathbb{E}[I(Y)] = 0$ and*

$$\mathbb{E}[|I(Y)|^2] = \int_a^b \mathbb{E}[|Y(t)|^2]\ dt$$

*where the expectation is taken with respect to the Weiner measure.*

PROOF.

$$\begin{aligned}
\mathbb{E}[\xi_{i-1}(B_{t_i} - B_{t_{i-1}})] &= \mathbb{E}[\mathbb{E}[\xi_{i-1}(B_{t_i} - B_{t_{i-1}})|\mathcal{F}_{t_{i-1}}]] \\
&= \mathbb{E}[\xi_{i-1}\,\mathbb{E}[(B_{t_i} - B_{t_{i-1}})|\mathcal{F}_{t_{i-1}}]] \\
&= \mathbb{E}[\xi_{i-1}\,\mathbb{E}[(B_{t_i} - B_{t_{i-1}})]] \\
&= 0.
\end{aligned} \tag{19}$$

Here, (19) follows from the fact that $B_{t_i} - B_{t_{i-1}}$ is independent of $\mathcal{F}_{t_{i-1}}$. As a consequence of this it is easily verified that $\mathbb{E}[I(Y)] = 0$. Moreover

$$|I(Y)|^2 = \sum_{i,j=1}^{n} \xi_{i-1}\,\xi_{j-1}(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}}).$$

Consider $i \neq j$. Without loss of generality we take $i < j$. Then the following holds

$$\begin{aligned}
&\mathbb{E}[\,\xi_{i-1}\,\xi_{j-1}(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})] \\
&\quad = \mathbb{E}[\mathbb{E}[\xi_{i-1}\,\xi_{j-1}(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})|\mathcal{F}_{t_{j-1}}]] \\
&\quad = \mathbb{E}[\xi_{i-1}\,\xi_{j-1}(B_{t_i} - B_{t_{i-1}})\mathbb{E}[(B_{t_j} - B_{t_{j-1}})|\mathcal{F}_{t_{j-1}}]] \qquad (20)\\
&\quad = 0. \qquad (21)
\end{aligned}$$

Whereby (20) once again follows due to the independence of the increment from the filtration.

Consider $i = j$. Then

$$\begin{aligned}
\mathbb{E}[\xi_{i-1}^2(B_{t_i} - B_{t_{i-1}})^2] &= \mathbb{E}[\mathbb{E}[\xi_{i-1}^2(B_{t_i} - B_{t_{i-1}})^2|\mathcal{F}_{t_{i-1}}]] \\
&= \mathbb{E}[\xi_{i-1}^2\,\mathbb{E}[(B_{t_i} - B_{t_{i-1}})^2]] \\
&= \mathbb{E}[\xi_{i-1}^2(t_i - t_{i-1})] \\
&= (t_i - t_{i-1})\mathbb{E}[\xi_{i-1}^2]. \qquad (22)
\end{aligned}$$

Then, the result follows as a result of (21) and (22). □

**2.2.2. Approximation Lemma.** In this section, we shall show an approximation lemma which will be crucial when we extend from step stochastic processes to general ones.

THEOREM 2.9. *Let $Y \in L^2_{ad}([a,b] \times \Omega)$. Then there exists a sequence $\{Y_n(t) : n \geq 1\}$ of step stochastic processes in $L^2_{ad}([a,b] \times \Omega)$ such that*

$$\lim_{n\to\infty} \int_a^b \mathbb{E}[|(Y(t) - Y_n(t))|^2]\; dt = 0. \qquad (23)$$

*Here the convergence is in $(L^2([a,b]\times\Omega), \mathcal{B}([a,b])\times\mathcal{F}, \lambda\times\mathbb{P})$ where $\lambda\times\mathbb{P}$ denotes the measure arising from the product of the Lebesgue measure with a probability measure.*

Proof. We split the proof in three cases.

Case 1. $\mathbb{E}[Y_tY_s]$ *is a continuous function of* $(t,s) \in [a,b]^2$.

We let $\Delta_n = \{t_0, t_1..., t_n\}$ be a partition of $[a,b]$ and we define the stochastic process $Y_n(t,\omega)$ by taking

$$Y_n(t,\omega) = Y(t_{i-1},\omega),\ t_{i-1} < t \leq t_i \tag{24}$$

Clearly, $\{Y_n(t,\omega)\}_{n=1}^{\infty}$ is a sequence of adapted step stochastic process. By assumption of continuity we have that

$$\lim_{s\to t} \mathbb{E}[|Y(t) - Y(s)|^2] = 0,$$

in particular we have that for $t \in [a,b]$

$$\lim_{s\to t} \mathbb{E}[|Y(t) - Y(s)|^2] = 0. \tag{25}$$

Using the inequality $|\alpha - \beta|^2 \leq 2(|\alpha|^2 + |\beta|^2)$ we obtain that

$$|Y(t) - Y_n(t)|^2 \leq 2(|Y_t|^2 + |Y_n(t)|^2).$$

Consequently for all $a \leq t \leq b$,

$$\begin{aligned}\mathbb{E}[|Y(t) - Y_n(t)|^2] &\leq 2(\mathbb{E}[|Y(t)|^2] + \mathbb{E}[|Y_n(t)|^2]) \\ &\leq 4 \sup_{a\leq s\leq b} \mathbb{E}[|Y(s)|^2]. \end{aligned} \tag{26}$$

Thus from (25) and (26) we are justified in applying the Lebesgue Dominated Convergence Theorem to obtain the following

$$\lim_{n\to\infty} \int_a^b \mathbb{E}[|Y(t) - Y_n(t)|^2]\, dt = 0.$$

Hence verifying the lemma in the first case.

Case 2. $Y$ *is bounded.*

In this case, we define a stochastic process $Z^n$ by taking

$$Z_n(t,\omega) = \int_0^{n(t-a)} e^{-\tau} Y(t - \frac{\tau}{n}, \omega)\, d\tau.$$

We trivially have that $\{Z_n(t)\}_{t\in\mathbb{R}^+}$ is adapted to $\mathcal{F}_t$ and $\int_a^b \mathbb{E}[|Z_n(t)|^2]\, dt < \infty$.

First we show that, for each $n$, $\mathbb{E}[Z_n(t)Z_n(s)]$ is a continuous function of $(t,s)$. To show this we perform the substitution $u = t - \frac{\tau}{n}$ which gives

$$Z_n(t,\omega) = \int_a^t ne^{-n(t-u)}Y(u,\omega)\ du.$$

This may be trivially used to verify that

$$\lim_{t\to s}\mathbb{E}[|Z_n(t) - Z_n(s)|^2] = 0$$

so that the claim follows.

Next we shall show that

$$\lim_{n\to\infty}\int_a^b \mathbb{E}[|Y(t) - Z_n(t)|^2]\ dt \to 0 \tag{27}$$

To show this, we first note that

$$Y(t) - Z_n(t) = \int_0^\infty e^{-\tau}(Y(t) - Y(t-\frac{\tau}{n})\ d\tau,$$

where $Y(t)$ is understood to be 0 for $t < a$. As $e^{-\tau}d\tau$ is a probability measure on $[0,\infty)$, we are justified in applying the Schwartz inequality to obtain

$$|Y(t) - Z_n(t)|^2 \leq \int_0^\infty |Y(t) - Y(t-\frac{\tau}{n})|^2 e^{-\tau}d\tau.$$

From this we proceed to conclude the following

$$\begin{aligned}
&\int_a^b \mathbb{E}[|Y(t) - Z_n(t)|^2]\ dt \\
&\qquad \leq \int_a^b\int_0^\infty e^{-\tau}\mathbb{E}[|Y(t) - Y(t-\frac{\tau}{n})|^2]d\tau dt \\
&\qquad = \int_0^\infty e^{-\tau}\Big(\int_a^b \mathbb{E}[|Y(t) - Y(t-\frac{\tau}{n})|^2]dt\Big)d\tau \\
&\qquad = \int_0^\infty e^{-\tau}\mathbb{E}\Big[\int_a^b |Y(t) - Y(t-\frac{\tau}{n})|^2 dt\Big]d\tau. \qquad (28)
\end{aligned}$$

Due to the assumption of boundedness of $Y$ we have that

$$\lim_{n\to\infty}\int_a^b |Y(t,.) - Y(t-\frac{\tau}{n},.)|^2 dt = 0 \tag{29}$$

almost surely. Thus (27) follows as a consequence of (28) and (29).

Now using the fact that $\mathbb{E}[Z_n(t)Z_n(s)]$ is a continuous function of $(t,s)$, we apply

Case 1 to $Z_n$ for each $n$ to choose an adapted stochastic process $Z_n(t,\omega)$ such that

$$\int_a^b \mathbb{E}[|Z_n(t) - Y_n(t)|^2]dt \leq \frac{1}{n}. \tag{30}$$

Hence by (27) and (30) we have that

$$\lim_{n\to\infty} \int_a^b \mathbb{E}[|Y(t) - Y_n(t)|^2]\ dt = 0,$$

completing the proof for the second case.

Case 3. *The general case for $Y \in L^2_{ad}([a,b]\times\Omega)$.*

Let $Y \in L^2_{ad}([a,b]\times\Omega)$. For each $n$, define

$$Z_n(t,\omega) = \begin{cases} Y(t,\omega) & \text{if } |Y(t,\omega)| \leq n, \\ 0 & \text{if } |Y(t,\omega)| > n. \end{cases}$$

Then by the Lebesgue dominated convergence theorem,

$$\lim_{n\to\infty} \int_a^b \mathbb{E}[|Y(t) - Z_n(t)|^2]\, dt = 0. \tag{31}$$

Now, for each $n$ we apply Case 2 to $Z_n$ to choose an adapted step stochastic process $Z_n(t,\omega)$ satisfying

$$\int_a^b \mathbb{E}\big(|Z_n(t) - Y_n(t)|^2\big)\, dt\ \leq\ \frac{1}{n}. \tag{32}$$

Hence (23) follows from (31) and (32) completing the proof. □

**2.2.3. Integral for General Stochastic Processes.** From what was established in 2.2.1 and 2.2.2 we may proceed to define the stochastic integral

$$\int_a^b Y(t)\ dB_t \quad Y \in L^2_{ad}([a,b]\times\Omega)$$

By applying the approximation lemma proved in 2.2.2 we may obtain a sequence $\{Y_n(t,\omega) : n \geq 1\}$ of adapted step stochastic processes such that (23) holds. For each $n$, $I(Y_n)$ is defined as in 2.2.1. By Lemma 2.3 the following is evident

$$\begin{aligned} \mathbb{E}[|I(Y_n) - I(Y_m)|^2] &= \int_a^b \mathbb{E}[|Y_n(t) - Y_m(t)|^2]\ dt \\ &\to 0 \quad \text{as } m,n\to\infty. \end{aligned} \tag{33}$$

Hence the sequence $\{I(Y_n)\}$ is Cauchy in $L^2(\Omega)$ and we define

$$I(Y) = \lim_{n\to\infty} I(Y_n) \quad \text{in } L^2(\Omega). \tag{34}$$

Definition 2.17. *The integral $I(Y)$ in (34) is said to be the stochastic integral of $Y$.*

It is trivially verified that the integral is linear. Also, by the discussion in 2.2.2 we have that Lemma 2.3 remains valid for any adapted stochastic process. We state this as a Theorem.

Theorem 2.10. *Let $Y \in L^2_{ad}([a,b]\times\Omega)$. Then the Itô integral $I(Y) = \int_a^b Y(t)\, dB_t$ is a random variable with $\mathbb{E}[I(Y)] = 0$ and*

$$\mathbb{E}[|I(Y)|^2] = \int_a^b \mathbb{E}[|Y(t)|^2]\, dt.$$

Hence by this theorem we have that $I : L^2_{ad}([a,b]\times\Omega) \to L^2(\Omega)$ is an isometry and this is indeed the well known Itô Isometry in stochastic integration.

CHAPTER 3

# Foundations of Quantum Probability

## 3.1. Elementary Notions of Quantum Probability

As previously mentioned, the relaxation of the assumption of commutativity of random variables necessitated new developments of probability. Our main focus shall be quantum probability and in this chapter we shall present elementary definitions related to the subject and prove Gleason's theorem, a crucial result in this area. The theorems and definitions related to quantum mechanics are taken from [19] unless otherwise specified.

In classical probability we work with measure spaces $(\Omega, \mathcal{F})$ where $\Omega$ is a set and $\mathcal{F}$ is a $\sigma$-algebra on $\Omega$. For quantum probability, this is no longer sufficient and instead we consider the space $(\mathcal{H}, \mathcal{A})$ whereby $\mathcal{H}$ is a separable Hilbert Space and $\mathcal{A}$ is the lattice of closed subspaces of $\mathcal{H}$. For the sake of completeness, we define the notion of a lattice which we take from [7].

Definition 3.1 (Lattice). *A poset $P$ is said to be a lattice if for any two elements $p_1, p_2 \in P$ there exists a greatest lower bound $p_1 \wedge p_2$ called the meet and a lower upper bound $p_1 \vee p_2$ called the join.*

In quantum probability, the role of random variables is carried out by observables which we define below. The definitions and results relating to observables in this chapter are found in [19].

Definition 3.2 (Observables). *Let $\Omega$ be a set, $\mathcal{F}$ be a $\sigma$-algebra over $\Omega$ and $\mathcal{P}(\mathcal{H})$ denote the set of orthogonal projections over a Hilbert Space. Then an observable is a spectral measure, that is a mapping $\pi : \mathcal{F} \to \mathcal{P}(\mathcal{H})$ satisfying:*

(i) $\pi(\emptyset) = \mathbf{0}$ *and* $\pi(\Omega) = I$*.*

(ii) *For a family of disjoint sets* $\{E_i\}_i$ *in* $\mathcal{F}$,

$$\pi(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \pi(E_i).$$

REMARK 3.1. *By using techniques similar to those used when working with classical measures, we may show that any spectral measure* $\pi$ *satisfies the following*

(i) $E_1 \subseteq E_2 \implies \pi(E_1) \leq \pi(E_2)$.

(ii) *For any* $E_1$ *and* $E_2$ *in* $\mathcal{F}$ :

$$\pi(E_1 \cup E_2) + \pi(E_1 \cap E_2) = \pi(E_1) + \pi(E_2).$$

*Then, by observing that* $\pi(E_1)\pi(E_1 \cup E_2) = \pi(E_1)$, *we may also conclude that the following holds*

$$\pi(E_1 \cap E_2) = \pi(E_1)\pi(E_2).$$

We may now proceed to discussing measures in the context of quantum probability.

**3.1.1. Measures.** It is clear that for a measure on the space $(\mathcal{H}, \mathcal{A})$ some adjustments to Definition 2.5 must be made. For example, it does not make sense to discuss disjointness seeing as the zero vector must necessarily be an element of every closed subspace. Instead of disjointness we shall use orthogonality. We are now in a position to give a formal definition of a quantum measure.

DEFINITION 3.3 (Quantum Probability Measure). *Consider the space* $(\mathcal{H}, \mathcal{A})$. *A quantum measure* $\mu$ *is a function defined on* $\mathcal{A}$ *which satisfies the following properties.*

(i) *For all* $A \in \mathcal{A}, \mu(A) \geq 0$.

(ii) $\mu(\boldsymbol{0}) = 0$. *Where* $\boldsymbol{0}$ *denotes the null projection.*

(iii) *For all countable pairwise orthogonal collections* $\{A_i\}_{i=1}^{\infty}$ *satisfying* $A_i \in \mathcal{A}$ *for all* $i$,

$$\mu\left(\bigvee_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

*whereby* $\vee$ *denotes the usual supremum induced by the lattice of projections.*

(iv) $\mu(I) = 1$.

Classical measures described in 2.5 may be easily constructed with several probability distributions being widely known and utilized. We would like to know if measures in a quantum probability setting assume a particular form and this is precisely what we will find in the next section.

### 3.2. Gleason's Theorem

Being one of the cornerstones of quantum probability, Gleason's Theorem shall give us that the probability of some closed subspace may be given as a trace of some operator. Before proving this, we shall need a few definitions and lemmas which we give below. The original result is found in [9], we shall follow the treatment given by [10].

Definition 3.4 (Unit Sphere of a Hilbert Space). *The unit sphere $S$ of $\mathcal{H}$ is the set $\{h \in \mathcal{H} : \|h\| = 1\}$ where $\|h\| = \sqrt{\langle h, h \rangle}$.*

Definition 3.5 (Great Circle). *A Great Circle is a circle formed on the surface of the sphere by intersecting a plane passing through the centre of the sphere with the sphere.*

Definition 3.6 (Frame Function). *A real-valued function $f$ defined on the unit sphere of $\mathcal{H}$ is said to be a frame function of weight $W$ if, for every orthonormal basis $\{x_i\}_i$ of $\mathcal{H}$, we have*

$$\sum_i f(x_i) = W.$$

Remark 3.2. *If $\{x_1, x_2, ...\}$ is an orthonormal basis then $\{\alpha x_1, \alpha x_2, ...\}$ is also an orthogonal basis. Thus, for a constants $\{\alpha_1, \alpha_2 \cdots\}$ satisfying $|\alpha_i| = 1$ for all $i$, a frame function $f$ must satisfy*

$$\sum_i f(\alpha x_i) = W. \tag{35}$$

Definition 3.7 (Regular Frame Function). *A frame function $f$ is regular if there exists a self-adjoint operator $\rho$ defined on $\mathcal{H}$ such that for all unit vectors*

*$x \in \mathcal{H}$ we have*

$$f(x) = x^T \rho x.$$

From this definition the following lemma is immediately clear.

LEMMA 3.1. *In a finite dimensional real Hilbert space, a frame function is regular if and only if it is the restriction of a quadratic form to the unit sphere.*

The next lemma, which we state without proof, shall be useful in a later theorem.

LEMMA 3.2. *The function on the unit circle in $\mathbb{R}^2$, $f(\theta) = \cos n\theta$ is a frame function if and only if $n = 0$ or $n = 2 \pmod 4$.*

In the next theorem, we shall start seeing sufficient conditions for a frame function to be regular which will later be used for another result.

THEOREM 3.1. *Every continuous frame function on the unit sphere in $\mathbb{R}^3$ is regular.*

PROOF. Let $Q_l$ denote the space of spherical harmonics $Y_{lm}$ of degree $l$. Let $F$ denote the set of continuous frame functions on the unit sphere $S$ in $\mathbb{R}^3$. It is known that every continuous function on $S$ may be expressed as a norm limit of a sum of $Y_{lm}$ s and we seek to find which of these $Y_{lm}$s may contribute to the sums of elements of $F$. Prior to this however, we recall that spherical harmonics may be written in terms of polar coordinates in the following manner:

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta) e^{im\phi}, \tag{36}$$

where $P_l^m$ denotes the associated Legendre polynomial of index $l, m \in \mathbb{N}$. Through (36) we may easily ascertain the following relations

$$Y_{lm}(\theta, \phi) = (-1)^{l+m} Y_{lm}(\pi - \theta, \phi) = (-1)^m Y_{lm}(\theta, \phi + \pi) = (-1)^l Y_{lm}(-\theta, -\phi). \tag{37}$$

$Q_0$ gives the space of constant functions and trivially, these are also frame functions. Hence $Q_0 \subseteq F$ meaning $Y_{00}$ may contribute to the sums of elements of $F$. $Q_1$ consists of linear functions on $\mathbb{R}^3$ restricted to $S$. By (37) these functions change sign under parity. Hence any function in $Q_1$ cannot lie in $F$ else its inclusion in $F$ would contradict (35). When considering $Q_2$, we need only note that this consists of quadratic forms of zero trace restricted to $S$. These are trivially frame functions and hence $Q_2 \subseteq F$.

Next we check if any $Q_l$ having $l > 2$ may be a subset of $F$. Suppose this is so. We recall that spherical harmonics satisfy the Laplace equation in $\mathbb{R}^3$ which, in cylindrical coordinates, is written in the following manner:

$$\frac{1}{r}\frac{\partial}{\partial r}(r\frac{\partial g}{\partial r}) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} + \frac{\partial^2 g}{\partial z^2} = 0. \tag{38}$$

Through substitution, we may easily see that $g_1 = r^l \cos l\theta$ and $g_2 = (r^2 - 2(l-1)z^2)r^{l-2}\cos((l-2)\theta)$ both satisfy (38). As these are of order $l$, $g_1$ and $g_2$ must be in $Q_l$ and consequently these may be written as a linear combination of $Y_{lm}$s.

Consider an orthonormal triple of vectors $\{x_1, x_2, x_3\}$ with any vector fixed which we, without loss of generality, take to be $x_1$. We let $f$ be a frame function defined on this set satisfying

$$f(x_1) + f(x_2) + f(x_3) = W.$$

For all orthogonal pairs of vectors $\{x_2, x_3\}$ lying in the plane orthogonal to $x_1$ we will have that

$$f(x_2) + f(x_3) = W - f(x_1),$$

hence giving that $f$ is a frame function on the great circle orthogonal to the fixed vector $x_1$. However, by Lemma 3.2 any function proportional to $\cos(n\theta)$ will only be a frame function if $n = 0$ or $n = 2 \pmod 4$ which,for $l > 2$, cannot be satisfied simultaneously by $l$ and $l-2$. This contradiction gives us that $Q_l \nsubseteq F$ for $l > 2$.

Consequently, we are justified in saying that $F$ is the closed linear span of $Q_0$ and $Q_2$ which, for this case, is equivalent to saying $F = Q_0 + Q_2$. Knowing this, we may proceed to show the statement of the theorem. As previously discussed functions of $Q_2$ are restrictions of a quadratic form to the unit sphere. Functions in $Q_0$ may also be seen as such by noting that, on $S$, $1 = x^2 + y^2 + z^2$

which is trivially a quadratic form. Hence, $F$ consists only of quadratic forms restricted to unit spheres and therefore, by Lemma 3.1 these are regular proving the theorem. $\square$

We shall now use the above result to show that all non-negative frame functions on $S$ are continuous and, due to the theorem we have just proven, this will mean that these frame functions are indeed regular. It is prudent at this juncture to define the notion of $\operatorname{osc}(f, X)$.

Definition 3.8. *Let $f$ be a bounded real-valued function defined on a set $X$. Then $\operatorname{osc}(f, X)$ is given by $\operatorname{osc}(f, X) = \sup_{x \in X} f(x) - \inf_{x \in X} f(x)$.*

The next lemma is highly related to great circles. We firstly observe that every point of the sphere lies on infinitely many great circles. Let $q$ be any point on the sphere and $\theta_q$ be the value of the polar angle $\theta$ at $q$. In the lemma, for some point $q$, we shall mainly be working with a great circle $C_q$ which has a tangent which coincides with that of a circle in the plane $\theta = \theta_q$. $C_q$ shall also be referred to as the Great Circle through $q$.

Lemma 3.3. *Let $N$ denote the set of points $n$ satisfying $\theta_n \leq \frac{\pi}{2}$. Let $p_3 \in N$ such that $p_3 \neq (0, 0, 1)$. Let $P_1$ be the set of all points $p_1 \in N$ such that for some point $p_2$ we have:*

*i) $p_2$ is on the Great Circle through $p_1$.*

*ii) $p_3$ is on the Great Circle through $p_2$.*

*Then the set $P_1$ has a non-empty interior.*

Proof. We shall show this lemma by deriving an expression for the set of points which satisfy *ii*). By suitable transformation of coordinates, we let $p_3$ have coordinates $(\sin\theta, 0, \cos\theta)$ in some orthonormal coordinate system and $p_2$ have coordinates $(a, b, c)$.

Consider the tangent $t_c$ at $p_2$ to the great circle connecting $p_2$ and $p_3$. Then $t_c$ is a vector in the plane in which $p_2$ and $p_3$ lie and it must also satisfy $t_c \perp p_2$.

From these conditions, we may find an expression for $t_c$ by taking

$$t_c = p_3 - \frac{\langle p_3, p_2 \rangle p_2}{\langle p_2, p_2 \rangle}.$$

We want to find the points $p_2$ such that this tangent coincides with the tangent $t_2$ of the circle of latitude given by $\theta = \theta_{p_2}$. It is clear that this tangent must lie on the $xy$-plane and it must be orthogonal to $p_2$. Hence, ignoring normalization, $t_2 = (-b, a, 0)$.

At $t_2$, $t_c$ and $t_2$ must be parallel giving us that there exists $\alpha \in \mathbb{R}$ such that $\alpha t_2 = t_c$. Equating the $z$ coordinates, the following equation is obtained

$$0 = (a^2 + b^2)\cos\theta - ac\sin\theta. \tag{39}$$

While two other equations may be derived, equation (39) is enough to specify the points $p_2$ which satisfy *ii*). Indeed, if the tangent of a great circle through $p_2$ lies in the $xy$-plane then it must necessarily be a tangent to the surface of the sphere in the $xy$-plane and hence parallel to the tangent of $C_{p_2}$. Hence the locus of points $L$ which satisfy *ii*) is given by

$$\iota = (a^2 + b^2)\cos\theta - ac\sin\theta = 0.$$

Let $x$ be any point for which $\iota > 0$. Then the great circle through $x$ must cross the locus $L$ since, on the equator, $\iota < 0$. Hence, $P_1$ contains at least the open set of points for which $\iota$ is positive and by taking the points $x = (\sin\kappa, 0, \cos\kappa)$ for $\theta < \kappa < \frac{\pi}{2}$ we verify that this open set is indeed non-empty. □

LEMMA 3.4. *Let $f$ be a frame function on the unit sphere $S$ in $\mathbb{R}^3$ such that for a certain neighbourhood $U$ of a point $p$ on $S$,* $\operatorname{osc}(f, U) = \alpha$*. Then, for every point on the great circle whose plane is orthogonal to $p$, we may find a neighbourhood $V$ satisfying* $\operatorname{osc}(f, V) \leq 2\alpha$*.*

PROOF. We choose a coordinate system such that $p = (0, 0, 1)$. We define the neighbourhood $U$ of $p$ by taking all points whose corresponding $\theta$ value satisfies $\theta < \vartheta$. Let $q_0$ be any point on the sphere having $\theta_{q_0} = \frac{\pi}{2}$ and $a$ be a point having the same $\phi$ coordinate as $q_0$ and having $\theta_a = \frac{\pi}{2} + \frac{\vartheta}{2}$ where $\theta_b$ denotes the value of $\theta$ for point $b$ using spherical polar coordinates.

Let $C_0$ be the unique great circle which connects $a$ and $q_0$ and let $a'$ and $q_0'$ be points in $N \cap C_0$ satisfying $a \perp a'$ and $q \perp q_0'$. Clearly, for both these points to lie in $U$ we must have $\vartheta \geq 0$ which is always satisfied and hence both $a'$ and $q_0'$ lie in $U$. The same argument holds if we substitute $q$, a point in some neighbourhood $V$ of $q_0$, for $q_0$ and hold $a$ fixed.

We repeat the above procedure for $q_1$ and $q_2$ in $V$. $C_i$ will be the great circle connecting $a$ and $q_i$ and choose points $a_i'$ and $q_i'$ in $N \cap C_i$ satisfying $a \perp a_i'$ and $q_i \perp q_i'$, $i \in \{1, 2\}$. Hence we will obtain two orthonormal bases $\{a, a_i'\}$ and $\{q, q_i'\}$ for $C_i$, $i \in \{1, 2\}$. Then, as $f$ is a frame function, we must have that

$$f(a) + f(a_i') = f(q_i) + f(q_i'),\ \ i \in \{1, 2\}.$$

Subtracting the equations we obtain by substituting $i = 1$ and $i = 2$, we obtain

$$\begin{aligned}|f(q_1) - f(q_2)| &= |f(a_1') - f(a_2') + f(q_2') - f(q_1')| \\ &\leq |f(a_1') - f(a_2')| + |f(q_2') - f(q_1')| \\ &\leq 2\alpha.\end{aligned}$$

In the last step, we used the fact that $\{a_1', a_2', q_1', q_2'\} \in U$. Hence, for any pair of points $q_1$ and $q_2$ both in $V$ $|f(q_1) - f(q_2)| \leq 2\alpha$ showing that $\operatorname{osc}(f, V) \leq 2\alpha$. $\square$

In the next lemma we generalize the result from a neighbourhood to any open set $U$.

Lemma 3.5. *Let $f$ be a frame function on the unit sphere $S$ in $\mathbb{R}^3$ such that for some non-empty open set $U$ we have $\operatorname{osc}(f, U) = \alpha$. Then every point of $S$ has a neighbourhood $W$ such that $\operatorname{osc}(f, W) \leq 4\alpha$.*

Proof. To prove this, we need only apply Lemma 3.4 twice. Recall that, from an arbitrary point in the sphere, we may reach any point in two steps of arc length $\frac{\pi}{2}$. Since $\operatorname{osc}(f, U) = \alpha$ where $U$ is the neighbourhood of some point $p$, there exists a point $q$ satisfying $q \perp p$ having a neighbourhood $V$ for which $\operatorname{osc}(f, V) \leq 2\alpha$. By the same reasoning, there is a neighbourhood $W$ of a point

$w$ orthogonal to $q$ for which $\mathrm{osc}(f, W) \leq 4\alpha$. As any point of $S$ may be reached in this manner, the lemma is proven. $\square$

Through these established lemmas, we shall now show that every non-negative frame function on $S$ in $\mathbb{R}^3$ is regular. In particular it is enough to show that these are continuous.

THEOREM 3.2. *Every non-negative frame function on $S$ in $\mathbb{R}^3$ is regular.*

PROOF. To start the proof, we first note an elementary property of frame functions. This being, that the addition of a constant to a frame function still results in a frame function albeit one with a different weight. Let $f$ be a non-negative frame function on $S$ having weight $W$. Due to the property mentioned, we may assume that $\inf_{x \in S} f(x) = 0$.

Let $\epsilon > 0$ and set $\eta = \frac{\epsilon}{88}$. As $\inf_{x \in S} f(x) = 0$ we may always find a point $p \in S$ satisfying $f(p) \leq \eta$. We denote by $\sigma$ the transformation, $\sigma x(\theta_x, \phi_x) = x(\theta_x, \phi_x + \frac{\pi}{2})$ and consequently define the function

$$g(x) = f(x) + f(\sigma x).$$

$\{x_1, x_2, x_3\}$ is an orthogonal triple if and only if $\{\sigma x_1, \sigma x_2, \sigma x_3\}$ is and hence

$$\sum_{i=1}^{3} g(x_i) = \sum_{i=1}^{3} (f(x_i) + f(\sigma x_i)) = 2W$$

showing $g$ is a frame function of weight $2W$.

We restrict our attention to vectors $q$ on the equator. For such a $q$, $q \perp \sigma q$ and in addition $p, q$ and $\sigma q$ are mutually orthogonal. From this we may deduce that $g$ satisfies

$$g(q) = f(q) + f(\sigma q) = f(q) + f(\sigma q) + f(p) - f(p) = W - f(p). \tag{40}$$

As this holds for any $q$ on the equator we may deduce that $g$ is constant on the equator.

We now take a point $r \in N \setminus \{p\}$ and let $C$ be the great circle through $r$

such that $r$ is the highest point on $C$. Due to this selection, the point at which $C$ intersects the equator will be orthogonal to $r$. We will call this point $p$. From this we may deduce that

$$g(r) + g(q) \leq 2W.$$

By (40) we have that $g(r)+g(q) = g(r)+W-f(p)$ and hence the we may deduce that

$$g(x) \leq W + f(p) \leq W + \eta \; \forall x \in N \setminus \{p\}$$

Now we take a point $s \in C \cap N$ and a point $t \in C \cap N$ orthogonal to $s$. As $s$ and $t$ span the plane of $C$ then so do $r$ and $q$. Consequently, the orthogonal vector $w$ orthogonal to the pairs is the same. Due to $g$ being a frame function, we have that $g(r) + g(q) + g(w) = g(s) + g(t) + g(w)$ and hence

$$g(r) + g(q) = g(r) + W - f(p) = g(s) + g(t) \leq g(s) + W + \eta.$$

From this we obtain that the following holds $\forall r \in N \setminus \{p\}$ and $s \in C \cap N$

$$g(r) \leq g(s) + 2\eta. \tag{41}$$

Let $\beta = \inf\{g(x) : x \in N \setminus \{p\}\}$ and we choose a point $z \in N \setminus \{p\}$ for which $g(z) \leq \beta + \eta$. This point is guaranteed to exist by definition of infimum. Next, we consider a vector $x$ which satisfies the conditions outlined by Lemma 3.3. This means that $x \in N \setminus \{p\}$ is such that for some $y$ we have

i) $y$ is on the Great Circle through $x$.

ii) $z$ is on the Great Circle through $y$.

By Lemma 3.3 we are guaranteed that such a point exists. By (41) we obtain that

$$g(x) \leq g(y) + 2\eta$$

as well as

$$g(y) \leq g(z) + 2\eta.$$

From these we may conclude that

$$\beta \leq g(x) \leq g(z) + 4\eta \leq \beta + 5\eta. \tag{42}$$

Hence, for the non-empty interior $U$ of the set $X$ composed of such vectors $x$, we have $\operatorname{osc}(g, U) \leq 5\eta$. By Lemma 3.5 there exists a neighbourhood $V$ of $p$ satisfying $\operatorname{osc}(g, V) \leq 20\eta$. Consider, $p = \sigma p$ and hence

$$g(p) = 2f(p) \leq 2\eta$$

from which we deduce that

$$\sup_{x \in V} g(x) \leq 22\eta. \tag{43}$$

We one again apply Lemma 3.5 to infer that every point $u$ on the sphere must have a neighbourhood $W$ such that

$$\operatorname{osc}(f(u), W) \leq 88\eta = \epsilon. \tag{44}$$

Continuity of $f$ on $S$ follows by noting that $\epsilon$ may be chosen to be arbitrarily small and by Theorem 3.1 we have that $f$ is regular as required. $\square$

Having proven the continuity of $f$, we now move on to the next part of the proof and to proceed we shall need the following definition.

Definition 3.9. *A real-linear subspace $\mathcal{H}'$ of a Hilbert space $\mathcal{H}$ is said to be completely real if the inner product on $\mathcal{H}'$ takes only real values on $\mathcal{H}' \times \mathcal{H}'$.*

We may immediately note that $S$ may be considered as a completely real subspace of a complex two-dimensional Hilbert space. Seeing as we have shown regularity of $f$ on $S$, we shall now investigate whether this has any implications of the regularity of $f$ on a larger Hilbert space.

Another observation we may make is that, by restriction of the inner product, a completely real subspace of any Hilbert space $\mathcal{H}$ is itself a Hilbert space. Consequently, a frame function for $\mathcal{H}$ remains a frame function when restricted to such a subspace. To start we shall prove a lemma related to frame functions on real Hilbert spaces.

Lemma 3.6. *If $f$ is a non-negative frame function of weight $W$ on a real Hilbert space, then for any unit vectors $x_1$ and $x_2$ we have*

$$|f(x_1) - f(x_2)| \leq 2W|x_1 - x_2|$$

PROOF. By regularity of $f$, we are guaranteed the existence of a self-adjoint operator $\rho$ satisfying $f(x) = x^T \rho x$. As $f$ is non-negative, $x^T \rho x \geq 0$ and due to the frame function condition we may additionally infer that $x^T \rho x \leq W$ for any unit vector $x$. Also, $|\rho| \leq W$. Since $\mathcal{H}$ is real, we must also have that $x_1^T \rho x_2 = x_2^T \rho x_1$. Then

$$(x_1 + x_2)^T \rho (x_1 - x_2) = x_1^T \rho x_1 - x_2^T \rho x_2 = f(x_1) - f(x_2). \tag{45}$$

Therefore we obtain

$$\begin{aligned} |f(x_1) - f(x_2)| &\leq |\rho||x_1 + x_2||x_1 - x_2| \\ &= |\rho|\sqrt{\langle x_1, x_1\rangle + 2\langle x_1, x_2\rangle + \langle x_2, x_2\rangle}|x_1 - x_2| \\ &\leq 2W|x_1 - x_2| \end{aligned}$$

as required. □

The next lemma has to do with the compactness of the unit sphere. We state this without proof.

LEMMA 3.7. *In a normed space $X$, the unit ball is compact if and only if $X$ is finite dimensional.*

LEMMA 3.8. *Let $f$ be a non-negative frame function on a two-dimensional complex Hilbert space $\mathcal{H}^2$ which is regular on every completely real subspace. Then $f$ is regular.*

PROOF. Let $f$ have weight $W$ and let $M = \sup_{x \in \mathcal{H}^2} f(x)$. Then we can choose a sequence of unit vectors $(x_n)_{n\in\mathbb{N}}$ such that $f(x_n) \to M$. By Lemma 3.7 and by using that any Hilbert space is a metric space, we may also arrange that $x_n \to y$. Construct $\lambda_n = \frac{\langle x_n, y\rangle}{|\langle x_n, y\rangle|}$. We may immediately note that $\langle \lambda_n x_n, y\rangle$ is real. By choice of $\lambda_n$ we have that $\lambda_n \to 1$ and $\lambda_n x_n \to y$.

Now we shall require a strengthening of the property in (35). In particular we will show that for any frame function $f$, unit vector $x$ and for any $\alpha$ satisfying $|\alpha| = 1$ we have

$$f(\alpha x) = f(x). \tag{46}$$

This is due to the fact that if $f$ is a frame function for a Hilbert space $\mathcal{H}$ then it is a frame function for any closed subspace $D$ of $\mathcal{H}$ by restriction. Hence by taking a one dimensional space $D$, the frame function condition gives (46). From this we may infer that $f(\lambda_n x_n) = f(x_n)$. We also have

$$\langle \lambda_n x_n, y\rangle = \frac{\langle y, x_n\rangle\langle x_n, y\rangle}{|\langle x_n, y\rangle|} = \frac{\langle y, x_n\rangle\langle y, x_n\rangle^*}{|\langle x_n, y\rangle|} \in \mathbb{R}$$

so $\lambda_n x_n$ and $y$ span a completely real subspace. We note that a closed, completely real subspace of a Hilbert space $\mathcal{H}$, by restriction of the inner product, is itself a Hilbert space. By use of Lemma 3.6 as well as (46) we conclude the following

$$\begin{aligned}
|f(y) - M| &= |f(y) - f(\lambda_n x_n) + f(x_n) - M| \\
&\leq |f(y) - f(\lambda_n x_n)| + |f(x_n) - M| \\
&\leq 2W|y - \lambda_n x_n| + |f(x_n) - M| \\
&< \epsilon
\end{aligned}$$

from which we conclude that $f(y) = M$.
To proceed we shall define the function $F$ on $\mathcal{H}$ as follows

$$F(v) := \begin{cases} 0 & \text{if } v = 0 \\ |v|^2 f(\frac{v}{|v|}) & \text{otherwise.} \end{cases} \tag{47}$$

Under the assumption that $f(x)$ is regular and hence may be written as $x^T\rho x$ on every completely real subspace, we have that $F$ is a quadratic form under restriction to any completely real subspace. In addition, by (46), we have that $F(\alpha v) = |\alpha|^2 F(v)$ for all scalars $\alpha$ and unit vectors $v$.

Let $z$ be a unit vector orthogonal to $y$ and we consider the completely real subspace spanned by $y$ and $z$. As $F$ restricts to be a frame function on this subspace, we have $F(y) = f(y) = M$ and $F(z) = W - f(y) = W - M$. On the $yz$-subspace, $F$ is a quadratic form by the assumption of regularity. By choice of $M$ as the supremum, $F$ attains its maximum value on the unit circle at $y$. Hence the matrix for $F$ relative to the basis $\{y, z\}$ is diagonal. For suppose not, we would obtain a larger value of $F$ at some point $\tilde{y}$ obtained by adding some component along the $z$ direction to $y$ contradicting $f(y) = M$. Hence, for $\alpha, \beta \in \mathbb{R}$,

$$F(\alpha y + \beta z) = \alpha^2 F(y) + \beta^2 F(z) = \alpha^2 M + \beta^2 (W - M). \tag{48}$$

Letting $\gamma$ and $\eta$ be non-zero complex numbers, we construct the vector

$$z' = \frac{\gamma}{|\gamma|}\frac{|\eta|}{\eta} z.$$

We may easily see that $z'$ will also be a unit vector orthogonal to $y$. Then by using (46) and (48) we have

$$F(\eta y + \gamma z) = F(\frac{|\eta|}{\eta}(\eta y + \gamma z)) = F(|\eta| y + |\gamma| z') = M|\eta|^2 + (W - M)|\gamma|^2.$$

As every vector in $\mathcal{H}$ may be expressed as a complex linear combination of $y$ and $z$ we have shown that for all $x \in \mathcal{H}$,

$$F(x) = x^T \rho x$$

where $\rho$ is the self adjoint operator with matrix representation

$$\begin{pmatrix} M & 0 \\ 0 & W - M \end{pmatrix}$$

relative to the basis $\{y, z\}$. Hence the proof is complete. $\square$

In the next lemma, we shall see that for regularity on the whole space, one need only have regularity on any two-dimensional subspace on $\mathcal{H}$. For this proof however, we shall require the following form of the Riesz Representation Theorem which we state without proof.

THEOREM 3.3 (Riesz Representation Theorem). *Let $\mathcal{H}_1, \mathcal{H}_2$ be two Hilbert spaces and let $h : \mathcal{H}_1 \times \mathcal{H}_2 \to k$ be a bounded sesquilinear form. Then $h$ has a representation $h(x, y) = x^T S' y$ where $S' : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded linear operator. We must also have that $S'$ is uniquely determined by $h$ and satisfies $|S'| = |h|$.*

LEMMA 3.9. *Let $f$ be a non-negative frame function on a Hilbert Space $\mathcal{H}$ which is regular when restricted to any two dimensional subspace of $\mathcal{H}$. Then $f$ is regular.*

PROOF. We recall the definition of function $F$ in (47). As $f$ is regular on any two dimensional subspace $Y$ then for a real or complex $\mathcal{H}$ space respectively there exists a bilinear and Hermitian form $A_Y$ respectively such that

$$F(x) = A_Y(x,x) \propto x^T \rho x$$

for all $x \in Y$. We define a form $A$ on all of $\mathcal{H} \times \mathcal{H}$ by taking

$$A(x,y) = A_Y(x,y) \tag{49}$$

for $x, y \in Y$. Such an assignment is justified due to the fact that any pair of vectors $x$ and $y$ may be considered as elements of the subspace they span. We will show that $A$ satisfies the required properties. The subspace $x$ and $y$ span will always be two dimensional except for the case when $x$ and $y$ are parallel meaning $y = \lambda x$ for some constant $\lambda$.

For such a case, we must have that the following holds

$$A_Y(x,y) = A_Y(x,\lambda x) = \lambda F(x)$$

irrespective of how $Y$ is chosen.

Through use of the polarization identity, the real or complex one depending on the nature of $\mathcal{H}$, we may write

$$\begin{aligned} A_Y(x,y) &= \frac{1}{4}\big(A_Y(x+y,x+y) - A_Y(x-y,x-y)\big) \\ &= \frac{1}{4}\big(F(x+y) - F(x-y)\big) \end{aligned} \tag{50}$$

for the real case and

$$\begin{aligned} A_Y(x,y) = \frac{1}{4}\big(&A_Y(x+y,x+y) - A_Y(x-y,x-y) \\ &+ iA_Y(x+iy,x+iy) - iA_Y(x-iy,x-iy)\big) \\ = \frac{1}{4}\big(&F(x+y) - F(x-y) + iF(x+iy) - iF(x-iy)\big) \end{aligned} \tag{51}$$

for the complex one.

Using the fact that $A_Y$ is linear in its second coordinate and conjugate linear in its first, we may easily verify using (49), (50) and (51) that

$$\begin{aligned} A(\alpha x, y) &= \alpha^* A(x, y) \\ A(x, y) &= A(y, x)^* \\ 4 \operatorname{Re} A(x, y) &= F(x + y) - F(x - y) \\ 2F(x) + 2F(y) &= F(x + y) + F(x - y). \end{aligned} \tag{52}$$

From the properties listed in (52) we may obtain that

$$\begin{aligned} 8 \operatorname{Re} A(x, z) + 8 \operatorname{Re} A(y, z) &= 2F(x + z) - 2F(x - z) + 2F(y + z) - 2F(y - z) \\ &= F(x + z + y + z) - F(x - z + y - z) \\ &+ F(x - y) - F(x - y) \\ &= 4 \operatorname{Re} A(x + y, 2z) = 8 \operatorname{Re} A(x + y, z). \end{aligned}$$

So that

$$\operatorname{Re} A(x, z) + \operatorname{Re} A(y, z) = \operatorname{Re} A(x + y, z).$$

Similarly, we may obtain that

$$\operatorname{Im} A(x, z) + \operatorname{Im} A(y, z) = \operatorname{Im} A(x + y, z).$$

From which we may conclude that

$$A(x, z) + A(y, z) = A(x + y, z). \tag{53}$$

From (53) and the first two properties listed in (52) we may conclude that $A$ is bilinear or Hermitian on $\mathcal{H} \times \mathcal{H}$.

For the next step of the proof we shall require to show that $A$ is bounded. To achieve this we let $x, y \in \mathcal{H}$ such that $|x| \leq 1$ and $|y| \leq 1$. We may always choose $\beta \in \mathbb{C}$ having $|\beta| = 1$ in such a way that $A(\beta x, y) \in \mathbb{R}$. From this we may conclude that

$$\begin{aligned} 4|A(x, y)| = 4A(\beta x, y) = 4 \operatorname{Re} A(\beta x, y) &= F(\beta x + y) - F(\beta x - y) \\ &\leq k(|\beta x + y|^2 + |\beta x - y|^2) = 2k(|\beta x|^2 + |y|^2) \\ &\leq 4k. \end{aligned}$$

This gives us that $A(x,y)$ is a bounded sesquilinear form.

Due to $A$ being such a form, we are justified in using 3.3 to guarantee the existence of a bounded self adjoint operator $\rho$ satisfying

$$A(x,y) = x^T \rho y \tag{54}$$

for all $x, y \in \mathcal{H}$. By noting that

$$f(x) = F(x) = A(x,x) = x^T \rho x$$

for all unit vectors $x \in \mathcal{H}$ the proof is concluded. □

At this juncture, we are now in a position to show that every non-negative frame function on a Hilbert space $\mathcal{H}^N$ of dimension $N \geq 3$ is regular.

THEOREM 3.4. *Every non-negative frame function on a Hilbert space having dimension at least three is regular.*

PROOF. We recall that a frame function on $\mathcal{H}$ remains a frame function when restricting to any completely real subspace of $\mathcal{H}$. As $\dim(\mathcal{H}) \geq 3$, any two dimensional completely real subspace may be embedded in a three dimensional one. Then, as this subspace will be isomorphic to $\mathbb{R}^3$, we may apply Theorem 3.2. From this, we conclude that any non-negative frame function $f$ is regular on every completely real two-dimensional subspace of $\mathcal{H}$. By Lemma 3.8 $f$ is then regular on all two dimensional subspaces and finally, by Lemma 3.9, we can conclude that is regular on the entirety of $\mathcal{H}$. □

Having proven the regularity of all frame functions on such Hilbert spaces, we are now in a position where we can now prove the main result of this section.

THEOREM 3.5 (Gleason's Theorem). *Let $\mu$ be a measure on the closed subsets of a separable Hilbert space $\mathcal{H}$ satisfying $\dim(\mathcal{H}) \geq 3$. Then there exists a positive, semi-definite, self-adjoint, trace-class operator $\rho$ such that for all closed subspaces $A$ of $\mathcal{H}$*

$$\mu(A) = \mathrm{Tr}(\rho P_A) \tag{55}$$

*where $P_A$ denotes the orthogonal projection of $\mathcal{H}$ onto $A$.*

Proof. First we consider the case of $C_x$, a one dimensional subspace spanned by the unit vector $x$. We may immediately see that $f(x) = \mu(C_x)$ defines a non-negative frame function $f$. By Theorem 3.4 we are guaranteed the existence of a self-adjoint operator $\rho$ for which the following holds

$$f(x) = \mu(C_x) = x^T \rho x.$$

As $x^T \rho x \geq 0$ for all unit vectors $x$, then $\rho$ must be positive semi-definite. For an orthonormal basis $\{x_i\}_i$ of $\mathcal{H}$, we must have that

$$\mu(\mathcal{H}) = \sum_i \mu(C_{x_i}) = \sum_i x_i^T \rho x_i.$$

As the sum on the right converges, we obtain that $\rho$ is in the trace class with $\operatorname{Tr} \rho = \mu(\mathcal{H})$.

Now we move on to any closed subspace $A$. If we let $A$ have an orthonormal basis $\{y_i\}_i$ then we can always extend this to a basis for $\mathcal{H}$ by adding vectors $\{z_j\}_j$ in such way that $\{y_i\}_i \cup \{z_j\}_j$ is an orthonormal basis for $\mathcal{H}$.

Taking $P_A$ to be the operator representing an orthogonal projection onto the subspace $A$, we must trivially have that $P_A y_i = y_i$ and $P_A z_j = 0$ for all $i$ and $j$. Hence we must have that

$$\begin{aligned}
\mu(A) &= \sum_i \mu(C_{y_i}) = \sum_i y_i^T \rho y_i \\
&= \sum_i (P_A y_i)^T \rho y_i + \sum_j (P_A z_j)^T \rho z_j \\
&= \operatorname{Tr}(\rho P_A)
\end{aligned}$$

$\square$

Through Gleason's Theorem, it is clear that there exists a correspondence between any quantum measure $\mu$ and the operator $\rho$ established by the theorem which we call the density operator. One may note that $\rho$ analogous to the probability density function in classical probability. Furthermore, as a result of the theorem, we must have that $\rho$ is of unit trace. Indeed, taking $A$ to be the entire space we obtain that $P_A = I$ so that $1 = \mu(A) = \operatorname{Tr}(\rho)$. Due to this, density operators are well-suited to represent states encountered in quantum mechanics. While the detailed reasoning behind this is beyond the scope of this dissertation, we note that, for this reason, we shall refer to $\rho$ as a state from here onward.

Hence in quantum probability, our main concern will be to determine the probability that an observable takes on a particular value, given that the system is in a specific state. We shall see an example later in Section 4.2. Quantum probability has numerous physical applications, one of the most prominent being its use in modelling qubits in quantum computing. This is beyond the scope of this thesis, however we refer the interested reader to [18] for further information on the topic.

CHAPTER 4

# Quantum Probability Spaces

To obtain an idea as to how quantum probability spaces are utilized, we shall first restrict ourselves to the finite case, that is when the observable is of finite dimension. As previously mentioned, with every state there is associated a probability space. Hence, for a particular state $\rho$ we consider the triple $(\Omega, \mu, \rho)$ and we shall call this a quantum probability space.

Firstly we note that, by Theorem 3.5, any measure on this space must be a trace class operator. The calculation of probability of a particular event in quantum probability is centred around the Spectral Theorem which we will devote the next section to proving. The proof for the Spectral Theorem is taken from [15].

## 4.1. Proving the Spectral Theorem

To start off we list two trivial properties of eigenvalues of Hermitian matrices.

THEOREM 4.1. *Let $G$ be a Hermitian matrix then*

*i) $G$ has real eigenvalues.*

*ii) The eigenvectors corresponding to differing eigenvalues are orthogonal.*

In the proof for the Spectral Theorem, we shall require that $G$ has at least one eigenvalue and we shall show this in the next theorem. For the coming theorems, we take $V$ to be a finite dimensional, complex inner product space.

THEOREM 4.2. *Let $G$ be a Hermitian matrix, then $G$ has at least one eigenvalue.*

Proof. We first introduce the following norm defined on the space of operators on $V$. Here $S$ shall denote the unit sphere in $V$.

$$\|T\| = \sup_{x \in S} \|Tx\|$$

for any operator $T$. We take $\lambda = \|G\|$ . As $\lambda$ is a supremum, we may take a sequence $\{x_k\}_{k\in\mathbb{N}}$, $x_k \in S$ such that $\|Gx_k\| \to \lambda$. By Lemma 3.7 we are guaranteed the existence of a subsequence of $\{x_k\}_{k\in\mathbb{N}}$ which converges to some limit $x_0 \in S$. With slight abuse of notation, we denote this subsequence by $x_k$. By continuity of the norm function, we may conclude that $\|Gx_0\| = \lambda$. We may easily verify that $\|G^2\| \geq \|G\|^2 = \lambda^2$, indeed

$$\lambda^2 = \|Gx_0\|^2 = \langle Gx_0, Gx_0 \rangle = \langle G^2x_0, x_0 \rangle \leq \left\|G^2x_0\right\| \|x_0\| \leq \left\|G^2\right\| \|x_0\|^2 = \left\|G^2\right\|.$$

In general, we also have that $\|G_1G_2\| \leq \|G_1\|\|G_2\|$ and hence we may conclude that $\lambda^2 = \|G^2\|$. In particular, we must have that $\langle G^2x_0, x_0\rangle = \|G^2x_0\| \|x_0\|$. By the Cauchy Schwartz inequality, we have that equality occurs if and only if $G^2x_0 = cx_0$ for some constant $c \in \mathbb{C}$. By noting that $\lambda^2 = \langle G^2x_0, x_0\rangle = \langle cx_0, x_0\rangle = c$, we may conclude that $c = \lambda^2$.

Then, we must have that $(G^2 - \lambda^2 I)x_0 = 0$. which in turn gives rise to two possible cases. If $(G - \lambda I)x_0 = 0$ then $\lambda$ is an eigenvalue of $G$ with corresponding eigenvector $x_0$. Else if $v = (G - \lambda I)x_0 \neq 0$ then $(G + \lambda I)v = 0$ and hence $-\lambda$ is an eigenvalue of $G$ with corresponding eigenvector $v$ finishing the proof. $\square$

In an effort to obtain a decomposition of the operator, we shall next show a decomposition of the space $V$ which shall help when decomposing the operator later on.

Theorem 4.3. *Let $G$ be a Hermitian operator and $V$ be an inner product space satisfying* $\dim(V) = n$. *Then there exist eigenvalues $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ of $G$ and eigen-subspaces $V_i = \{x \in V : Gx = \lambda_i x\}$ such that $V = \bigoplus_{i=1}^n V_i$.*

Proof. We shall proceed by induction on $n = \dim(V)$.
For the base case, we consider a one dimensional space $V$ spanned by the vector $v$. By Theorem 4.2, we are guaranteed the existence of an eigenvalue $\lambda_1$ and

a corresponding eigenvector $v_1$. As $V$ is one dimensional, we must have that $V = \text{span}(v) = \text{span}(v_1) = V_1$ hence showing that the base case holds.

For the inductive step, we consider an $n$-dimensional space $V$. By similar reasoning to the base case, we let the space $V_1$ be the space generated by $v_1$ and we let $W = V_1^{\perp}$. It is easily seen that $W$ is invariant under $G$. Using the fact that $G$ is Hermitian we may note that the following holds

$$\langle Gw, v\rangle = \langle w, Gv\rangle = \langle w, \lambda_1 v\rangle = 0 \text{ for } v \in V_1, w \in W.$$

Consequently, $V = V_1 \bigoplus W$. By using the fact that the restricted operator $G : W \to W$ is still Hermitian and that $W$ is $n-1$ dimensional, the induction hypothesis establishes the required result. □

At this juncture, we are now ready to prove the main result of this section.

Theorem 4.4 (Spectral Theorem). *A matrix $G$ is Hermitian if and only if it admits the following eigen-decomposition*

$$G = \sum_{i=1}^{k} \lambda_i G_{\lambda_i},$$

*where $\lambda$ is an eigenvalue of $G$ and $G_{\lambda_i}$ denotes the orthogonal projection of the Hilbert space $\mathcal{H}$ onto the subspace generated by the eigenvector(s) corresponding to eigenvalue $\lambda_i$. Furthermore, $k$ is the number of distinct eigenvalues of $G$.*

Proof. The backward implication is trivial by noting that, due to Theorem 4.1, $G_{\lambda_i} G_{\lambda_j} = \mathbf{0}$ for $i \neq j$ and that any projection $P$ satisfies $P = P^2 = P^*$.

To prove the forward implication, we shall show that the sum $\sum_{i=1}^{k} G_{\lambda_i}$ is a resolution of the identity that is $\sum_{i=1}^{k} G_{\lambda_i} = I$. To show this we shall first note that by Theorem 4.3 any vector $v \in V$ may be written as a direct sum

$v = v_1 + v_2 + ... + v_k$. Thus we obtain

$$
\begin{aligned}
\left(\sum_{i=1}^{k} G_{\lambda_i}\right) v &= \left(\sum_{i=1}^{k} G_{\lambda_i}\right)(v_1 + v_2 + ... + v_k) \\
&= \sum_{i=1}^{k} G_{\lambda_i} v_i \qquad (56) \\
&= \sum_{i=1}^{k} v_i \\
&= v.
\end{aligned}
$$

(56) is justified by noting that $G_{\lambda_i} v_j = 0$ for $i \neq j$ due to Theorem 4.1. This shows that the sum is a resolution of the identity as required. From this we may conclude the following for any $v \in V$,

$$
\begin{aligned}
Gv &= \left(\sum_{i=1}^{k} G_{\lambda_i} G\right)(v_1 + v_2 + ... + v_k) \\
&= \left(\sum_{i=1}^{k} G_{\lambda_i} G v_i\right) \\
&= \sum_{i=1}^{k} \lambda_i G_{\lambda_i} v_i \\
&= \left(\sum_{i=1}^{k} \lambda_i G_{\lambda_i}\right) v.
\end{aligned}
$$

From this we conclude that $G = \sum_{i=1}^{k} \lambda_i G_{\lambda_i}$ as required. □

REMARK 4.1. *We should note that this theorem may be extended. Indeed, one may generalise it to both operators in infinite-dimensional vector spaces and unbounded operators. This result is beyond the scope of the dissertation. We direct the interested reader to [22] for further discussion on the subject.*

As a consequence of the Spectral Theorem, finite-dimensional observables are in one-to-one correspondence with self-adjoint matrices. Therefore, in the finite-dimensional setting, it is justified to also refer to observables as self-adjoint matrices.

## 4.2. Finite Probability Spaces

As any bounded linear operator may be represented by a matrix then, when working in finite dimensions, any observable must be a Hermitian matrix. We will give an example below as to illustrate how one may go about calculating the probability of a particular event in the context of quantum probability.

EXAMPLE 4.1. Consider the observable $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and the state $\rho = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Through simple calculations, we may verify that $X$ has eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 1$ with corresponding unit eigenvectors $v_1 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and $v_2 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$.

Then, we may verify that $X$ admits the following eigen-decomposition.

$$\begin{aligned} X &= (-1)\frac{1}{2}\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -1 & 1 \end{smallmatrix}\right) + (1)\frac{1}{2}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \end{smallmatrix}\right) \\ &= (-1)\left(\begin{smallmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{smallmatrix}\right) + (1)\left(\begin{smallmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{smallmatrix}\right). \end{aligned}$$

We may take $\Omega$ to be the Borel $\sigma$-algebra on $\mathbb{R}$. Now to ascertain the probability of observing $-1$ and $1$ we must determine the measure of the space generated by the projections to which they are mapped. For the state $\rho$ we may conclude, by Theorem 3.5, that the probability of $\omega \in \Omega$ being mapped to $-1$ and $1$ is precisely $\mathrm{Tr}(\rho\left(\begin{smallmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{smallmatrix}\right)) = \mathrm{Tr}(\rho\left(\begin{smallmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{smallmatrix}\right)) = \frac{1}{2}$.

**4.2.1. Interfering Observables and Commuting Families.** At this juncture, we may discuss the notion of interfering variables. To define this notion however we must first introduce the commutator. The proofs for this section are taken from [12].

DEFINITION 4.1 (Commutator). *For two observables $X_1$ and $X_2$, the commutator $[X_1, X_2]$ is given by $X_1X_2 - X_2X_1$.*

From this, we may define the notion of a commutating family which shall be useful later on.

DEFINITION 4.2 (Commuting Family). *A family of matrices $\mathfrak{F}$ is said to be a commuting family if and only if for any $F_1, F_2 \in \mathfrak{F}$ we have that $[F_1, F_2] = 0$.*

Definition 4.3. *Two observables $X_1$ and $X_2$ are said to be interfering if* $[X_1, X_2] \neq 0$.

We shall see the reasoning behind the use of the term "interfering" in section 4.2.2. We may also note that in a classical setting, all random variables are non-interfering. To motivate the discussion in this section we consider the following example.

Example 4.2. Consider two observables $X_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $X_2 = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$. Taking $\Omega$ in a similar manner to Example 4.1 and performing a similar calculation, we may conclude that, in the state $\rho$, $X_1$ and $X_2$ assume values $-1$ and $1$ each with probability $\frac{1}{2}$.

However, taking the observable $X = X_1 + X_2$ and performing the eigen- decomposition, we find that $X$ assumes the values $\sqrt{2}$ and $-\sqrt{2}$ with probability $\frac{1}{2}$ each. This contrasts behaviour of random variables in classical probability for if $Y$ and $Y'$ are random variables which assume values from the sets $\{y_1, y_2, ..., y_n\}$ and $\{y'_1, y'_2, ..., y'_m\}$ respectively then $Y + Y'$ is another random variable which assumes values from the set $\{y_i + y'_j : 1 \leq i \leq n, 1 \leq j \leq m\}$.

Hence one may ask when indeed do observables obey such a relation. We may provide a partial answer to this, however first we must introduce the notion of simultaneously diagonalizable matrices and results related to them. We start off by giving the definition of simultaneously diagonalizable matrices.

Definition 4.4 (Simultaneously Diagonalizable Matrices). *A family of matrices $\mathfrak{F}$ is said to be simultaneously diagonalizable if and only if there exists a matrix $S$ such that for all $F \in \mathfrak{F}$, $S^{-1}FS$ is a diagonal matrix.*

In the first lemma, we shall see what constraints a matrix $F$ must obey in order for it to commute with a particular type of block matrix which we define next.

Definition 4.5. *A block matrix $D'$ is a matrix which assumes the following form,*

$D' = \begin{pmatrix} D'_1 & & \\ & \ddots & \\ & & D'_s \end{pmatrix}$ *where each $D'_i$ is an $n_i \times n_i$ square matrix and $s$ is said to be the number of blocks. In addition, if the only non-zero entries of $D'$ are $D'_1, \cdots, D'_s$ then $D'$ is said to be a block diagonal matrix.*

LEMMA 4.1. *Let $D'$ be a block diagonal matrix, where each diagonal block has the form*

$$D'_i = \lambda_i I_{n_i}, \quad i = 1, \dots, s,$$

*with distinct scalars $\lambda_1, \dots, \lambda_s$ (that is, $\lambda_i \neq \lambda_j$ for $i \neq j$). If a matrix $F$ commutes with $D'$, then all off-diagonal blocks of $F$ must vanish. In other words, $F_{ij} = 0$ for all $i \neq j$, so $F$ is itself block diagonal.*

PROOF. Partition $F$ conformally with $D'$, so that

$$F = \begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1s} \\ F_{21} & F_{22} & \cdots & F_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ F_{s1} & F_{s2} & \cdots & F_{ss} \end{pmatrix},$$

where each block $F_{ij}$ has size $n_i \times n_j$.

Since $D'F = FD'$, looking at the $(i, j)$-block gives

$$\lambda_i F_{ij} = F_{ij} \lambda_j, \quad \text{for all } i, j \in \{1, \dots, s\}.$$

Equivalently,

$$(\lambda_i - \lambda_j) F_{ij} = 0.$$

Because the scalars $\lambda_1, \dots, \lambda_s$ are distinct, it follows that $F_{ij} = 0$ whenever $i \neq j$. Thus, $F$ is block diagonal. □

Next we shall obtain a relation between a block diagonal matrix $D'$ and its individual blocks. We state the theorem without proof.

Theorem 4.5. *Let $D'$ be a block diagonal matrix. Then $D'$ is diagonalizable if and only if $D'_i$ is diagonalizable for each $i \in \{1, ..., s\}$.*

With these results established, we are now in a position to prove the main result of this section.

Theorem 4.6. *Let $\mathfrak{F}$ be a family of diagonalizable matrices each of dimension $n \times n$. Then $\mathfrak{F}$ is a simultaneously diagonalizable family if and only if it is a commuting family.*

Proof. We let $\mathfrak{F}$ be a simultaneously diagonalizable family, then we are guaranteed the existence of a matrix $P$ such that $P^{-1}F_iP = D_i$ for all $F_i \in \mathfrak{F}$. Here, $D_i$ is a diagonal matrix. Let $F_1, F_2 \in \mathfrak{F}$. By writing $F_i = PD_iP^{-1}$ and by noting that diagonal matrices commute with one another, it follows trivially that $\mathfrak{F}$ is a commuting family.

We now establish the converse by induction on the dimension $n$. The base case $n = 1$ follows trivially. We assume that for $d \in \{1, ..., n-1\}$ any commuting family of $d \times d$ diagonalizable matrices is simultaneously diagonalizable. We now consider $n \times n$ matrices.

If every $F \in \mathfrak{F}$ is of the form $aI$ for some $a \in \mathbb{C}$ then the result follows trivially. Then we assume that $F \in \mathfrak{F}$ is a $n \times n$ diagonalizable matrix having distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$, $k \geq 2$ with each $\lambda_i$ having multiplicity $n_i$, $F$ satisfies $FF' = F'F$ for all $F' \in \mathfrak{F}$ and that each $F' \in \mathfrak{F}$ is diagonalizable.

We consider the matrix $\hat{P}$ which diagonalizes $F$ (but not necessarily all other $F' \in \mathfrak{F}$) and we apply it to $F$. Hence we obtain the following

$$\hat{F} = \hat{P}^{-1}F\hat{P} = \begin{pmatrix} \lambda_1 I_{n_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_k I_{n_k} \end{pmatrix}$$

and we obtain $\hat{F}'$ in a similar manner. As $\hat{F}\hat{F}' = \hat{F}'\hat{F}$ for all $F' \in \mathfrak{F}$, Lemma 4.1 ensures that each $\hat{F}'$ assumes the following form

$$\hat{F}' = \begin{pmatrix} \hat{F}'_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \hat{F}'_k \end{pmatrix}$$

where $\hat{F}'_i$ is a $n_i \times n_i$ matrix. As $\hat{F}'$ is diagonalizable, Theorem 4.5 gives us that each $\hat{F}'_i$ is diagonalizable. As these are all of dimension less than $n$, we are guaranteed the existence of $k$ similarity matrices $P_1, P_2, .., P_k$ of the appropriate size each of which diagonalize the corresponding block $\hat{F}'_i$. Then we define the matrix

$$P = \begin{pmatrix} P_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & P_k \end{pmatrix}.$$

We need only note that, as $P_i^{-1}\lambda_i I_{n_i} P_i = \lambda_i I_{n_i}$, then both $P^{-1}\hat{P}^{-1}F\hat{P}P$ and $P^{-1}\hat{P}^{-1}\hat{F}\hat{P}P$ are diagonal making the matrix $\hat{P}P$ the required similarity matrix. □

At this juncture, we may now return to our discussion of the observables. Our main aim shall be to utilize Theorem 4.6 and to this end we need only show that Hermitian matrices are diagonalizable. We state the following theorem without proof.

THEOREM 4.7. *Idempotent matrices are diagonalizable.*

THEOREM 4.8. *Hermitian matrices are diagonalizable.*

PROOF. By Theorem 4.4 we may decompose the Hermitian matrix $G$ by taking $G = \sum_{i=1}^{k} \lambda_i G_{\lambda_i}$. The $G_{\lambda_i}s$ are idempotent as they are projections and they commute as $G_{\lambda_i}G_{\lambda_j} = G_{\lambda_j}G_{\lambda_i} = \mathbf{0}$. Hence by Theorem 4.6 these are simultaneously diagonalizable by some invertible matrix $P$. So we must have that the following holds.

$$\begin{aligned} G &= \sum_{i=1}^{k} \lambda_i G_{\lambda_i} \\ &= \sum_{i=1}^{k} \lambda_i P D_i P^{-1} \\ &= P(\sum_{i=1}^{k} \lambda_i D_i) P^{-1}. \end{aligned}$$

By noting that the sum of diagonal matrices is again diagonal, we have obtained that $G$ is diagonalizable. □

Let $X_1$ and $X_2$ be two commuting observables. Then by Theorem 4.6 we must have that these are simultaneously diagonalizable. That is $X_1 + X_2 = PD_{X_1}P^{-1} + PD_{X_2}P^{-1} = P(D_{X_1} + D_{X_2})P^{-1}$. From here, it may easily be seen that $X_1 + X_2$ is similar to a diagonal matrix composed of entries of the form $\lambda_{X_1} + \lambda_{X_2}$ and hence the values assumed by the observable $X_1 + X_2$ will be from the set $\{\lambda_{X_1} + \lambda_{X_2} : \lambda_{X_1} \in \sigma(X_1), \lambda_{X_2} \in \sigma(X_2)\}$ as we have required to show.

It should be noted that commuting observables do not make up all the observables which obey the mentioned relation. We consider an example found in [23] which shows this.

Example 4.3. Consider the observables

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 6-\sqrt{2} & 2 \\ 0 & 2 & 4+2\sqrt{2} \end{pmatrix}, \; X_2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then the spectra of $X_1, X_2$ and $X_1+X_2$ are $\{0, 2+\sqrt{2}, 8\}, \{-4, 0, 4\}$ and $\{0, 4, 6+\sqrt{2}\}$ respectively. Clearly, $X_1$ and $X_2$ satisfy the relation discussed however it is easily verified that $X_1X_2 \neq X_2X_1$.

Further conditions regarding when the eigenvalues of $X_1 + X_2$ are given by entries of the set $\sigma(X_1) + \sigma(X_2)$ may be found in [23]. Next we shall discuss two central properties of quantum probability which, like in classical probability, are of great importance.

**4.2.2. Expectation and Variance.** We will now proceed to define both the expectation and the variance of an observable $X$. We note that as every state induces a probability space, then $X$ can have a different expectation and variance for each state. Unless otherwise stated, the definitions and notations in this section are found in [19].

We have already seen the expectation of a random variable $Y$. We shall now define the analgoue of this in the context of quantum probability.

DEFINITION 4.6 (Expectation of an Observable). *The expectation of an observable $X$ in state $\rho$ is given by*

$$\mathbb{E}[X] = \operatorname{Tr}(\rho X).$$

With $\rho$ assuming the role of the classic probability distribution, we may clearly observe the relation between the two formulae. The same exact observation may be made for the variance of an observable which we define next.

DEFINITION 4.7 (Variance of an Observable). *The variance of an observable $X$ in state $\rho$ is given by*

$$\operatorname{Var}_\rho(X) = \operatorname{Tr}(\rho X^2) - (\operatorname{Tr}(\rho X))^2 = \operatorname{Tr}(\rho(X - \operatorname{Tr}\rho X)^2).$$

Through the next theorem, we shall see how notions of the variance may be used to mathematically describe a well studied phenomenon in quantum mechanics, this being the Heisenberg Uncertainty Principle. Prior to this however, we must define a notion of a pure state.

DEFINITION 4.8 (Pure State). *A state is called a pure state if it cannot be written as a convex combination of other states. We may easily verify that a state is pure if and only if is of the form $\rho = uu^T$ where $u$ is a unit vector in some Hilbert Space $\mathcal{H}$.*

DEFINITION 4.9 (Mixed State). *A state is called a mixed state if it is not pure.*

REMARK 4.2. *We may note that if $\rho = uu^T$ is a pure state corresponding to unit vector $u$, then $\operatorname{Tr}(\rho X) = \langle u, Xu \rangle$.*

From this we first proceed to the following lemma.

LEMMA 4.2. *The variance of an observable $X$ at state $\rho = \sum_{i=1}^{k} p_i u_i u_i^T$ where $\sum_{i=1}^{k} p_i = 1$, satisfies the following.*

$$\mathrm{Var}_\rho(X) = \sum_{i=1}^{k} p_i ||(X - \mathrm{Tr}(\rho X) I)||^2.$$

PROOF. Let $\rho$ be a state given by $\rho = \sum_{i=1}^{k} p_i u_i u_i^T$ where $\sum_{i=1}^{k} p_i = 1$. For simplicity, we write $m = \mathrm{Tr}(\rho X)$. Then the following holds,

$$\begin{aligned}
\mathrm{Var}_\rho(X) &= \mathrm{Tr}(\rho(X - mI)^2) \\
&= \sum_{i=1}^{k} p_i \,\mathrm{Tr}(u_i u_i^T (X - mI)^2) \\
&= \sum_{i=1}^{k} p_i \langle u_i, (X - mI)^2 u_i\rangle \\
&= p_i \sum_{i=1}^{k} p_i \langle (X - mI) u_i, (X - mI) u_i\rangle \\
&= \sum_{i=1}^{k} p_i ||(X - mI) u_i||^2.
\end{aligned}$$

□

THEOREM 4.9. *For any pure state $uu^T$ and observables $X_1, X_2$ the following inequality must hold*

$$\mathrm{Var}_{uu^T}(X_1)\,\mathrm{Var}_{uu^T}(X_2) \geq \frac{1}{4} |\langle u, i[X_1, X_2] u\rangle|^2.$$

PROOF. For notational simplicity, we let $a = \langle u, X_1 u\rangle$ and $b = \langle u, X_2 u\rangle$ and it is immediately seen that $a$ and $b$ are real scalars. Then consider the following,

$$\begin{aligned}
\langle u, i[X_1, X_2] u\rangle &= \langle u, i[X_1 - aI, X_2 - bI] u\rangle \\
&= i(\langle (X_1 - aI) u, (X_2 - bI) u\rangle - \langle (X_2 - bI) u, (X_1 - aI) u\rangle) \\
&= -2\,\mathrm{Im}(\langle (X_1 - aI) u, (X_2 - bI) u\rangle).
\end{aligned}$$

Proceeding with the Cauchy Schwartz inequality and Lemma 4.2, the following is evident

$$\begin{aligned}\frac{1}{2}\langle u, i[X_1, X_2]u\rangle &\leq (||(X_1 - aI)u||)(||(X_2 - bI)u||)\\ &= (\operatorname{Var}_{uu^T}(X_1)\operatorname{Var}_{uu^T}(X_2))^{\frac{1}{2}}.\end{aligned}$$

□

Now consider the case when $X_1$ and $X_2$ do no commute with one another, then it follows as a result of the Spectral Theorem that there exists some unit vector $u$ and a real scalar $\delta \neq 0$ such that $i[X_1, X_2]u = \delta u$. By Theorem 4.9, we have that $X_1$ and $X_2$ satisfy

$$\operatorname{Var}_{uu^T}(X_1)\operatorname{Var}_{uu^T}(X_2) \geq \frac{1}{4}\delta^2. \tag{57}$$

This means that, for the inequality to hold, any decrease in variance in one observable must be balanced by a corresponding increase in variance of the other. Physically, this may be interpreted as the impossibility of concurrently measuring both $X_1$ and $X_2$ with exact accuracy and this is precisely what is described by the Heisenberg Uncertainty Principle. In classical probability, all variables commute and consequently the right hand side of (57) becomes 0. In this case, given that the variance is always positive, (57) will impose no additional restriction on the variance of the random variables.

CHAPTER 5

# Multiple Probability Spaces

## 5.1. Tensor Product of Hilbert Spaces

In probability, for example when dealing with multivariate random variables, it is often desired to combine information regarding different random variables into one probability space. In classical probability, this may be easily achieved. Indeed, if $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ are two probability spaces, we may combine them into a single probability space by considering the space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$ where $\mathcal{F}_1 \times \mathcal{F}_2$ and $\mathbb{P}_1 \times \mathbb{P}_2$ denote the product $\sigma$ algebra and product measure respectively. To achieve the same notion in quantum probability we need to utilize tensor products which we shall discuss in this section. We shall follow the construction due to [19]. This construction is reliant on kernels which we define in the following definition. We also give an important result which will be used in this section without proof from [19].

Theorem 5.1. *For $i \in \{1, 2\}$, let $S_i$ be a total subspace of a Hilbert Space $\mathcal{H}_i$; that is, $S_i^{\perp} = \{0\}$. Let $U_0 : S_1 \to S_2$ be a scalar product preserving map. Then there exists a linear isometry $U : \mathcal{H}_1 \to \mathcal{H}_2$ which extends $U_0$.*

Definition 5.1 (Kernel). *Let $H$ be a set. A kernel $K : H \times H \to \mathbb{C}$ is a map where for every $\alpha_i \in \mathbb{C}, x_i \in H, i, j = 1, 2, ..., n$ the following holds:*

$$\sum_{i,j} \bar{\alpha}_i \alpha_j K(x_i, x_j) \geq 0.$$

For the case when the underlying set $H$ is finite, the kernel is simply a positive semi-definite matrix. We shall denote the space of kernels on $H$ by $K(H)$. For the construction, we shall require that $K(H)$ is closed under pointwise multiplication. To this end we proceed with the following two results. The first result is a standard result which we state without proof.

Theorem 5.2 (Cholesky Decomposition). *If $A$ is a positive semi-definite Hermitian matrix, then there exists a matrix $C$ such that $A = CC^*$.*

Theorem 5.3. *If $A$ and $A'$ are two positive semi-definite matrices then their Hadamard product, that is the matrix obtained by element wise multiplication of $A$ and $A'$, is also positive semi-definite.*

Proof. To show this, it is enough to show that the Hadamard product of $A$ and $A'$ is a variance covariance matrix as, by Theorem 2.5 this will give us that it is positive semi-definite. Choose a matrix $C$ such that $A = CC^*$, this is guaranteed to exist by Theorem 5.2. Let $\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_n$ be any $2n$ independent and identically distributed standard normal variables. Write $\gamma_j = 2^{-\frac{1}{2}}(\alpha_j + i\beta_j)$ and $\underline{\xi} = C\underline{\gamma}$ whereby $\underline{\gamma}$ is a column vector having $\gamma_j$ as its components. Then it may be easily verified that $\underline{\xi}$ is a complex valued Gaussian random vector which satisfies $\mathbb{E}[\underline{\xi}] = 0,\ \mathbb{E}[\underline{\xi}\underline{\xi}^*] = CC^* = A$.

By the same reasoning, we may obtain a pair of independent complex Gaussian random vectors $\underline{\eta}, \underline{\upsilon}$ such that $\mathbb{E}[\eta_j] = \mathbb{E}[\upsilon_j] = 0$ and $\mathbb{E}[\overline{\eta}_i\eta_j] = A_{ij},\ \mathbb{E}[\overline{\upsilon}_i\upsilon_j] = A'_{ij}$ where $1 \leq i, j \leq n$. Let $\zeta_j = \eta_j\upsilon_j$, then $\mathbb{E}[\zeta_j] = 0$ and $\mathbb{E}[\overline{\zeta}_i\zeta_j] = A_{ij}A'_{ij}$. Hence it follows that the Hadamard product of $A$ and $A'$ is the variance covariance matrix of the complex random vector $\underline{\zeta}$ and is thus positive semi definite by Theorem 2.5. ☐

Remark 5.1. *Theorem 5.3 is known as the Schur Product Theorem. Although it can be proven using purely linear-algebraic techniques as shown in [12], we chose to present the proof using a Gaussian random vector argument, since similar reasoning will be employed in Theorem 5.6.*

This leads us to the following corollaries which will be important when defining the tensor product.

Corollary 5.4. *The space $K(H)$ of all kernels on the set $H$ is closed under pointwise multiplication.*

Proof. To prove this, we need only observe that $K$ is a kernel on $H$ if and only if for any finite set $\{x_1, x_2, ..., x_n\} \subseteq H$ the matrix $A$ whose entries are given by $A_{ij} = K(x_i, x_j)$ is positive semi definite. This result along with Theorem 5.3 give the result. $\square$

REMARK 5.2. *By the same observation made in the proof of Corollary 5.4 we may easily verify that the inner product of a Hilbert Space $\mathcal{H}$ is indeed a kernel acting on $H$ where $H$ denotes the elements of $\mathcal{H}$ as a set.*

COROLLARY 5.5. *Let $H_i$, $i = 1, 2, .., n$ be sets and let $K_i \in K(H_i)$ for each $i$. Define $H = H_1 \times H_2 \times ... \times H_n$ and*

$$K(\underline{x}, \underline{y}) = \prod_{i=1}^{n} K_i(x_i, y_i).$$

*For $\underline{x} = (x_1, ..., x_n), \underline{y} = (y_1, ..., y_n)$ where $x_i, y_i \in H_i$. Then $K \in K(H)$.*

PROOF. Let $\underline{x}^{(r)} = (x_{1r}, x_{2r} ..., x_{nr}) \in H$ for $1 \leq r \leq m$. Define the matrices $A^{(1)}, A^{(2)}, ... A^{(n)}$ by taking $A^{(i)}_{rs} = K_i(x_{ir}, x_{is})$. Thus we may observe that

$$K(\underline{x}^{(r)}, \underline{x}^{(s)}) = \prod_{i=1}^{n} A^{(i)}_{rs}.$$

By noting that each $A^{(i)}$ is positive semi definite, the corollary follows from Theorem 5.3. $\square$

THEOREM 5.6. *Let $H$ be any set and let $K \in K(H)$. Then there exists a Hilbert space $\mathcal{H}$ and a map $\lambda : H \to \mathcal{H}$ which satisfy the following properties*

*i) The set $\{\lambda(x) : x \in H\}$ is total in $\mathcal{H}$.*

*ii) $K(x, y) = \langle \lambda(x), \lambda(y) \rangle$ for all $x, y \in H$.*

*In addition, if $\mathcal{H}'$ is another Hilbert space and $\lambda' : H \to \mathcal{H}'$ is another map satisfying i) and ii) then there exists a unitary isomorphism $U : \mathcal{H} \to \mathcal{H}'$ such that $U\lambda(x) = \lambda'(x)$ for all $x \in H$.*

PROOF. For any finite set $F_1 = \{x_1, x_2, ..., x_n\} \subseteq H$ it follows by a similar argument to Theorem 5.3 that there is some complex Gaussian random vector $(\zeta_1^{F_1}, \zeta_2^{F_1}, ..., \zeta_n^{F_1})$ such that for $1 \leq i, j \leq n$,

$$\mathbb{E}[\zeta_i^{F_1}] = 0, \ \mathbb{E}[\overline{\zeta_i^{F_1}} \zeta_j^{F_1}] = K(x_i, x_j).$$

Now if $F_1 \subset F_2 = \{x_1, x_2, ..., x_n, x_{n+1}\}$ then the marginal distribution of $(\zeta_1^{F_2}, \zeta_2^{F_2}, ..., \zeta_n^{F_2})$ derived from that of $(\zeta_1^{F_2}, \zeta_2^{F_2}, ..., \zeta_n^{F_2}, \zeta_{n+1}^{F_2})$ is the same as that of

$(\zeta_1^{F_1}, \zeta_2^{F_1}, ..., \zeta_n^{F_1})$. Then, by Kolmogorov's Consistency Theorem, we are guaranteed the existence of a family of complex Gaussian random variables $\{\zeta_x : x \in H\}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that for $x, y \in H$,

$$\mathbb{E}[\zeta_x] = 0, \ \mathbb{E}[\overline{\zeta_x}\zeta_y] = K(x, y).$$

By taking $\mathcal{H}$ to be the closed linear span of $\{\zeta_x : x \in H\}$ in $L^2(\mathbb{P})$ and $\lambda(x) = \zeta_x$ we note that the first part of the theorem holds.

To establish the required isomorphism we let $S = \{\lambda(x) : x \in H\}$ and $S' = \{\lambda'(x) : x \in H\}$. Define $U_0 : S \to S'$ by taking $U_0(\lambda(x)) = \lambda'(x)$. Trivially, $U_0$ is scalar product preserving and $S$ and $S'$ are total in $\mathcal{H}$ and $\mathcal{H}'$ respectively. By Theorem 5.1, $U_0$ may be extended to an isomorphism $U : \mathcal{H} \to \mathcal{H}'$ hence giving the result. □

Definition 5.2 (Gelfand Pair). *Let $(\mathcal{H}, \lambda)$ be the pair determined uniquely up to a unitary isomorphism by kernel $K$ on the set $H$. Then $(\mathcal{H}, \lambda)$ is said to be a Gelfand pair associated with $K$.*

Through the Gelfand pair, we may now define the notion of the tensor product between Hilbert spaces.

Definition 5.3 (Tensor Product of Hilbert Spaces). *Let $\mathcal{H}_i$ for $1 \leq i \leq n$ be Hilbert spaces. Take $H = \mathcal{H}_1 \times ... \times \mathcal{H}_n$, the Cartesian product obtained when treating the $\mathcal{H}_i$s as sets. Define a kernel on $H$ by taking*

$$K(\underline{u}, \underline{v}) = \prod_{i=1}^{n} \langle u_i, v_i \rangle, \ \underline{u} = (u_1, ..., u_n), \ \underline{v} = (v_1, ..., v_n)$$

*where $u_i, v_i \in \mathcal{H}_i$. Consider a Gelfand pair $(\mathcal{H}, \lambda)$ associated with $K$ which satisfies i) and ii). Then $\mathcal{H}$ is said to be a tensor product of $\mathcal{H}_i$, $i = 1, 2, ..., n$. We write*

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_n = \bigotimes_{i=1}^{n} \mathcal{H}_i,$$

$$\lambda(\underline{u}) = u_1 \otimes u_2 \otimes ... \otimes u_n = \bigotimes_{i=1}^{n} u_i$$

*and we call $\lambda(\underline{u})$ the tensor product of vectors $u_i, 1 \leq i \leq n$.*

For the sake of completeness, we shall prove some elementary properties of the tensor products.

THEOREM 5.7. *The tensor product satisfies the following properties:*

*i)* $\langle \bigotimes_{i=1}^n u, \bigotimes_{i=1}^n v \rangle = \prod_{i=1}^n \langle u_i, v_i \rangle$.

*ii) The map* $(u_1, u_2, ..., u_n) \to u_1 \otimes u_2 \otimes ... \otimes u_n$ *is multilinear.*

*iii) The set* $\{\bigotimes_{i=1}^n u_i : u_i \in \mathcal{H}_i, i = 1, 2, ..., n\}$ *is total in* $\bigotimes_{i=1}^n \mathcal{H}_i$.

PROOF. Firstly, consider

$$\begin{aligned}
\langle \lambda(\underline{u}), \lambda(\underline{v}) \rangle &= \langle \zeta_{\underline{u}}, \zeta_{\underline{v}} \rangle \\
&= \mathbb{E}[\overline{\zeta}_{\underline{u}} \zeta_{\underline{v}}] \\
&= K(\underline{u}, \underline{v}) \\
&= \prod_{i=1}^n \langle u_i, v_i \rangle
\end{aligned}$$

which proves $i$).

To show $ii$) we need only use $i$) as well as the sesquilinearity of the inner product. Indeed, let $i$ be fixed.

$$\begin{aligned}
(58) \qquad & \|u_1 \otimes ... \otimes u_{i-1} \otimes (\alpha u_i + \beta v_i) \otimes u_{i+1} \otimes ... \otimes u_n - \alpha \bigotimes_{j=1}^n u_j \\
& - \beta u_1 \otimes ... \otimes u_{i-1} \otimes v_i \otimes u_{i+1} \otimes ... \otimes u_n\|^2
\end{aligned}$$

Now for ease of notation we let

$$\begin{aligned}
A &= u_1 \otimes ... \otimes u_{i-1} \otimes (\alpha u_i + \beta v_i) \otimes u_{i+1} \otimes ... \otimes u_n, \\
B &= \alpha \bigotimes_{j=1}^n u_j, \\
C &= \beta u_1 \otimes ... \otimes u_{i-1} \otimes v_i \otimes u_{i+1} \otimes ... \otimes u_n
\end{aligned}$$

Then the norm in (58) becomes

(59)

$$\langle A, A \rangle - \langle A, B \rangle - \langle A, C \rangle + \langle B, B \rangle - \langle B, A \rangle + \langle B, C \rangle - \langle C, A \rangle + \langle C, B \rangle + \langle C, C \rangle$$

By expanding the above using *i*) and noting that the inner products only differ in the *ith* vector, we obtain that (59) becomes equivalent to the following expression

$$\begin{aligned}
&\prod_{j\neq i} \|u_j\|^2 \left(\|\alpha u_i + \beta v_i\|^2 + |\alpha|^2 \|u_i\|^2 + |\beta|^2 \|v_i\|^2\right) \\
&- 2\,\mathrm{Re}(\langle \alpha u_i + \beta v_i, \alpha u_i\rangle + \langle \alpha u_i + \beta v_i, \beta v_i\rangle + \langle \alpha u_i, \beta v_i\rangle) \\
&= 0.
\end{aligned}$$

Finally, *iii*) is evident by the construction. □

We end this section by giving an orthonormal basis for $\bigotimes_{i=1}^n \mathcal{H}_i$.

THEOREM 5.8. *Let* $\{e_{ij} : j = 1, 2, ...\}$ *be an orthonormal basis in* $\mathcal{H}_i$ *for* $i = 1, 2, ..., n$. *Then the set*

$$\{e_{1j_1} \otimes e_{2j_2} \otimes ... \otimes e_{nj_n} : j_1 = 1, 2, ..., j_2 = 1, 2, ..., j_n = 1, 2, ...\}$$

*is an orthonormal basis for* $\bigotimes_{i=1}^n \mathcal{H}_i$. *In particular,*

$$\bigotimes_{i=1}^n u_i = \sum_{j_1, j_2 ..., j_n} \{\prod_{i=1}^n \langle e_{ij_i}, u_i\rangle\} \bigotimes_{i=1}^n e_{ij_i}$$

*where the sum on the right hand side is strongly convergent in* $\bigotimes_{i=1}^n \mathcal{H}_i$.

PROOF. First we establish orthonormality of the set. Let $x = e_{1j_1} \otimes e_{2j_2} \otimes ... \otimes e_{nj_n}$ and $y = e_{1j_1'} \otimes e_{2j_2'} \otimes ... \otimes e_{nj_n'}$. Then by Theorem 5.7*i*, the following holds

$$\begin{aligned}
\langle x, y\rangle &= \prod_{i=1}^n \langle e_i j_i, e_{ij_i'}\rangle \\
&= \prod_{i=1}^n \delta_{j_i j_i'}. \qquad (60)
\end{aligned}$$

Thus we note that the inner product in (60) is 0 if and only if $j_i = j_i'$ for all $i$ and hence orthonormality of the set follows.

Next we show that the set is a basis. By assumption, we have that each $u_i \in \mathcal{H}_i$ may be written as

$$u_i = \sum_{j_i} \langle e_{ij_i}, u_i\rangle e_{ij_i}$$

and hence we have that

$$\begin{aligned}\bigotimes_{i=1}^{n} u_i &= \bigotimes_{i=1}^{n}\left(\sum_{j_i}\langle e_{ij_i}, u_i\rangle e_{ij_i}\right)\\ &= \sum_{j_1}\sum_{j_2}\cdots\sum_{j_n}\left(\prod_{i=1}^{n}\langle e_{ij_i}, u_i\rangle\right)\bigotimes_{i=1}^{n} e_{ij_i}.\end{aligned}$$

The last step follows by multilinearity of the tensor product.

To conclude the proof we need only show that the sum on the right hand side is strongly convergent. To establish it is enough to show that $(S_n)_{n\in\mathbb{N}}$ where

$$S_N = \sum_{j_1=1}^{N}\cdots\sum_{j_n=1}^{N}\left(\prod_{i=1}^{n}\langle e_{ij_i}, u_i\rangle\right)\bigotimes_{i=1}^{n} e_{ij_i}$$

is Cauchy. However this is a direct consequence of the fact that $u_i = \sum_{j_i}\langle e_{ij_i}, u_i\rangle e_{ij_i}$ converges strongly in $\mathcal{H}_i$ and the fact that the tensor product norm satisfies

$$\left\|\bigotimes_{i=1}^{n} v_i\right\| = \prod_{i=1}^{n}\|v_i\|.$$

$\square$

Corollary 5.9. *If* $\dim(\mathcal{H}_i) = m_i < \infty$ *then*

$$\dim(\bigotimes_{i=1}^{n}\mathcal{H}_i) = \prod_{i=1}^{n} m_i.$$

## 5.2. Tensor Product of Operators

In order to extend the notion of joining multiple probability spaces to a quantum context, we shall also need to define the notion of a tensor product of operators and this is precisely what shall be discussed in this section. The results in this section are due [19].

To give an intuition as to how we can define such a tensor product, we shall consider the following case. For $i = 1, 2, \cdots, n$ let $\mathcal{H}_i$ be a Hilbert space satisfying $\dim(\mathcal{H}_i) = m_i < \infty$ and $\mathcal{H} = \bigotimes_{i=1}^{n}\mathcal{H}_i$. Furthermore, let $T_i$ be a self adjoint operator in $\mathcal{H}_i$ with eigenvalues $\{\lambda_{ij}, 1 \le j \le m_i\}$ and corresponding orthonormal set of eigenvectors $\{e_{ij} : 1 \le j \le m_i\}$ so that we obtain

$$T_i e_{ij} = \lambda_{ij} e_{ij},\ 1 \le j \le m_i,\ i = 1, 2, ..., n.$$

By Theorem 5.8 we may define a self adjoint operator $T$ on $\mathcal{H}$ by first defining

$$T \bigotimes_{i=1}^{n} e_{ij_i} = \prod_{i=1}^{n} \lambda_{ij_i} \bigotimes_{i=1}^{n} e_{ij_i} \tag{61}$$

for $1 \le j_i \le m_i$ and then extending by linearity on all of $\mathcal{H}$.

It may be easily verified that the operator $T$ has eigenvalues $\prod_{i=1}^{n} \lambda_{ij_i}$ and satisfies

$$T \bigotimes_{i=1}^{n} u_i = \bigotimes_{i=1}^{n} T_i u_i \quad \text{for all } u_i \in \mathcal{H}_i,\ 1 \le i \le n.$$

Furthermore,

$$\begin{aligned}
\|T\| &= \max\left(\left|\prod_{i=1}^{n} \lambda_{ij_i}\right|,\ 1 \le j_i \le m_i\right) \\
&= \prod_{i=1}^{n} \max\left(|\lambda_{ij}|,\ 1 \le j \le m_i\right) \\
&= \prod_{i=1}^{n} \|T_i\|.
\end{aligned}$$

In particular, the following identity holds

$$\begin{aligned}
\left|\sum_{1 \le j,k \le N} \overline{\alpha}_j \alpha_k \prod_{i=1}^{n} \langle u_{ij}, T_i u_{ik}\rangle\right| &= \left|\left\langle \sum_{j=1}^{N} \alpha_j \bigotimes_{i=1}^{n} u_{ij},\ T \sum_{j=1}^{N} \alpha_j \bigotimes_{i=1}^{n} u_{ij} \right\rangle\right| \\
&\le \prod_{i=1}^{n} \|T_i\| \left\| \sum_{j=1}^{N} \alpha_j \bigotimes_{i=1}^{n} u_{ij} \right\|^2
\end{aligned} \tag{62}$$

for all scalars $\alpha_j$, $u_{ij} \in \mathcal{H}_i$, $1 \le j \le N$, $1 \le i \le n$, $N = 1, 2, \cdots$. We write $T = \bigotimes_{i=1}^{n} T_i = T_1 \otimes \cdots \otimes T_n$ and this is said to be the tensor product of operators $T_i,\ 1 \le i \le n$. In the next theorem, we shall extend this idea to all bounded operators on the Hilbert Space $\mathcal{H}$. We recall that $\mathcal{B}(\mathcal{H})$ denotes the set of bounded operators on a Hilbert Space $\mathcal{H}$.

THEOREM 5.10. *Let $\mathcal{H}_i$, $1 \le i \le n$ be Hilbert spaces and let $T_i \in \mathcal{B}(\mathcal{H}_i)$ for each $i$. Then there exists a unique bounded operator $T$ in $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i$ which* *satisfies*

$$T\bigotimes_{i=1}^{n}u_i=\bigotimes_{i=1}^{n}T_iu_i. \tag{63}$$

*where $u_i\in\mathcal{H}_i, 1\le i\le n$. Furthermore we have that $\|T\|=\prod_{i=1}^{n}\|T_i\|$.*

Proof. Let $S=\{\bigotimes_{i=1}^{n}u_{ij}\mid u_{ij}\in\mathcal{H}_i,\ 1\le j\le N\}$. For all scalars $\alpha_j$, $1\le j\le N$ and product vectors $\bigotimes_{i=1}^{n}u_{ij}\in S$, $1\le j\le N$ the following holds

$$\begin{aligned}\left\|\sum_{j=1}^{N}\alpha_j\bigotimes_{i=1}^{n}T_iu_{ij}\right\|^2&=\sum_{1\le j,k\le N}\overline{\alpha_j}\alpha_k\prod_{i=1}^{n}\langle u_{ij},T_i^*T_iu_{ik}\rangle\\ &=\sum_{1\le j,k\le N}\overline{\alpha_j}\alpha_k\prod_{i=1}^{n}\langle u_{ij},P_iT_i^*T_iP_iu_{ik}\rangle \end{aligned}\tag{64}$$

where $P_i$ is the projection on the finite dimensional subspace $M_i$ spanned by $\{u_{ij},\ 1\le j\le N\}$ in $\mathcal{H}_i$ for each $i$. Since $P_iT_i^*T_iP_i$ is a positive operator in $M_i$, it follows from (62) as well as (64) that

$$\begin{aligned}\left\|\sum_{j=1}^{N}\alpha_j\bigotimes_{i=1}^{n}T_iu_{ij}\right\|^2&\le\prod_{i=1}^{n}\|P_iT_i^*T_iP_i\|\left\|\sum_{j=1}^{N}\alpha_j\bigotimes_{i=1}^{n}u_{ij}\right\|^2\\ &\le\prod_{i=1}^{n}\|T_i^*T_i\|\left\|\sum_{j=1}^{N}\alpha_j\bigotimes_{i=1}^{n}u_{ij}\right\|^2. \end{aligned}\tag{65}$$

Thus we define $T$ on the set $S$ through (63). As a consequence of (65), we obtain that $T$ extends uniquely as a linear map on the linear subspace generated by $S$. Since $S$ is total in $\mathcal{H}$ this linear extension can be closed to define a bounded operator $T$ on $\mathcal{H}$ satisfying

$$\|T\|\le\prod_{i=1}^{n}\|T_i^*T_i\|^{1/2}=\prod_{i=1}^{n}\|T_i\|.$$

For the opposite inequality, we let $0<\varepsilon<1$ be arbitrary. Choose unit vectors $u_i\in\mathcal{H}_i$ such that $\|T_iu_i\|\ge(1-\varepsilon)\|T_i\|$ for each $i$, assuming $T_i\ne 0$. Then $\bigotimes_{i=1}^{n}u_i$ is a unit vector in $\mathcal{H}$ and

$$\left\|\bigotimes_{i=1}^{n}T_iu_i\right\|=\prod_{i=1}^{n}\|T_iu_i\|\ge(1-\varepsilon)^n\prod_{i=1}^{n}\|T_i\|.$$

Thus $\|T\|\ge\prod_{i=1}^{n}\|T_i\|$. □

The operator $T$ in Theorem 5.10 shall be defined to be the tensor product of the operators $T_i$ for $1 \leq i \leq n$ and we write $T = \bigotimes_{i=1}^{n} T_i$.

For the sake of completeness, we shall state the following properties without proof seeing as these follow trivially from the definitions.

Theorem 5.11 (Properties of Tensor Products of Operators).

*i)* *The mapping* $(T_1, T_2, ..., T_n) \to T$ *is a multilinear mapping from* $\mathcal{B}(\mathcal{H}_1) \times \cdots \times \mathcal{B}(\mathcal{H}_n)$ *into* $\mathcal{B}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$.

*ii)* $ST = \bigotimes_{i=1}^{n} S_i T_i$ *and* $T^* = \bigotimes_{i=1}^{n} T_i^*$.

*iii)* *If each* $T_i$ *has a bounded inverse then* $T$ *also has a bounded inverse* $T^{-1} = \bigotimes_{i=1}^{n} T_i^{-1}$.

*iv)* *If* $T_i$ *is self adjoint, unitary, normal or a projection operator then* $T$ *is also respectively a self adjoint, unitary, normal or a projection operator.*

*v)* *If* $T_i$ *is positive for all* $1 \leq i \leq n$ *then* $T$ *is also positive.*

*vi)* $\mathrm{Tr}(T) = \prod_{i=1}^{n} \mathrm{Tr}(T_i)$.

Corollary 5.12. *If* $\rho_i$ *is a state in* $\mathcal{H}_i$ *for* $1 \leq i \leq n$ *then* $\rho = \bigotimes_{i=1}^{n} \rho_i$ *is a state in* $\bigotimes_{i=1}^{n} \mathcal{H}_i$.

Through the established results, we may now see how to combine different probability spaces into one in the context of quantum probability.

Definition 5.4 (Product of Quantum Probability Spaces). *Let* $(\mathcal{H}_i, \mathcal{P}(\mathcal{H}_i), \rho_i)$ *be a quantum probability space for each* $i = 1, 2, \cdots, n$ *Let* $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i$ *and* $\rho = \bigotimes_{i=1}^{n} \rho_i$. *Then the quantum probability space* $(\mathcal{H}, \mathcal{P}(\mathcal{H}), \rho)$ *is said to be the product of the spaces* $(\mathcal{H}_i, \mathcal{P}(\mathcal{H}_i), \rho_i)$ $1 \leq i \leq n$.

We may note that if $X_i$ is an observable in $\mathcal{H}_i$ then $X = \bigotimes_{i=1}^{n} X_i$ is an observable in $\mathcal{H}$. As in the case of subsection 4.2.2, we may also define an expectation in this space.

Definition 5.5 (Expectation in a Product of Quantum Probability Spaces). *The expectation of an observable* $X = \bigotimes_{i=1}^{n} X_i$ *in the product state* $\rho = \bigotimes_{i=1}^{n} \rho_i$

*is given by*

$$\mathbb{E}[X] = \prod_{i=1}^{n} \mathrm{Tr}(\rho_i X_i).$$

### 5.3. Symmetric and Antisymmetric Tensor products

We have seen how the tensor product may be used to construct the product of probability spaces. However, some applications require that additional structure is imposed on the tensor product such as when dealing with Fock spaces present in Section 5.4. The two most common structures generally imposed are symmetry and antisymmetry of the tensor product which we shall discuss in brief. Let $S_n$ denote the group of permutations of the set $\{1, 2, \cdots, n\}$ and for each $\sigma \in S_n$ we let $U_\sigma$ be the operator defined on the product vectors in $\mathcal{H}^{\otimes^n} = \bigotimes_{i=1}^n \mathcal{H}$ by

$$U_\sigma u_1 \otimes \cdots \otimes u_n = u_{\sigma^{-1}(1)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}$$

Then $U_\sigma$ is a scalar product preserving the map of the total set of product vectors in $\mathcal{H}^{\otimes^n}$ onto itself. Hence by Theorem 5.1 this may be uniquely extended to a unitary operator on $\mathcal{H}^{\otimes^n}$. With slight abuse of notation, we denote this operator by $U_\sigma$. Now we define the following closed subspaces.

Definition 5.6. *Let*

$$\mathcal{H}^{\otimes_s^n} = \{u \in \mathcal{H}^{\otimes^n} : U_\sigma u = u \text{ for all } \sigma \in S_n\}$$

$$\mathcal{H}^{\otimes_a^n} = \{u \in \mathcal{H}^{\otimes^n} : U_\sigma u = \varepsilon(\sigma) u \text{ for all } \sigma \in S_n\}$$

*where $\varepsilon(\sigma) = \pm 1$, the signature of $\sigma$ assumes value $1$ if $\sigma$ is even and $-1$ if $\sigma$ is odd. Then $\mathcal{H}^{\otimes_s^n}$ and $\mathcal{H}^{\otimes_a^n}$ are said to be the $n$-fold symmetric and antisymmetric tensor products of $\mathcal{H}$ respectively.*

We shall now investigate properties related to these two spaces. We start by defining two important operators related to them.

Theorem 5.13. *Let $\mathcal{S}_n$ and $\mathcal{S}_n^a$ be operators in $\mathcal{H}^{\otimes^n}$ given by*

$$\mathcal{S}_n = \frac{1}{n!} \sum_{\sigma \in S_n} U_\sigma$$

$$\mathcal{S}_n^a = \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) U_\sigma.$$

Then $\mathcal{S}_n$ and $\mathcal{S}_n^a$ are projections onto $\mathcal{H}^{\otimes_s^n}$ and $\mathcal{H}^{\otimes_a^n}$ respectively. In addition, for any $u_i, v_i \in \mathcal{H}, 1 \le i \le n$, the following relations hold.

$$\langle \mathcal{S}_n u_1 \otimes \cdots \otimes u_n, \mathcal{S}_n v_1 \otimes \cdots \otimes v_n \rangle = \langle u_1 \otimes \cdots \otimes u_n, \mathcal{S}_n v_1 \otimes \cdots \otimes v_n \rangle \tag{66}$$

$$= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^{n} \langle u_i, v_{\sigma(i)} \rangle, \tag{67}$$

$$\langle \mathcal{S}_n^a u_1 \otimes \cdots \otimes u_n, \mathcal{S}_n^a v_1 \otimes \cdots \otimes v_n \rangle = \langle u_1 \otimes \cdots \otimes u_n, \mathcal{S}_n^a v_1 \otimes \cdots \otimes v_n \rangle$$

$$= \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} \langle u_i, v_{\sigma(i)} \rangle.$$

$$= \frac{1}{n!} \det((\langle u_i, v_j \rangle)) \tag{68}$$

Proof. To start we note the elementary result $U_\sigma U_{\sigma'} = U_{\sigma\sigma'}$. From this, we may ascertain the following

$$\mathcal{S}_n^* = \frac{1}{n!} \sum_{\sigma \in S_n} U_\sigma^* = \frac{1}{n!} \sum_{\sigma \in S_n} U_{\sigma^{-1}} = \mathcal{S}_n,$$

$$\mathcal{S}_n U_\sigma = U_\sigma \mathcal{S}_n = \mathcal{S}_n^2,$$

$$\mathcal{S}_n^{a*} = \frac{1}{n!} \sum_{\sigma \in S_n} \epsilon(\sigma) U_\sigma^* = \frac{1}{n!} \sum_{\sigma \in S_n} \epsilon(\sigma^{-1}) U_{\sigma^{-1}} = \mathcal{S}_n^a,$$

$$\mathcal{S}_n^a U_\sigma = U_\sigma \mathcal{S}_n^a = \varepsilon(\sigma) \mathcal{S}_n^a, \quad \mathcal{S}_n^{a2} = \mathcal{S}_n^a.$$

This shows that both $\mathcal{S}_n$ and $\mathcal{S}_n^a$ are projections. Furthermore, by definition of $\mathcal{H}^{\otimes_s^n}$ and $\mathcal{H}^{\otimes_a^n}$ we have $\mathcal{S}_n u = u$ if and only if $u \in \mathcal{H}^{\otimes_s^n}$ and $\mathcal{S}_n^a v = v$ if and only if $v \in \mathcal{H}^{\otimes_a^n}$. This shows the first part of the theorem.

The first equality in (66) follows by noting that $\mathcal{S}_n$ is a projection. The second follows due to the following equality

$$\langle u_1 \otimes \cdots \otimes u_n, \mathcal{S}_n v_1 \otimes \cdots \otimes v_n \rangle = \frac{1}{n!} \sum_{\sigma \in S_n} \langle u_1 \otimes \cdots \otimes u_n, v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \rangle$$

$$= \frac{1}{n!} \sum_{\sigma \in S_n} \prod_i \langle u_i, v_{\sigma(i)} \rangle.$$

The case for $\mathcal{S}_n^a$ follows similarly. $\square$

It shall be of use to us to establish an orthonormal basis for each of these spaces.

THEOREM 5.14. *Let $\{e_i : i = 1, 2, \cdots\}$ be an orthonormal basis for $\mathcal{H}$. We let $u_i \in \mathcal{H}, i = 1, 2, \cdots, k$ and $r_1, \cdots, r_k$ be positive integers satisfying $\sum_{i=1}^{k} r_i = n$. Then we denote by $\bigotimes_{i=1}^{k} u_i^{\otimes r_i}$ the element $u_1 \otimes \cdots \otimes u_1 \otimes u_2 \otimes \cdots \otimes u_2 \otimes \cdots \otimes u_k \otimes \cdots \otimes u_k$ where $u_i$ is repeated $r_i$ times for each $i$.Let $\mathcal{S}_n$ and $\mathcal{S}_n^a$ be the projections given in Theorem 5.13. Then the sets*

$$\{\left(\frac{n!}{r_1! \cdots r_k!}\right)^{1/2} \mathcal{S}_n \bigotimes_{j=1}^{k} e_{i_j}^{\otimes r_j} : i_1 < i_2 < \cdots < i_k,\ r_j \geq 1 \text{ for each } 1 \leq j \leq k,$$

$$r_1 + r_2 + \cdots + r_k = n,\ k = 1, 2, \ldots, n\},$$

$$\left\{(n!)^{1/2} \mathcal{S}_n^a e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_n} : i_1 < i_2 < \cdots < i_n\right\}$$

*are orthonormal bases in $\mathcal{H}^{\otimes_s^n}$ and $\mathcal{H}^{\otimes_a^n}$ respectively. In particular, if $\dim \mathcal{H} = N < \infty$ then*

$$\dim \mathcal{H}^{\otimes_s^n} = \binom{N+n-1}{n}, \qquad \dim \mathcal{H}^{\otimes_a^n} = \begin{cases} \binom{N}{n} & \text{if } n \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. The first part of the theorem follows by (67) and (68). Indeed, substitute the given vectors in their respective expressions and we note that when the multi-indices match the resulting sum is 1 and 0 otherwise. The formulae for the dimensions may be ascertained by noting that these may be identified with the problem of placing $n$ indistinguishable balls in $N$ cells and the problem of placing $n$ indistinguishable balls in $N$ cells with the added constraint that no cell has more than one ball respectively. □

### 5.4. Fock Space

Having defined tensor products, we will now focus our attention on the Fock space. This space is of great use in quantum mechanics and we shall see its relation to the subject later on. At this juncture, we may proceed to the definition of the Fock space. The results and definitions in this section are due to [19].

DEFINITION 5.7 (Fock Spaces). *Let $\mathcal{H}$ be a Hilbert Space. Denote by $\mathcal{H}^{\otimes^n}$, $\mathcal{H}^{\otimes_s^n}$ and $\mathcal{H}^{\otimes_a^n}$ the n-fold tensor, symmetric tensor and antisymmetric tensor product* *of $\mathcal{H}$ respectively. Then the Hilbert spaces*

$$\Gamma_{fr}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes^n}, \ \Gamma_s(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes^n_s}, \ \Gamma_a(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes^n_a}$$

*are said to be the free, boson and fermion Fock Spaces respectively.*

In the above, the 0-fold product is simply $\mathbb{C}$ and the 1-fold product is the Hilbert Space $\mathcal{H}$.

It may be easily seen that there are projections which map $\Gamma_{fr}(\mathcal{H})$ onto $\Gamma_s(\mathcal{H})$ and $\Gamma_a(\mathcal{H})$. Indeed, these projections may be given respectively by

$$\mathcal{S} = \bigoplus_{n=1}^{\infty} \mathcal{S}_n, \ \mathcal{S}^a = \bigoplus_{n=1}^{\infty} \mathcal{S}_n^a \tag{69}$$

where $\mathcal{S}_n$ and $\mathcal{S}_n^a$ are the projections described in Theorem 5.13.

For ease of reference we refer to the $n$-th direct summand as the $n$-th particle subspace and any element of it an $n$-particle vector. As it will be useful later on, we shall also define the notion of a vacuum vector.

Definition 5.8 (Vacuum Vector). *In every Fock Space, the vacuum vector $\phi_0$ is given by*

$$\phi_0 = 1 \oplus 0 \oplus ... \oplus 0. \tag{70}$$

Furthermore we shall denote by $\Gamma_{fr}^0(\mathcal{H}), \Gamma_s^0(\mathcal{H})$ and $\Gamma_a^0(\mathcal{H})$ the dense linear subspace generated by all the $n$-particle vectors $n = 1, 2, ...$ in the respective Fock spaces. For each of the Fock spaces, we let $P_n$ denote the projection on the $n$-particle subspace for each $n$. Next we proceed to define a tensor product on each of the finite particle Fock spaces.

Definition 5.9. *For the spaces $\Gamma_{fr}(\mathcal{H})$, $\Gamma_s(\mathcal{H})$ and $\Gamma_a(\mathcal{H})$ respectively, we may define a tensor product $(u, v) \to u \otimes v, uv, u \wedge v$ by taking*

$$u \otimes v = \bigoplus_{n=0}^{\infty} \sum_{r+s=n} P_r u \otimes P_s v, \ uv = \mathcal{S}(u \otimes v) \text{ and } u \wedge v = \mathcal{S}^a(u \otimes v)$$

*where $\mathcal{S}$ and $\mathcal{S}^a$ are given by (69).*

Should an orthonormal basis be identified in $\mathcal{H}$, then one may easily establish orthonormal bases for each of the defined Fock Spaces.

THEOREM 5.15. *Let $\{e_j : j \in \mathbb{N}\}$ be an orthonormal basis in $\mathcal{H}$. Then the sets*

$$\{\phi_0\} \cup \{e_{i_1} \otimes \cdots \otimes e_{i_n} : i_j = 1, 2, \dots;\ j = 1, \dots, n;\ n = 1, 2, \dots\},$$

$$\{\phi_0\} \cup \{\left(\frac{n!}{r_1! \cdots r_k!}\right)^{1/2} e_{i_1}^{r_1} e_{i_2}^{r_2} \cdots e_{i_k}^{r_k} : r_1 + \cdots + r_k = n$$
$$1 \le i_1 < i_2 < \cdots < i_k,\ \text{where } n, k = 1, 2, \cdots\},$$

$$\{\phi_0\} \cup \{\ (n!)^{1/2} e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_n} : 1 \le i_1 < \cdots < i_n,\ n = 1, 2, \dots\}$$

*are orthonormal bases of $\Gamma_{fr}(\mathcal{H})$,$\Gamma_s(\mathcal{H})$ and $\Gamma_a(\mathcal{H})$ respectively.*

REMARK 5.3 (Dimensionality of the Fock space). *From the above we may deduce that, when $\mathcal{H}$ is of infinite dimension, all the Fock spaces will also be of infinite dimension. For the case when $\mathcal{H}$ is of finite dimension $N$, while the dimensions of the boson and free Fock spaces remain infinite, the direct sum in $\Gamma_a(\mathcal{H})$ terminates at the $N^{th}$ stage consequently making the dimension $\Gamma_a(\mathcal{H})$ finite. We may find this dimension to explicitly be*

$$\dim(\Gamma_a(\mathcal{H})) = \sum_{n=0}^{N} \binom{N}{n} = 2^N.$$

*This is mathematical description of what is known in quantum mechanics as the Pauli exclusion principle which we state below.*

THEOREM 5.16 (Pauli Exclusion Principle). *A fermionic state where two particles occupy the same one-particle state is identically zero meaning it cannot exist physically.*

CHAPTER 6

# Quantum Stochastic Integration

## 6.1. Creation and Annihilation Operators

Just as Brownian Motion is fundamental in stochastic integration there exists an analogous operator to this in stochastic quantum integration. This operator is defined using the creation and annihilation operators of the fermion Fock space and these are precisely what we shall be discussing in this section. The definitions and results in this section are taken from [19].

To define these operators, we need first define the following operators on the $n$ particle subspaces $\mathcal{H}^{\otimes_a^n}$.

THEOREM 6.1. *For any $u \in \mathcal{H}$ define $M_n(u) : \mathcal{H}^{\otimes_a^n} \to \mathcal{H}^{\otimes_a^{n+1}}$ by*

$$M_n(u)v_1 \wedge v_2 \cdots \wedge v_n = u \wedge v_1 \wedge v_2 \wedge \cdots \wedge v_n$$

*where $v_j \in \mathcal{H}$ for $j \in \{1, 2, \cdots, n\}$. The adjoint of $M_n(u)^*$ of $M_n(u)$ is then given by*

$$M_n^*(u)v_1 \wedge \cdots \wedge v_{n+1} = \frac{1}{n+1} \sum_{j=1}^{n+1} (-1)^{j-1} \langle u, v_j \rangle v_1 \wedge \cdots \wedge \hat{v}_j \wedge \cdots \wedge v_{n+1}$$

*where $\hat{v}_j$ denotes the omission of $v_j$ from the product.*

PROOF. We have by (68) and Definition 5.9 that the following holds

$$\langle u_1 \wedge \cdots \wedge u_{n+1},\, u \wedge v_1 \wedge \cdots \wedge v_n \rangle =$$

$$\frac{1}{n+1!} \begin{vmatrix} \langle u_1,u\rangle & \langle u_1,v_1\rangle & \cdots & \langle u_1,v_n\rangle \\ \langle u_2,u\rangle & \langle u_2,v_1\rangle & \cdots & \langle u_2,v_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_{n+1},u\rangle & \langle u_{n+1},v_1\rangle & \cdots & \langle u_{n+1},v_n\rangle \end{vmatrix}$$

$$= \frac{1}{(n+1)!} \sum_{j=1}^{n+1} (-1)^{j-1} \langle u_j, u\rangle \begin{vmatrix} \langle u_1,v_1\rangle & \cdots & \langle u_1,v_n\rangle \\ \vdots & & \vdots \\ \langle u_{j-1},v_1\rangle & \cdots & \langle u_{j-1},v_n\rangle \\ \langle u_{j+1},v_1\rangle & \cdots & \langle u_{j+1},v_n\rangle \\ \vdots & & \vdots \\ \langle u_{n+1},v_1\rangle & \cdots & \langle u_{n+1},v_n\rangle \end{vmatrix}$$

$$= \frac{1}{n+1} \sum_{j=1}^{n+1} (-1)^{j-1} \langle u_j, u\rangle \langle u_1 \wedge \cdots \wedge \widehat{u}_j \wedge \cdots \wedge u_{n+1},\, v_1 \wedge \cdots \wedge v_n \rangle.$$

□

Our goal in the next few proofs shall be to establish what are known as the canonical anticommutation relations which are fundamental when working in fermion Fock spaces.

THEOREM 6.2. *Let $u \in \mathcal{H}$, we define the linear operators $a(u)$ and $a^\dagger(u)$ on the domain of all finite particle vectors in $\Gamma_a(\mathcal{H})$ by taking*

$$a(u)\phi_0 = 0,\;\; a(u)|_{\mathcal{H}^{\otimes_a^n}} = \sqrt{n} M^*_{n-1}(u),$$

$$a^\dagger(u)\phi_0 = u,\;\; a^\dagger(u)|_{\mathcal{H}^{\otimes_a^n}} = \sqrt{n+1} M_n(u),\;\; n \in \mathbb{N},$$

*where $M_n(u)$ and $M^*_n(u)$ are as defined in Theorem 6.1. Then*

$$[a(u)a^\dagger(v) + a^\dagger(v)a(u)]\phi = \langle u, v\rangle \phi$$

*for every finite particle vector $\phi \in \Gamma_a(\mathcal{H})$.*

PROOF. By definition of the operators as well as Theorem 6.1 the following holds

$$a^\dagger(v)w_1 \wedge \cdots \wedge w_n = \sqrt{n+1}\, v \wedge w_1 \wedge \cdots \wedge w_n,$$

$$(71) \quad a(u)a^\dagger(v)w_1 \wedge \cdots \wedge w_n = \langle u, v\rangle w_1 \wedge \cdots \wedge w_n$$

$$+\sum_{j=1}^{n}(-1)^j\langle u, w_j\rangle v \wedge w_1 \wedge \cdots \wedge \widehat{w}_j \wedge \cdots \wedge w_n,$$

$$a(u)w_1 \wedge \cdots \wedge w_n = n^{-\frac{1}{2}}\sum_{j=1}^{n}(-1)^{j-1}\langle u, w_j\rangle w_1 \wedge \cdots \wedge \widehat{w}_j \wedge \cdots \wedge w_n,$$

$$(72) \quad a^\dagger(v)a(u)w_1 \wedge \cdots \wedge w_n = \sum_{j=1}^{n}(-1)^{j-1}\langle u, w_j\rangle v \wedge w_1 \wedge \cdots \wedge \widehat{w}_j \wedge \cdots \wedge w_n.$$

The addition of (71) and (72) gives us the result. $\square$

We shall now show the existence of the annihilation and creation operators in the fermion Fock space $\Gamma_a(\mathcal{H})$.

THEOREM 6.3. *In the fermion Fock space $\Gamma_a(\mathcal{H})$ there exists a unique family $\{a(u), a^\dagger(u) : u \in \mathcal{H}\}$ of bounded operators which satisfy the following*

*i)* $a(u)\phi_0 = 0$ *and* $a^\dagger(u)\phi_0 = u$ *where* $\phi_0$ *is the vacuum vector and* $u$ *is a 1-particle vector.*

*ii)* $a(u)v_1 \wedge \cdots \wedge v_n = n^{-\frac{1}{2}}\sum_{j=1}^{n}(-1)^{j-1}\langle u, v_j\rangle v_1 \wedge \cdots \wedge \widehat{v}_j \wedge \cdots \wedge v_n$,
$a^\dagger(u)v_1 \wedge \cdots \wedge v_n = (n+1)^{1/2}u \wedge v_1 \wedge \cdots \wedge v_n$ *for all* $v_j \in \mathcal{H},\ 1 \le j \le n,\ n = 1, 2, \ldots$

*iii)* $a^\dagger(u) = a(u)^*$ *for all* $u \in \mathcal{H}$.

PROOF. First define $a(u)$ and $a^\dagger(u)$ on the dense linear subspace $\Gamma_a^0(\mathcal{H})$ of all finite particle vectors in $\Gamma_a(\mathcal{H})$ through the use of Theorem 6.2. Also by Theorem 6.1, $a(u)$ and $a^\dagger(u)$ are adjoint to each other on $\Gamma_a^0(\mathcal{H})$. As a result of Theorem 6.2 we have for any $\phi \in \Gamma_a^0(\mathcal{H})$

$$\|a(u)\phi\|^2 + \|a^\dagger(u)\phi\|^2 = \langle \phi, a^\dagger(u)a(u) + a(u)a^\dagger(u)\phi\rangle = \|u\|^2\|\phi\|^2.$$

In particular,

$$\|a(u)\phi\| \le \|u\|\|\phi\|, \quad \|a^\dagger(u)\phi\| \le \|u\|\|\phi\|, \quad \phi \in \Gamma_a^0(\mathcal{H}).$$

Hence we can extend $a(u)$ and $a^\dagger(u)$ uniquely to bounded operators on $\Gamma_a(\mathcal{H})$ and with slight abuse of notation we denote them by the same symbols so that

$$\|a(u)\| \leq \|u\|, \quad \|a^\dagger(u)\| \leq \|u\|.$$

The rest of the proof follows trivially. □

We call $a(u)$ and $a^\dagger(u)$ the annihilation and creation operators respectively. Finally, we end the discussion on these operators by proving the canonical anti-commutation relations we had previously mentioned.

THEOREM 6.4 (Canonical Anticommutation Relations). *The operators $a(u)$ and $a^\dagger(u)$ given in Theorem 6.3 satisfy the following properties*

*i)* $a(u)a(v) + a(v)a(u) = 0 = a^\dagger(u)a^\dagger(v) + a^\dagger(v)a^\dagger(u)$.

*ii)* $a(u)a^\dagger(v) + a^\dagger(v)a(u) = \langle u, v\rangle I$ *for all* $u, v \in \mathcal{H}$.

PROOF. By Theorem 6.3*i* as well as 6.3*ii* the vectors $\{a^\dagger(u)a^\dagger(v)+a^\dagger(v)a^\dagger(u)\}\phi$ and $\{a^\dagger(u)a^\dagger(v) + a^\dagger(v)a^\dagger(u)\}w_1 \wedge \cdots \wedge w_n$ are respectively scalar multiples of $u \wedge v + v \wedge u$ and $(u \wedge v + v \wedge u) \wedge w_1 \wedge \cdots \wedge w_n$ which vanish. Using the fact that finite particle vectors are dense in $\Gamma_a(\mathcal{H})$ as well as the fact that the operators $a^\dagger(u)$ are bounded for every $u \in \mathcal{H}$ we obtain the second relation in *i*). The first relation in *i*) follows from the second by taking adjoints. □

In the coming sections, we shall make use of Von Neumann algebras. Hence we give their definition below.

DEFINITION 6.1 (Von Neumann Algebra). *A Von Neumann Algebra is a subalgebra of $\mathcal{B}(\mathcal{H})$ which is closed under involution and the weak order topology.*

Let $J : \mathcal{H} \to \mathcal{H}$ be a conjugation on $\mathcal{H}$. We define the operator $\Psi(z) = a^\dagger(z) + a(Jz)$. From these operators we may obtain $\mathcal{C}_1$, the algebra generated by the set of operators $\{\Psi(z) : z \in \mathcal{H}\}$. By observing that $\Psi(z)^* = \Psi(Jz)$ we obtain that $\mathcal{C}_1$ is self adjoint and hence its weak closure $\mathcal{C}$ will be a Von Neumann Algebra. We recall that $\mathcal{C}_1$ is self adjoint if it satisfies $A \in \mathcal{C}_1 \iff A^* \in \mathcal{C}_1$

## 6.2. Probability Gauge Spaces and Non Commutative $L^p$ Spaces

While a classic probability space was sufficient when dealing with stochastic integration, we require a generalization when dealing with quantum stochastic integration. This is precisely achieved by the regular probability gauge spaces we shall discuss in this section. We start off by giving the definition which is taken from [28].

Definition 6.2 (Probability Gauge Space). *A regular probability gauge space is a triple $(\mathcal{H}, \mathfrak{U}, m)$ where $\mathcal{H}$ is a complex Hilbert Space, $\mathfrak{U}$ is a Von Neumann Algebra and the gauge $m$ is a faithful, central and normal state on $\mathfrak{U}$. That is, $m : \mathfrak{U} \to \mathbb{C}$ is a linear map which satisfies the following*

*i) $m(I) = 1$; If $A \in \mathfrak{U}$ such that $A \geq 0$ then $m(A) \geq 0$.*

*ii) If $A \geq 0$ and $m(A) = 0$ then $A = 0$.*

*iii) For all $A_1, A_2 \in \mathfrak{U}$, $m(A_1A_2) = m(A_2A_1)$.*

*iv) In $\{P_\alpha\}$ is a family of mutually orthogonal projections in $\mathfrak{U}$ then*

$$m\left(\sum_\alpha P_\alpha\right) = \sum_\alpha m(P_\alpha).$$

In order to recover the probability space we use in the commutative setting, we need only take the special case $\mathfrak{U} = L^\infty(\Omega, \mathbb{P})$ and $m(f) = \int_\Omega f d\mathbb{P}$ and we treat equality in this context as meaning equality almost everywhere.

Property $i$) follows trivially by the monotonicity of the integral as well as by noting that $\int_\Omega 1 d\mathbb{P} = \mathbb{P}(\Omega) = 1$. Property $iii$) follows by the commutativity of the underlying algebra.

For property $ii$), we let $f \in \mathfrak{U}$ satisfy the given conditions. We let $E_n = \{x \in \Omega : f(x) \geq \frac{1}{n}\}$, then the following holds

$$0 = \int_\Omega f d\mathbb{P} \geq \int_{E_n} f d\mathbb{P} \geq \frac{1}{n}\mathbb{P}(E_n).$$

Thus $\mathbb{P}(E_n) = 0$ for all $n$ so that $\mathbb{P}(\{x : f(x) > 0\}) = \mathbb{P}(\cup_n E_n) = 0$ showing $f$ is 0 almost everywhere as desired.

Finally, to verify property $iv$), we first note that projections in $L^\infty$ are indeed indicator functions that is $P_\alpha = \mathbb{1}_{E_\alpha}$ for some measurable set $E_\alpha$. Secondly note the trivial fact that, as the projections are orthogonal, $\sum_\alpha \mathbb{1}_{E_\alpha} = \mathbb{1}_{\cup_\alpha E_\alpha}$. Hence, using countable additivity of $\mathbb{P}$, the following is evident

$$m(\sum_\alpha P_\alpha) = \int \mathbb{1}_{\bigcup_\alpha E_\alpha} d\mathbb{P} = \mathbb{P}\Big(\bigcup_\alpha E_\alpha\Big) = \sum_\alpha \mathbb{P}(E_\alpha) = \sum_\alpha \int_\Omega \mathbb{1}_{E_\alpha}$$
$$= \sum_\alpha m(P_\alpha)$$

hence showing that *iv*) holds.

For the purposes of the quantum stochastic integral, we shall require a particular probability gauge space which we shall be discussing next. The proof of this result is taken from [11].

THEOREM 6.5. *Let $m : \mathcal{C} \to \mathbb{C}$ be the map defined as $m(A) = \langle \phi_0, A\phi_0 \rangle$, Then the space $(\Gamma_a(\mathcal{H}), \mathcal{C}, m)$ is a probability gauge space.*

PROOF. Define the inner product $\langle x, y\rangle' = \langle Jx, y\rangle$ as this will ease notation in the proof. We have $m(I) = \langle \phi_0, \phi_0\rangle = 1$ and clearly, for $A$ positive, we have $m(A) > 0$ and so *i*) holds.

Let $A$ be a strictly positive operator then it is Hermitian. Hence, through continuous function calculus, we may find an operator $B$ such that $A = B^2$. Consequently the following holds

$$m(A) = m(B^2) = \langle \phi_0, B^2\phi_0\rangle = \langle B\phi_0, B\phi_0\rangle = \|B\phi_0\|^2 > 0$$

whereby the inequality follows since $\phi_0$ is separating. Thus *ii*) holds by the contrapositive.

Next we prove that $m$ satisfies *iii*). We note that from the anticommutation relations one may easily derive that $a(Jx)\Psi(y) + \Psi(y)a(Jx) = \langle x, y\rangle' I$. This relation, along with the fact that $a(Jx)\phi_0 = 0$ leads us to deduce the following

$$a(Jx)\Psi(x_1)\Psi(x_2)\cdots\Psi(x_n)\phi_0 = \sum_{j=1}^{n}(-1)^{j-1}\langle x, x_j\rangle'\Psi(x_1)\cdots\hat{\Psi}(x_j)\cdots\Psi(x_n)\phi_0$$

where $\hat{\Psi}(x_j)$ denotes omission of the factor. Furthermore, by noting that $a^{\dagger *}(x)\phi_0 = 0$ the following is evident

$$\begin{aligned}\langle \phi_0, \Psi(x)\Psi_1(x_1)\cdots\Psi_n(x_n)\phi_0\rangle &= \langle \phi_0, (a^\dagger(x) + a(Jx))\Psi(x_1)\cdots\Psi(x_n)\phi_0\rangle \\ &= \langle \phi_0, a(Jx)\Psi(x_1)\cdots\Psi(x_n)\phi_0\rangle \\ &= \sum_{j=1}^{n}(-1)^{j-1}\langle x, x_j\rangle'\langle \phi_0, \Psi(x_1)\cdots\hat{\Psi}(x_j)\cdots\Psi(x_n)\phi_0\rangle.\end{aligned}$$

By a similar argument we obtain that

$$\begin{aligned}\Psi(x_1)\cdots\Psi(x_n)\Psi(x)\phi_0 &= \Psi(x_1)\cdots\Psi(x_n)a^\dagger(x)\phi_0 \\ &= \sum_{j=1}^{n}(-1)^{n-j}\langle x, x_j\rangle'\Psi(x_1)\cdots\hat{\Psi}(x_j)\cdots\Psi(x_n)\phi_0 \\ &\pm a^\dagger(x)\Psi(x_1)\cdots\Psi(x_n)\phi_0.\end{aligned}$$

Hence,

(73)

$$\langle \phi_0, \Psi(x_1)\cdots\Psi(x_n)\Psi(x)\phi_0\rangle = \sum_{j=1}^{n}(-1)^{n-j}\langle x, x_j\rangle'\langle \phi_0, \Psi(x_1)\cdots\hat{\Psi}(x_j)\cdots\Psi(x_n)\phi_0\rangle.$$

If $n$ is odd then $(-1)^{n-j} = (-1)^{j-1}$ and hence

$$\langle \phi_0, \Psi(x)\Psi(x_1)\cdots\Psi(x_n)\phi_0\rangle = \langle \phi_0, \Psi(x_1)\cdots\Psi(x_n)\Psi(x)\phi_0\rangle. \tag{74}$$

Else, if $n$ is even, then $\Psi(x_1)\cdots\Psi(x_n)\Psi(x)\psi_0$ and $\Psi(x)\Psi(x_1)\cdots\Psi(x_n)\psi_0$ are both tensors whose non zero homogenous components are of odd rank and hence orthogonal to $\phi_0$. This would give us that both sides of (74) are 0 and hence equal. Hence (74) holds for every $\Psi \in \mathcal{C}_1$. Now, both sides of (74) are continuous functions of $\Psi$ in the weak order topology and thus (74) holds for all $\Psi \in \mathcal{C}$. Hence *iii*) holds.

Finally, we shall show that *iv*) holds. Let $\{P_n\}_{n\in\mathbb{N}}$ be a family of mutually orthogonal projections. Due to the mutual orthogonality of the projections, the following is evident

$$m\left(\sum_{n=1}^{\infty} P_n\right) = \langle \phi_0, \left(\sum_{n=1}^{\infty} P_n\right)\phi_0\rangle$$
$$= \sum_{n=1}^{\infty}\langle \phi_0, P_n\phi_0\rangle$$
$$= \sum_{n=1}^{\infty} m(P_n).$$

Thus we have shown that *iv*) holds, completing the proof. □

Now we shall show how the operator $\Psi$ may be seen as a non-commutative analogue of Brownian Motion in the Fermion Fock space. To this end, we shall define the notions of expectation and covariance in the context of a probability gauge space.

DEFINITION 6.3. *Let $X$ be an observable in a probability gauge space. Then*

$$\mathbb{E}[X] = m(X) \text{ and } \mathrm{Cov}(X_1, X_2) = m(X_1^* X_2) - m(X_1)m(X_2).$$

We calculate these quantities for $\Psi(z)$ where $z \in \mathcal{H}$. Firstly we have that

$$\mathbb{E}[\Psi(z)] = m(\Psi(z))$$
$$= \langle \phi_0, \Psi(z)\phi_0\rangle$$
$$= \langle \phi_0, \Psi(z)\phi_0\rangle$$
$$= \langle \phi_0, a^\dagger(z)\phi_0\rangle + \langle \phi_0, a(Jz)\phi_0\rangle$$

Now $a(Jz)\phi_0 = 0$ by definition and $a^\dagger(z)\phi_0$ is orthogonal to $\phi_0$ and hence $\mathbb{E}[\Psi(z)] = 0$.

Next we shall investigate the covariance. Let $z_s, z_t \in \mathcal{H}$ where $s, t \in \mathbb{R}^+$.

$$\begin{aligned}
\operatorname{Cov}(\Psi(z_s), \Psi(z_t)) &= m(\Psi(z_s)\Psi(z_t)) \\
&= \langle \phi_0, \Psi(z_s)\Psi(z_t)\phi_0 \rangle \\
&= \langle \phi_0, (a(Jz_s) + a^\dagger(z_s))(a(Jz_t) + a^\dagger(z_t))\phi_0 \rangle \\
&= \langle \phi_0, a(Jz_s)a(Jz_t)\phi_0 \rangle + \langle \phi_0, a(Jz_s)a^\dagger(z_t)\phi_0 \rangle \\
&\quad + \langle \phi_0, a^\dagger(z_s)a(Jz_t)\phi_0 \rangle + \langle \phi_0, a^\dagger(z_s)a^\dagger(z_t)\phi_0 \rangle. \qquad (75)
\end{aligned}$$

Using the fact that $a(u)\phi_0 = 0$ for any $u \in \mathcal{H}$ we have that the first and third term in (75) cancel out. Also for $v \in \mathcal{H}$, $a^\dagger(u)a^\dagger(v)\phi_0$ is orthogonal to $\phi_0$ and hence the fourth term cancels out as well. Also

$$\begin{aligned}
a(Jz_s)a^\dagger(z_t)\phi_0 &= (a(Jz_s)a^\dagger(z_t) + a^\dagger(z_t)a(Jz_s))\phi_0 \\
&= \langle Jz_s, z_t \rangle \phi_0 \qquad (76)
\end{aligned}$$

by the anticommutation relations. Substituting (76) in (75) we obtain

$$\operatorname{Cov}(\Psi(z_s), \Psi(z_t)) = \langle Jz_s, z_t \rangle.$$

For a Hilbert Space $\mathcal{H}$ satisfying $\langle z_s, z_t \rangle = \min(s, t)$ we have that $(\Psi_t)_{t \in \mathbb{R}^+}$ is a sequence of observables having mean 0 and the same covariance structure as Brownian motion. Hence these observables may be thought of as a non-commutative analogue of classical Brownian Motion. It is easily verified that for $\mathcal{H} = L^2(\mathbb{R}^+, dt)$ and $x_t = \mathbb{1}_{[0,t]}$ for $t \in \mathbb{R}^+$, the required covariance structure is attained. We shall next discuss briefly the notion of non commutative $L^p$ spaces.

Just as $L^p$ spaces are defined through the use of an integral with respect to some measure, we may define their non commutative analogues through the use of gauges.

Definition 6.4 (Non-Commutative $L^p$ Spaces). *For $1 \le p < \infty$, $L_p(\mathcal{C})$ is the completion of $\mathcal{C}$ with respect to the norm $\|u\|_p = m(|u|^p)^{\frac{1}{p}} = \langle \phi_0, (u{*}u)^{\frac{p}{2}}\phi_0 \rangle^{\frac{1}{p}}$. For the case when $p = \infty$ we must take the completion with respect to the operator norm.*

Many results of the commutative $L^p$ spaces carry over to the non commutative ones. For our purposes we shall require the fact that Hölder's inequality holds in this space. We state the result taken from [16] without proof.

THEOREM 6.6. *Let $(\mathcal{H}, \mathcal{C}, m)$ be a regular probability gauge space with corresponding $L^p$ spaces defined as in Definition 6.4. If $A_1 \in L^p$ and $A_2 \in L^q$ where $\frac{1}{p} + \frac{1}{q} = 1$ then*

$$\|A_1\|_p = \sup\{|m(A_1 S)| : S \in \mathfrak{U}, \|S\|_q \leq 1\}.$$

*and*

$$|m(A_1 A_2)| \leq \|A_1\|_p \|A_2\|_q \tag{77}$$

Clearly (77) is Hölder's inequality for gauge spaces. Applying this inequality to $m(|A|^p)$, we obtain that

$$m(|A|^p) \leq m(|A|^{\frac{pq}{p}})^{\frac{p}{q}} m(I) = \|A\|_q^p .$$

This shows that the classic inequality $\|\cdot\|_p \leq \|\cdot\|_q$ for $1 \leq p \leq q \leq \infty$ holds for the non-commutative $L^p$ spaces as well.

Just as conditional expectation was used for the classical stochastic integral, we shall also require it for the non-commutative one. To this end, we shall establish a non-commutative Radon-Nikodym theorem. We quote the result from [28] without proof.

THEOREM 6.7. *Let $(\mathcal{H}, \mathcal{C}, m)$ be a regular probability gauge space and let $\mathcal{B}$ be a Von Neumann subalgebra of $\mathcal{C}$. Then $(\mathcal{H}, \mathcal{B}, m|_{\mathfrak{U}_1})$ is also a regular probability gauge space. Let $L^p(\mathcal{C})$ and $L^p(\mathfrak{B})$ for $1 \leq p \leq \infty$ be the completions of $\mathcal{C}$ and $\mathcal{B}$ respectively. Then if $A \in L^p(\mathcal{C})$ there exists a unique $\hat{A} \in L^p(\mathcal{B})$ such that*

$$m(AA') = m(\hat{A}A') \tag{78}$$

*for all $A' \in L_1^\infty$.*

REMARK 6.1. *By Hölder's inequality we may note that Theorem 6.7 remains valid for all $A' \in L^q(\mathcal{B})$ where $q$ satisfies $\frac{1}{p} + \frac{1}{q} = 1$.*

Definition 6.5 (Non Commutative Conditional Expectation). *The map $A \mapsto \hat{A}$ given in Theorem 6.7 is said to be the conditional expectation with respect to $\mathcal{B}$ and is denoted by $m(\cdot|\mathcal{B})$ and enjoys the following properties*

*i) $m(\cdot \mid \mathcal{B})$ is a contraction from $L^p(\mathcal{C})$ onto $L^p(\mathcal{B})$ for all $1 \leqslant p \leqslant \infty$.*

*ii) $m(\cdot \mid \mathcal{B})$ is positivity preserving.*

*iii) $m(vu \mid \mathcal{B}) = vm(u \mid \mathcal{B})$ for all $u \in L^p(\mathcal{C}), v \in L^{p'}(\mathcal{B})$.*

*iv) If $\mathcal{B}_1 \subseteq \mathcal{B}_2$, then $m\left(m\left(\cdot \mid \mathcal{B}_2\right) \mid \mathcal{B}_1\right) = m\left(\cdot \mid \mathcal{D}_1\right)$.*

*v) $m(\cdot \mid \mathcal{B})$ on $L^2(\mathcal{C})$ is the projection onto the subspace $L^2(\mathcal{B})$.*

We shall also give the definition of a quantum martingale as this shall be useful when defining the integral later on. We let $\mathcal{H} = L^2(\mathbb{R}^+, ds)$ where $ds$ is the Lebesgue measure and let $J$ be the complex conjugation on $\mathcal{H}$. For a given $t \in \mathbb{R}^+ \cup \{0\}$ we take $\mathcal{C}_t$ to be the Von Neumann subalgebra of $\mathcal{C}$ generated by the operators $\Psi(u)$ whereby $u \in L^2(\mathbb{R}^+, ds)$ having $\operatorname{ess\,supp} u \subseteq [0, t]$. Clearly, $\mathcal{C}_s \subseteq \mathcal{C}_t$ and for $0 \leq s \leq t$ we have that $\mathcal{C}$ is generated by $\mathcal{C}_t$ for $t \in \mathbb{R}^+$. For ease of notation, we shall denote $m(\cdot|\mathcal{C}_s)$ by $M_s(\cdot)$ for $s \in \mathbb{R}^+$.

Definition 6.6 (Quantum Martingale). *An $L^p$ martingale adapted to the family $\{\mathcal{C}_t : t \in \mathbb{R}^+\}$ is a collection $\{X_t : t \in \mathbb{R}^+\}$ of operators on $\Gamma_a(\mathcal{H})$ such that $X_t \in L^p(\mathcal{C}_t)$ for $t \in \mathbb{R}^+$ and $M_s(X_t) = X_s$ for any $0 \leq s \leq t$.*

## 6.3. The Itô-Clifford Integral

Having defined the necessary spaces and structures, we shall now proceed to the construction of the Itô-Clifford integral. While this is beyond the scope of the dissertation, this integral may be used to solve quantum stochastic differential equations. These may be applied to, among others, non-equilibrium thermo field dynamics. The interested reader is directed to [14] for further information on this application. The results in this section are from [4].

As in the case of the classical Itô Integral, the notion of an adapted process plays a crucial role in quantum stochastic calculus. We proceed to the definition.

Definition 6.7 (Quantum Adapted Process). *Let $0 \leq t_0 \leq t$. A map $f : [t_0, t] \to L^1(\mathcal{C})$ is said to be a quantum adapted process if $f(s) \in L^1(\mathcal{C}_s)$ for all*

*$s \in [t_0, t]$. An $L^p$ quantum adapted process is a quantum adapted process $f$ such that $f(s) \in L^p(\mathcal{C}_s)$ for all $s$.*

Hereafter we shall deal with these types of processes exclusively, hence we shall drop the adjectives and refer to the maps in Definition 6.7 as processes. As in the Itô integral, the notion of a simple process plays a central role in defining the integral. We proceed with their definition.

Definition 6.8 (Simple Process). *A process $h$ on $[t_0, t]$ is said to be simple if it assumes the form*

$$h = \sum_{k=1}^{n} h_{k-1} \chi_{[t_{k-1}, t_k)}$$

*on $[t_0, t)$ for some $t_0 \leq t_1 \leq \cdots \leq t_n = t$ and $h_k \in L^1(\mathcal{C}), 1 \leq k \leq n$.*

Since $h$ is a process, we have that $h(s) = h_k \in L^1(\mathcal{C})$ for all $t_{k-1} \leq s < t_k$. That is, $h_{k-1} \in L^1(\mathcal{C}_{t_{k-1}})$. Let $u \in L^2_{\text{loc}}(\mathbb{R}^+, ds)$, the space of locally square integrable functions on the positive real line with respect to the Lebesgue measure. We shall denote $\Psi(u\chi_{[0,t]})$ by $\Psi_t$. Clearly, $\Psi_t$ is a self adjoint operator and satisfies $\Psi_t \in L^\infty(\mathcal{C}_t)$ for all $t \in \mathbb{R}^+$. Furthermore, due to the anticommutation relations, we have that for $0 \leq s \leq t$

$$(\Psi_t - \Psi_s)^2 = \Psi(u\chi_{[s,t]})^2 = \int_s^t |u(x)|^2 dx. \tag{79}$$

We shall now proceed to the definition of the integral in case of simple processes.

Definition 6.9 (Itô-Clifford Integral for Simple Processes). *Let $h$ be a simple process as in Definition 6.8 then the Itô-Clifford integral of $h$ over $[t_0, t]$ is given by*

$$I(h) = \int_{t_0}^{t} h(s) d\Psi_s = \sum_{k=1}^{n} h_{k-1}(\Psi_{t_k} - \Psi_{t_{k-1}}) = \sum_{k=1}^{n} h_{k-1} \Delta \Psi_k.$$

Trivially, we have that $I(h) \in L^1(\mathcal{C}_t)$ and is independent of the decomposition of $h$ as a sum of step-functions. The next lemma is of great importance, both for showing a particular property of the integral and also for extending it for more general processes later on.

LEMMA 6.1. *Let $\{X_t : t \in \mathbb{R}^+\}$ be an $L^p$ martingale and let $0 \le s \le t$. Then $m(f(X_t - X_s)g) = 0$ for any $f \in L^q(\mathcal{C}_s), g \in L^r(\mathcal{C}_s)$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. In particular, if $f \in L^q(\mathcal{C}_s)$ and $g \in L^r(\mathcal{C}_s)$ where $\frac{1}{q} + \frac{1}{r} = 1$ we have that $m(f(\Psi_t - \Psi_s)g) = 0$.*

PROOF. We have

$$\begin{aligned} m(f(X_t - X_s)g) &= m(gf(X_t - X_s)) \\ &= m(gfM_s(X_t - X_s)) \\ &= 0 \end{aligned}$$

whereby the first equality follows from the fact that $m$ is central in $\mathcal{C}$ and the second equality follows by the definition of conditional expectation given in (78). The last part of the lemma follows due to the fact that $\Psi_t \in L^\infty$ for all $t$. $\square$

We shall now discuss some important properties this integral for simple processes.

THEOREM 6.8. *The Itô-Clifford Integral of simple processes satisfies the following:*

*i) $\int_{t_0}^t (\alpha h(s) + \beta g(s))d\Psi_s = \alpha \int_{t_0}^t h(s)d\Psi_s + \beta \int_{t_0}^t g(s)d\Psi_s$ for simple processes $g, h$ and $\alpha, \beta \in \mathbb{C}$.*
*ii) $m(\int_{t_0}^t h(s)d\Psi_s) = 0$ for simple $h$.*
*iii) If $h$ is a simple $L^2$ process on $[t_0, t]$ that is $h(s) \in L^2(\mathcal{C}_s)$ for all $t_0 \le s \le t$ then $\int_{t_0}^t h(s)d\Psi_s \in L^2(\mathcal{C}_t)$ and $|| \int_{t_0}^t h(s)d\Psi_s||_2^2 = \int_{t_0}^t ||h(s)||_2^2 |u(s)|^2 ds$.*

PROOF. Linearity of the integral follows trivially.

For *ii*), consider

$$\begin{aligned} m\left(\int_{t_0}^t h(s)d\Psi_s\right) &= \sum_{k=1}^n m(h_{k-1}(\Psi_{t_k} - \Psi_{t_{k-1}})) \\ &= 0. \end{aligned}$$

Whereby, in the above, $t_0 \le t_1 \le \cdots \le t_n$ is a suitable partition of $[t_0, t]$ and $h_{k-1} \in L^1(\mathcal{C}_{t_{k-1}})$. The final equality follows as a result of Lemma 6.1.

For *iii*),

$$\left\| \int_{t_0}^{t} h d\Psi_s \right\|_2^2 = m\left( \left( \sum_{k=1}^{n} h_{k-1}\Delta\Psi_k \right)^* \left( \sum_{j=1}^{n} h_{j-1}\Delta\Psi_j \right) \right)$$
$$= \sum_{k,j} m(\Delta\Psi_k h_{k-1}^* h_{j-1}\Delta\Psi_j).$$

By Lemma 6.1 the off diagonal terms vanish. Then, for a general non-zero term of the summation, the following holds

$$\begin{aligned} m(\Delta\Psi_k h_{k-1}^* h_{k-1}\Delta\Psi_k) &= m(h_{k-1}^* h_{k-1}(\Delta\Psi_k)^2) \\ &= m(h_{k-1}^* h_{k-1}) \int_{t_{k-1}}^{t_k} |u(s)|^2 ds \\ &= \int_{t_0}^{t} m(h_{k-1}^* h_{k-1})\chi_{[t_{k-1},t_k]}(s)|u(s)|^2 ds \end{aligned}$$

wherein the second equality follows from (79). Summing over $k$ gives the required result. □

We may call 6.8*iii* the isometry property of the Itô-Clifford integral. The analogy with the Itô isometry of classical stochastic integration is clear.

In the next few results, we shall be extending this integral for more general processes. For the sake of simplicity, we shall denote the measure $|u(s)|^2 ds$ on $\mathbb{R}^+$ by $d\mu$.

THEOREM 6.9. *Let $g$ be a continuous $L^2$ process on $[t_0, t]$ that is, $s \to g(s)$ is continuous from $[t_0, t]$ into $L^2(\mathcal{C})$. Then, for any $\epsilon > 0$, there exists a simple $L^2$ process $h$ on $[t_0, t]$ such that*

$$\int_{t_0}^{t} \|g(s) - h(s)\|_2^2 \, d\mu(s) < \epsilon.$$

PROOF. Let $\epsilon' > 0$ be given. As $g$ is continuous on a bounded interval $[t_0, t]$ then it is uniformly continuous on that interval. Then it must be the case that there is $t_0 \le t_1 \le \cdots \le t_n = t$ such that

$$\|g(s) - g(t_{k-1})\|_2^2 < \epsilon'$$

whenever $t_{k-1} \leq s \leq t_k$. By taking $h = \sum_{k=1}^{n} g(t_{k-1})\chi_{[t_{k-1},t_k)}$ we may observe that $\|g(s) - h(s)\|_2^2 < \epsilon'$ for all $t_0 \leq s \leq t$ and hence

$$\int_{t_0}^{t} \|g(s) - g(t_{k-1})\|_2^2 \, d\mu(s) < \epsilon' \int_{t_0}^{t} d\mu(s)$$

giving us the result. □

We state the next lemma without proof.

LEMMA 6.2. *For a fixed $X \in L^2(\mathcal{C})$, $s \to M_s(X)$ is a continuous map from $\mathbb{R}^+$ into $L^2(\mathcal{C})$.*

For the final few steps, we shall need to introduce two spaces. The first is the space $L^2([t_0,t], d\mu; L^2(\mathcal{C}))$ which is the complex Hilbert Space of $L^2(\mathcal{C})$-valued measurable maps on $[t_0,t]$ which are square integrable with respect to $d\mu$. The second space is the Hilbert Space tensor product, $L^2([t_0,t], d\mu) \otimes L^2(\mathcal{C})$. The latter is the completion of the algebraic tensor product $C([t_0,t]) \otimes L^2(\mathcal{C})$ where $C([t_0,t])$ denotes the space of complex-valued functions on $[t_0,t]$. What is of interest to us is that these two spaces are isomorphic. This may be easily verified by identifying the element $\phi \otimes X$ with $\phi(\cdot)X$ where $\phi \in L^2([t_0,t], d\mu)$ and $X \in L^2(\mathcal{C})$.

Naturally, not every $f \in L^2([t_0,t], d\mu; L^2(\mathcal{C}))$ is a process, seeing as not every $f$ satisfies $f(s) \in L^2(\mathcal{C}_s)$ $\mu$ *a.e.* We denote by $\mathfrak{h}[t_0,t]$ the space of processes in $L^2([t_0,t], d\mu; L^2(\mathcal{C}))$. We now prove some properties related to this space.

THEOREM 6.10. *$\mathfrak{h}[t_0,t]$ is a closed subspace of $L^2([t_0,t], d\mu; L^2(\mathcal{C}))$. In particular, $\mathfrak{h}[t_0,t]$ is a Hilbert Space.*

PROOF. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathfrak{h}[t_0,t]$ such that $f_n \to f$ in $L^2([t_0,t], d\mu; L^2(\mathcal{C}))$. By passing to a subsequence we may assume that $f_n(s) \to f(s)$ $\mu$ *a.e.* in $L^2(\mathcal{C})$, this is justified by Corollary 2.1 in [16]. However, $f_n(s) \in L^2(\mathcal{C}_s)$ $\mu$ *a.e* so that $f(s) \in L^2(\mathcal{C}_s)$ $\mu$ *a.e* showing $f \in \mathfrak{h}[t_0,t]$ as required. Hence, $\mathfrak{h}[t_0,t]$ is a Hilbert Space as a closed subspace of a Hilbert Space. □

THEOREM 6.11. *Let $f \in \mathfrak{h}[t_0,t]$. Then $f$ may be approximated arbitrarily closely in $\mathfrak{h}[t_0,t]$ using simple processes in $\mathfrak{h}[t_0,t]$.*

PROOF. Let $\epsilon > 0$ be given. Due to the isomorphism between the two spaces, we are guaranteed the existence of $n \in \mathbb{N}$, $\phi_j \in C[t_0, t]$ and $X_j \in L^2(\mathcal{C})$ for $1 \leq j \leq n$ such that

$$\int_{t_0}^{t} \left\| f(s) - \sum_{j=1}^{n} \phi_j(s) X_j \right\|_2^2 d\mu(s) < \epsilon^2.$$

As $f$ is a process and $M_s : L^2(\mathcal{C}) \to L^2(\mathcal{C}_s)$ is a contraction, the following holds

$$\begin{aligned} \int_{t_0}^{t} \left\| f(s) - \sum_{j=1}^{n} \phi_j(s) M_s(X_j) \right\|_2^2 d\mu(s) \\ = \int_{t_0}^{t} \left\| M_s\Big( f(s) - \sum_{j=1}^{n} \phi_j(s) X_j \Big) \right\|_2^2 d\mu(s) \\ < \epsilon^2 \end{aligned}$$

Through Lemma 6.2, each $\phi_j(\cdot)M(X_j)$ is a continuous $L^2$ process. In turn, this means that they belong to $\mathfrak{h}[t_0, t]$ and by Theorem 6.9 we may assert that they may be approximated by simple processes. □

We are now in a position to construct the integral $I(f)$ for non-simple processes and this is exactly what is achieved in the next theorem.

THEOREM 6.12. *For $f \in \mathfrak{h}[t_0, t]$, there is a sequence $(h_n)_{n\in\mathbb{N}}$ of simple processes converging to $f$ in $\mathfrak{h}[t_0, t]$ and there exists $I(f) \in L^2(\mathcal{C}_t)$ such that $\int_{t_0}^{t} h_n(s) d\Psi_s$ converges to $I(f)$ in $L^2(\mathcal{C}_t)$. Furthermore, $I(f)$ is independent of the particular sequence $(h_n)_{n\in\mathbb{N}}$ converging to $f$.*

PROOF. A sequence of $(h_n)_{n\in\mathbb{N}}$ converging to $f$ in $\mathfrak{h}[t_0, t]$ is guaranteed to exist by virtue of Theorem 6.12. By the isometry property, since $(h_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathfrak{h}[t_0, t]$, $(\int_{t_0}^{t} h_n d\Psi)$ is Cauchy in $L^2(\mathcal{C}_t)$. The existence of the required $I(f)$ follows by completeness of $L^2(\mathcal{C}_t)$. The required independence is also easy to see. □

We may now formally define the Itô-Clifford Integral.

DEFINITION 6.10 (The Itô Clifford Integral). *For $f \in \mathfrak{h}[t_0, t]$, the Itô-Clifford Integral $\int_{t_0}^{t} f(s) d\Psi_s$ of $f$ is given by $I(f)$ where $I(f)$ is as given in Theorem 6.12 and satisfies $I(f) \in L^2(\mathcal{C}_t)$. If $0 \leq t_0 \leq t_1 \leq t_2$ and $f \in \mathfrak{h}[t_0, t_2]$ then the* *restrictions of $f$ to the intervals $[t_0, t_1]$ and $[t_1, t_2]$ belong to $\mathfrak{h}[t_0, t_1]$ and $\mathfrak{h}[t_1, t_2]$ respectively and*

$$\int_{t_0}^{t_2} f d\Psi_s = \int_{t_0}^{t_1} f d\Psi_s + \int_{t_1}^{t_2} f d\Psi_s.$$

We conclude this section by discussing some properties of the Itô-Clifford Integral.

Theorem 6.13 (Properties of the Itô Clifford Integral).

*i)* $\int_{t_0}^{t} (\alpha f + \beta g) d\Psi_s = \alpha \int_{t_0}^{t} f(s) d\Psi_s + \beta \int_{t_0}^{t} g(s) d\Psi_s$ *for* $f, g \in \mathfrak{h}[t_0, t]$ *and* $\alpha, \beta \in \mathbb{C}$.

*ii)* $m(\int_{t_0}^{t} f d\Psi_s) = 0$ *for any* $f \in \mathfrak{h}[t_0, t]$.

*iii)* $\left\| \int_{t_0}^{t} f d\Psi_s \right\|_2 = \|f\|_{\mathfrak{h}}$ *for any* $f \in \mathfrak{h}[t_0, t]$.

*iv)* $M_{t_0}(\int_{t_0}^{t} f d\Psi_s) = 0$ *for any* $f \in \mathfrak{h}[t_0, t]$.

*v) If* $f \in \mathfrak{h}[0, t]$ *for all* $t \in \mathbb{R}^+$, *then* $\{\int_0^t f d\Psi_s\}_{t \in \mathbb{R}^+}$ *is an* $L^2$*-martingale adapted to* $\{\mathcal{C}_t\}_{t \in \mathbb{R}^+}$.

Proof. *i*), *ii*) and *iii*) follow directly from Theorem 6.8.

To show *iv*) it is enough that we verify the property for simple processes. Then, let $h = \sum_{k=1}^{n} h_{k-1} \chi_{[t_{k-1}, t_k]}$ with $t_0 \le t_1 \le \cdots \le t_n = t$ where $h_{k-1} \in L^2(\mathcal{C}_{t_{k-1}})$ be a simple process in $\mathfrak{h}[t_0, t]$. By letting $g \in L^2(\mathcal{C}_{t_0})$, we obtain as a consequence of Lemma 6.1 that the following is true

$$\begin{aligned} m\Big( g \int_{t_0}^{t} h d\Psi_s \Big) &= \sum_{k=1}^{n} m(g h_{k-1} \Delta\Psi_k) \\ &= 0. \end{aligned}$$

Finally, for *iv*) we note the following

$$\begin{aligned} M_{t_0}\Big( \int_0^t f(s) d\Psi_s \Big) &= M_{t_0}\Big( \int_0^{t_0} f(s) d\Psi_s + \int_{t_0}^{t} f(s) d\Psi_s \Big) \\ &= M_{t_0}\Big( \int_0^{t_0} f(s) d\Psi_s \Big) \\ &= \int_0^{t_0} f(s) d\Psi_s. \end{aligned}$$

□

CHAPTER 7

# Conclusion

## 7.1. Summary of Results

Throughout this dissertation, we have seen that many elements of classical probability have analogues in the quantum framework. While classical probability is often grounded in set theory, quantum probability tends to rely more heavily on operator algebras.

This is seen on multiple occasions such as in the form a measure assumes in both frameworks. In particular, we have that classically a measure is realized as an integral over a set, and in the quantum framework as a trace of an operator. Another example of this is seen in Multiple Probability Spaces wherein the cross product was not sufficient in the quantum framework rather we had to utilize the tensor product. The construction of the tensor product was probabilistic in nature, relying on the use of Gaussian variables. Finally we provided the background of the quantum stochastic integral and also showed that it is constructed in a very similar to its classical analogue with an isometry theorem also being proven.

## 7.2. Further Research

In this dissertation we delved deeply into the realm of quantum probability and provided a lot of comparisons between classical probability and quantum probability. That being said, the dissertation is far from exhaustive, and there is a lot of room for future discussion and research.

With regards to the Boson Fock space, we note that there is much more to be discussed. Indeed we note that many classical distributions and stochastic processes may be recovered by measuring observables in so called coherent states.

The creation and annihilation operators may also be discussed in this context and these involve the use of Weyl operators. Further detail may be obtained in [19].

In addition, there are also several research directions when considering quantum stochastic integration. The construction we considered was centred around the fermion Fock space however an analogous integral based on the Boson Fock space may be constructed. This was indeed the first quantum stochastic integral to be considered and its construction may be found in [13].

Our discussion both in the classical and in the quantum case focused on the Itô Integral however this is not the only type of stochastic integral with the Stratonovich integral being another example. This type of integral is discussed briefly in [4] however the discussion is not exhaustive. The most common application of stochastic integrals are to solve stochastic differential equations and in [24] we see an extension of these to the quantum framework.

Finally, we propose a possible research direction which to our knowledge would be novel. In classical probability, we initially integrate with respect to Brownian Motion. One generalization of this is integration with respect to Fractional Brownian Motion which does not have martingale property and hence makes the construction of the integral much more mathematically involved. To our knowledge, both the construction of a Fractional Brownian Motion and a quantum stochastic integral with respect to it in a quantum probabilistic setting are unexplored and may present challenging and interesting avenues for future research.

## Bibliography


1. L. Accardi, *Quantum probability and the Foundations of Quantum Theory*, Statistics in Science: The Foundations of Statistical Methods in Biology, Physics and Economics, Springer, 1990, pp. 119–147.
2. E.R. Ackermann, J.P. De Villiers, and P. Cilliers, *Nonlinear Dynamic Systems Modelling using Gaussian Processes: Predicting Ionospheric Total Electron Content over South Africa*, Journal of Geophysical Research: Space Physics **116** (2011), no. A10.
3. A.K. Aringazin and M.I. Mazhitov, *Stochastic Models of Lagrangian Acceleration of Fluid Particle in Developed Turbulence*, International Journal of Modern Physics B **18** (2004), no. 23n24, 3095–3168.
4. C. Barnett, R.F. Streater, and I.F. Wilde, *The Itô-Clifford Integral*, Journal of Functional Analysis **48** (1982), no. 2.
5. J. Bell, *The Kolmogorov Continuity Theorem, Hölder Continuity, and Kolmogorov–Chentsov Theorem*, Lecture Notes, University of Toronto (2015).
6. P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Statistics, Wiley, 1995.
7. G. Birkhoff, *Lattice Theory*, vol. 25, American Mathematical Soc., 1940.
8. S.H. Chan, *Introduction to Probability for Data Science*, Michigan Publishing, 2021.
9. A.M. Gleason, *Measures on the Closed Subspaces of a Hilbert Space*, The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution, Springer, 1975, pp. 123–133.
10. H. Granström, *Gleason's Theorem*, Master's Thesis, Stockholm University, Stockholm, Sweden (2006).
11. L. Gross, *Existence and Uniqueness of Physical Ground States*, Journal of Functional Analysis **10** (1972), no. 1, 52–109.

12. R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 2012.
13. R.L. Hudson and K.R. Parthasarathy, *Quantum Itô's Formula and Stochastic Evolutions*, Communications in Mathematical Physics **93** (1984), no. 3, 301–323.
14. A.E. Kobryn, T. Hayashi, and T. Arimitsu, *Quantum Stochastic Differential Equations for Boson and Fermion Systems—method of non-Equilibrium Thermo Field Dynamics*, Annals of Physics **308** (2003), no. 2, 395–446.
15. S. Kumaresan, *Short Courses in Mathematics*, Universities Press, 2004.
16. R.A. Kunze, *$L_p$ Fourier Transforms on Locally Compact Unimodular Groups*, Transactions of the American Mathematical Society **89** (1958), no. 2, 519–540.
17. H.H. Kuo, *Introduction to Stochastic Integration*, Springer, 2006.
18. M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge university press, 2010.
19. K.R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, vol. 85, Birkhäuser, 2012.
20. B. Pinier, E. Mémin, S. Laizet, and R. Lewandowski, *Stochastic Flow Approach to model the Mean Velocity Profile of Wall-Bounded Flows*, Physical Review E **99** (2019), no. 6, 063101.
21. O. Raimond and G. Riabov, *Strong Measurable Continuous Modifications of Stochastic Flows.*, Ukrainian Mathematical Journal **75** (2024), no. 11.
22. W. Rudin, *Functional Analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill, 1991.
23. W. So, *Commutativity and Spectra of Hermitian Matrices*, Linear Algebra and its Applications **212** (1994), 121–129.
24. R.F. Streater, *The Itô-Clifford Integral part II*, Quantum Probability and Applications II: Proceedings of a Workshop held in Heidelberg, West Germany, October 1–5, 1984, Springer, 2006, pp. 493–503.
25. B.F. Vajargah and M. Moradi, *Period Dependent Branching Process and its Applications in Epidemiology*, Infection, Genetics and Evolution **11** (2011), no. 6, 1225–1228.

26. D.V. Voiculescu, *Free Probability Theory*, Fields Institute Monographs, American Mathematical Society, 1997.
27. J. Von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition*, Princeton University Press, 2018.
28. I.F. Wilde, *The Free Fermion Field as a Markov Field*, Journal of Functional Analysis **15** (1974), no. 1, 12–21.

APPENDIX A

# Conditional Expectation

In this appendix we shall be discussing conditional expectation and stating a few of its results without proof. These are taken from [17]. For this discussion we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $L^p(\Omega)$ the set of random variables $Y$ having $\mathbb{E}[|Y|] < \infty$. It may be shown that this is a Banach space with norm $\|Y\|_p = (\mathbb{E}[|Y|^p])^{\frac{1}{p}}$. We shall use $L^p(\Omega, \mathcal{F})$ when we want to emphasize the underlying $\sigma$-algebra. Consider the case when we have another $\sigma$-algebra $\mathcal{G}$ satisfying $\mathcal{G} \subset \mathcal{F}$. Now let $Y \in L^1(\Omega)$ and we define a real valued function $\mu$ on $\mathcal{G}$ by taking

$$\mu(A) = \int_A Y(\omega) d\mathbb{P}(\omega),\ A \in \mathcal{G}$$

Firstly, we emphasize the fact that $|\mu(A)| \leq \int_A |X| d\mathbb{P} \leq \int_\Omega |X| d\mathbb{P} = \mathbb{E}[X]$ for all $A \in \mathcal{G}$. Secondly we also have that $\mu$ satisfies the definition of a signed measure which is absolutely continuous with respect to $\mathbb{P}$. There properties allow us to apply the Radon–Nikodym as to obtain a random variable $Z$ which is $\mathcal{G}$ measurable and satisfies $\mathbb{E}[|Z|] < \infty$. Furthermore we have that

$$\mu(A) = \int_A Z(\omega) d\mathbb{P}(\omega)\ \forall A \in \mathcal{G}. \tag{80}$$

Now if $\tilde{Z}$ is another random variable which is $\mathcal{G}$ measurable, has finite expectation and satisfies (80) we obtain that $\int_A (Z - \tilde{Z}) d\mathbb{P} = 0$ for all $A \in \mathcal{G}$ meaning $Z = \tilde{Z}$ almost surely. From this discussion, we may proceed to the definition.

Definition A.1 (Conditional Expectation). *Let $Y \in L^1(\Omega, \mathcal{F})$. Let $\mathcal{G}$ be a $\sigma$-algebra and $\mathcal{G} \subset \mathcal{F}$. The conditional expectation of $Y$ given $\mathcal{G}$, $\mathbb{E}[Y|\mathcal{G}]$ is the unique random variable $Z$ up to $\mathbb{P}$ measure 1 which is $\mathcal{G}$ measurable and satisfies*

$$\int_A Y d\mathbb{P} = \int_A Z d\mathbb{P}\ \forall A \in \mathcal{G} \tag{81}$$

Theorem A.1 (Properties of the Conditional Expectation).

*i)* $\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]] = \mathbb{E}[Y]$.

*ii) For* $Y$ $\mathcal{G}$ *measurable,* $\mathbb{E}[Y|\mathcal{G}] = Y$.

*iii) For* $Y$ *and* $\mathcal{G}$ *independent,* $\mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y]$.

*iv) If* $Z$ *is* $\mathcal{G}$ *measurable and* $\mathbb{E}[|YZ|] < \infty$ *then* $\mathbb{E}[YZ|\mathcal{G}] = Y\mathbb{E}[Z|\mathcal{G}]$.

*v) If* $\mathcal{G}_1 \subset \mathcal{G}_2$ *then* $\mathbb{E}[Y|\mathcal{G}_1] = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}_2]|\mathcal{G}_1]$.