SOME CONSTRUCTIONS OF MUTUALLY UNBIASED BASES OVER FINITE FIELDS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÖKHAN ELMAS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
CRYPTOGRAPHY

FEBRUARY 2024

Approval of the thesis:

SOME CONSTRUCTIONS OF MUTUALLY UNBIASED BASES OVER FINITE FIELDS

submitted by **GÖKHAN ELMAS** in partial fulfillment of the requirements for the degree of **Master of Science in Cryptography Department, Middle East Technical University** by,

Prof. Dr. A. Sevtap Selçuk Kestel _____
Dean, Graduate School of **Applied Mathematics**

Assoc. Prof. Dr. Oğuz Yayla _____
Head of Department, **Cryptography**

Assist. Prof. Dr. Buket Özkaya _____
Supervisor, **Cryptography, METU**

Prof. Dr. Ferruh Özbudak _____
Co-supervisor, **Mathematics, FENS, Sabancı University**

Examining Committee Members:

Assoc. Prof. Dr. Oğuz Yayla
Cryptography, METU _____

Assist. Prof. Dr. Buket Özkaya
Cryptography, METU _____

Dr. Markus Grassl
ICTQT, University of Gdansk _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: GÖKHAN ELMAS

Signature :

ABSTRACT

SOME CONSTRUCTIONS OF MUTUALLY UNBIASED BASES OVER FINITE FIELDS

ELMAS, Gökhan
M.S., Department of Cryptography
Supervisor : Assist. Prof. Dr. Buket Özkaya
Co-Supervisor : Prof. Dr. Ferruh Özbudak

February 2024, 86 pages

Mutually unbiased bases, as a mathematical concept, has important implications in quantum information theory where the information is encoded as linear combinations of vectors in Hilbert spaces instead of as arrays of digits. Offering a designation on the preparation and measurement of the quantum states, mutually unbiased bases provide mathematics based security to many quantum computation protocols including the famous quantum key distribution protocol named BB84. The construction of mutually unbiased bases, however, is not straighforward and it requires an extensive mathematical approach based on the properties of the finite fields.

Keywords: mutually unbiased bases, quantum computation, Hilbert spaces, bent functions

ÖZ

KARŞILIKLI TARAFSIZ BAZLARIN SONLU CİSİMLER ÜZERİNDEKİ BAZI İNŞALARI

ELMAS, Gökhan
Yüksek Lisans, Kriptografi Bölümü
Tez Yöneticisi : Yrd. Doç. Dr. Buket Özkaya
Ortak Tez Yöneticisi : Prof. Dr. Ferruh Özbudak

Şubat 2024, 86 sayfa

Karşılıklı tarafsız bazlar, matematiksel bir kavram olarak, bilginin basamak dizileri yerine Hilbert uzaylarındaki vektörlerin lineer birleşimleri üzerine kodlandığı kuantum enformasyon teorisinde önemli çıkarımlara sahiptir. Kuantum hallerin hazırlanması ve ölçülmesi aşamalarına bir tasvir sağlamasına bağlı olarak, karşılıklı tarafsız bazlar, yaygın bilinen kuantum anahtar dağıtım protokolü BB84 de dahil olmak üzere bir çok kuantum hesaplama protokolüne matematik tabanlı güvenlik sağlamaktadır. Karşılıksız tarafsız bazların inşa edilmesi doğrudan gerçekleştirilememekte ve sonlu cisimlerin özellikleri üzerinde kapsamlı bir matematiksel yaklaşımı gerektirmektedir.

Anahtar Kelimeler: karşılıklı tarafsız bazlar, kuantum hesaplama, Hilbert uzayları, bent fonksiyonlar

Chapter 0!

ACKNOWLEDGMENTS

Foremost, I would like thank to my advisor, Dr. Buket Özkaya for her support and inspring insights during my masters degree studies. Her continuous support and suggestions made this thesis possible. I also would like to present my gratitude to my co-advisor Prof. Dr. Ferruh Özbudak who made a great contribution to this study by directing my transitions between quantum computing and mathematics with his immense knowledge. I am also grateful to the committee members, Assoc. Prof. Dr. Oğuz Yayla and Dr. Markus Grassl for their valuable comments and advices on my thesis studies.

Besides, I am very grateful to Assoc. Prof. Dr. Emre Yüce and his dear students, from METU Physics Department, for letting me dive into the astonishing world of optics and leading me to learn a lot with the experimental research in his laboratory. I am also thankful to Dr. Mahdi Hosseini for sharing his expertise with me on optical computational tasks.

Lastly, I am forever thankful to my family, for their continuous support and unconditional love.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

QFT	Quantum Fourier Transform
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
POVM	Positive Operator-Valued Measure
MUB	Mutually Unbiased Bases

CHAPTER 1

INTRODUCTION

One of the most frequent problems in mathematics, physics and engineering sciences is to provide equations with coordinate systems in which better analyses and solutions are possible.Therefore, as a coordinate transformation, Fourier Transformation is the one of the most crucial transformations with many modern applications in today's technology. For instance, a more computationally efficient version of Fourier Transform, called Fast Fourier Transform, is employed in global communication networks enabling transfer of knowledge all around the globe. Apart from the technological and applied side, Fourier transform also enables a comprehensable framework for quantum computation by directly producing bases vectors for Hilbert spaces in which qubits are described. Its quantum version, so-called Quantum Fourier Transform (QFT) is the backbone of many quantum algorithms which surpass existing classical algorithms in terms of computational complexity.

In this chapter, we introduce Hilbert functions and investigate the basic structures on Hilbert spaces by introducing the Fourier transforms and describe how it acts on the basis of Hilbert spaces.

1.1 Hilbert Spaces

We introduce the Hilbert spaces since we will investigate various forms of Fourier transforms via Hilbert spaces.

Definition 1.1.1. *Given a complex vector space* $\mathbf{H}$*, a map* $\langle\ ,\ \rangle : \mathbf{H} \times \mathbf{H} \longrightarrow \mathbb{C}$ *is called a Hermitian inner product on* $\mathbf{H}$ *if*

- $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$
- $\langle x, y\rangle = \overline{\langle y, x\rangle}$
- $\langle x, x\rangle \geq 0$
- $\langle x, x\rangle = 0 \iff x = 0$

for all $x, y, z \in \mathbf{H}$ *and for all* $a, b \in \mathbb{C}$.

As usual, one can define a norm on $\mathbf{H}$ via the formula

$$|x| = \sqrt{\langle x, x\rangle} \tag{1.1}$$

for all $x \in \mathbf{H}$.

Definition 1.1.2. *A complex vector space* $\mathbf{H}$ *with a Hermitian inner product is called Hilbert space if* $\mathbf{H}$ *is complete with respect to the norm induced by the inner product.*

Here, completeness means that every Cauchy sequence in the vector space $\mathbf{H}$ converges to a point in $\mathbf{H}$.

Example 1.1.3. *Let* $\mathbf{H}_1$ *be the most obvious complex vector space, namely* $\mathbb{C}^n$. *For any* $u = (u_1, ..., u_n), v = (v_1, ..., v_n) \in \mathbb{C}^n$, *set*

$$\langle u, v\rangle_1 = \sum_{i=1}^{n} \overline{u_i} v_i \tag{1.2}$$

is an inner product on $\mathbf{H}_1$ *with the induced norm*

$$\|u\|_1 = \sqrt{\langle u, u\rangle} = \sqrt{\sum_{i=1}^{n} \overline{u_i} u_i} = \sqrt{\sum_{i=1}^{n} |u_i|^2} \tag{1.3}$$

$\mathbf{H_1}$ *is a Hilbert space with the inner product* $\langle\ ,\ \rangle_1$ *and its induced norm.*

Here we note that if take $u, v \in \mathbf{H}_1$ as column matrices, like

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

the statement

$$\langle u, v\rangle_1 = \sum_{i=1}^{n} \overline{u_i} v_i$$

can be translated to the language of matrices as

$$\langle u, v\rangle_1 = u^\dagger v \tag{1.4}$$

where $u^\dagger$ indicates conjugate transpose matrix of u. More explicitly,

$$u^\dagger = \left[\overline{u_1}, \ldots, \overline{u_n}.\right] \tag{1.5}$$

Example 1.1.4. *Take $\mathbf{H}_2 = \mathrm{C}([a, b])$, the vector space of complex-valued continuous functions on the closed interval $[a, b]$. An inner product on $\mathbf{H}_2$ can be defined as*

$$\langle f, g\rangle_2 = \int_a^b \overline{f} g \tag{1.6}$$

with the induced norm

$$||f||_2 = \int_a^b \overline{f} f. \tag{1.7}$$

However, $\mathbf{H}_2$ is not a Hilbert space with this setting considering the fact that the norm (1.7) can diverge.

Although $\mathbf{H}_1$ is a Hilbert space and $\mathbf{H}_2$ is not, it is still possible to note a correspondence between $\langle\ ,\ \rangle_1$ and $\langle\ ,\ \rangle_2$. For this correspondence, we need to introduce vector representation of functions.

Given two functions $f, g \in \mathrm{C}([a, b])$, assume we generate an equi-spaced partition of the interval $[a, b]$ as

$$\tau = \{x_0 = a, x_1 = a + \Delta, \ldots, x_k = a + k\Delta, \ldots, x_n = b\}$$

where

$$\Delta = \frac{b - a}{n}.$$

For all $1 \leq i \leq i$, let $f_i = f(x_i)$. Then the column vector

$$f_\tau = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

is the vector representation of f with respect to the partition τ. With this setting, we can consider the inner product $\langle f_\tau, g_\tau \rangle_1$ as

$$\langle f_\tau, g_\tau \rangle_1 = f_\tau^\dagger g_\tau = \sum_{i=0}^{n} \overline{f_i} g_i. \tag{1.8}$$

Set

$$S_\tau = \langle f_\tau, g_\tau \rangle_1 \Delta.$$

As $n \to \infty$, we get the finest partition of $[a, b]$ and we have $\Delta \to 0$. Moreover, S_τ would be the Riemannian sum which can be restated as

$$S_\tau = \lim_{\Delta \to 0} \langle f_\tau, g_\tau \rangle_1 \Delta = \int_a^b \overline{f} g = \langle f, g \rangle_2.$$

A useful utilization of inner products in vector spaces is the determination of the orthonormal base vectors for the vector space. Note that two vectors u, v in a vector space V are called orthonormal if they are of unit length ($||u|| = ||v|| = 1$) and their inner product vanishes. ($\langle u, v \rangle = 0$).

Let $\mathbf{V}$ be a complex vector space of dimension n with an inner product $\langle\ ,\ \rangle$ and assume $\mathbf{V}$ has an orthonormal basis $\{e_1, e_2, ..., e_n\}$. Then we have $\langle e_i, e_j \rangle = 0$ for $i \neq j$ and $\langle e_i, e_i \rangle = 1$ for all $1 \leq i \leq n$. Any vector $v \in \mathbf{V}$ can be uniquely written as a linear combination of basis vectors with

$$v = v_1 e_1 + ... + v_n e_n = \sum_{i=1}^{n} v_i e_i \tag{1.9}$$

where $v_i \in \mathbb{C}$. Since the expression in (1.9) is unique with respect to the choice of basis vectors, each of the values v_i are well-defined.

Moreover, for each $i \in \{1, ..., n\}$,

$$\langle e_i, v\rangle = \langle e_i, \sum_{j=1}^{n} v_j e_j\rangle = \sum_{j=1}^{n} v_j \langle e_i, e_j\rangle = \sum_{j=1, j\neq i}^{n} v_j \langle e_i, e_j\rangle + v_i \langle e_i, e_i\rangle = v_i. \quad (1.10)$$

Therefore, given a linear combination of orthonormal basis vectors $v = \sum_{i=0}^{n} v_i e_i$, the coefficients v_i can be determined using the inner product $\langle\ ,\ \rangle$ as

$$v_i = \langle e_i, v\rangle. \quad (1.11)$$

Inner products are still useful when we have an orthogonal basis but not an orthonormal one. Let $\{f_1, ..., f_n\}$ be an orthogonal but not orthonormal basis for a complex vector space $\mathbf{V}$. Therefore $\langle f_i, f_j\rangle = 0$ when $i \neq j$ but $\langle f_i, f_i\rangle \neq 1$ for at least one i. Since $\{f_1, ..., f_n\}$ is still a basis, any vector $v \in \mathbf{V}$ can be uniquely stated as

$$v = v_1 f_1 + ... + v_n f_n = \sum_{i=1}^{n} v_i f_i. \quad (1.12)$$

Then for each i, we calculate

$$\langle f_i, v\rangle = \langle f_i, \sum_{j=1}^{n} v_j f_j\rangle = \sum_{j=1}^{n} v_j \langle f_i, f_j\rangle = v_i \langle f_i, f_i\rangle. \quad (1.13)$$

From (1.13), one can conclude

$$v_i = \frac{\langle f_i, v\rangle}{\langle f_i, f_i\rangle}. \quad (1.14)$$

Plugging (1.14) into (1.12), for any $v \in \mathbf{V}$, we have,

$$v = \sum_{i=1}^{n} v_i f_i = \sum_{i=1}^{n} \frac{\langle f_i, v\rangle}{\langle f_i, f_i\rangle} f_i. \quad (1.15)$$

1.2 Function Bases

In this section, we move our concentration from generic vector spaces to specific vector spaces that will be used through out the study. Since Fourier transform takes functions to functions, we work on vector spaces where the points are not n-tuples but continuous functions. Since each bases element is already an element of the vector space, we will construct a base consisting of functions.

As we introduced in Example 1.1.4, the set of complex-valued continuous functions on the closed real interval $[a, b]$, denoted by $\mathbf{C}([a, b])$, is a vector space with the inner product

$$\langle f, g\rangle = \int_a^b \overline{f} g.$$

Although $\mathbf{C}[a, b]$ is not a Hilbert space, it is a separable topological space and by separability we mean the condition that the space contains a countable dense subset. By Weierstrass approximation theorem, the polynomials in $\mathbf{C}[a, b]$ with rational coefficients forms a countable dense subset in $\mathbf{C}[a, b]$ and therefore, $\mathbf{C}[a, b]$ is separable.

Taking $a = -\pi$ and $b = \pi$, for now we fix $\mathbf{V} = \mathbf{C}([-\pi, \pi])$. Let $E = \{\varphi_n\}_{n\in\mathbb{Z}} \subset \mathbf{V}$ be an infinite family of functions where

$$\begin{aligned} \varphi_k \colon [-\pi,\pi] &\to \mathbb{C} \\ x &\to e^{ikx} = \cos(kx) + i\sin(kx). \end{aligned}$$

We now observe that the set E is actually a basis for $\mathbf{V}$. In order to see this, we first assume $k \neq j$ and calculate

$$\begin{aligned} \langle \varphi_k, \varphi_j\rangle &= \int_{-\pi}^{\pi} \overline{\varphi_k}(x)\varphi_j(x)dx = \int_{-\pi}^{\pi} e^{-ikx}e^{ijx}dx = \int_{-\pi}^{\pi} e^{i(j-k)x}dx \\ &= \int_{-\pi}^{\pi} \cos((j-k)x) + i\sin((j-k)x)dx \\ &= \int_{-\pi}^{\pi} \cos((j-k)x)dx + i\int_{-\pi}^{\pi} \sin((j-k)x)dx. \end{aligned} \tag{1.16}$$

For the first part of the integral,

$$\begin{aligned} \int_{-\pi}^{\pi} \cos((j-k)x)dx &= \frac{\sin((j-k)x)}{i(j-k)}|_{-\pi}^{\pi} \\ &= \frac{1}{i(j-k)}(\sin((j-k)\pi) + \sin((j-k)\pi)) \\ &= 0 + 0 = 0. \end{aligned} \tag{1.17}$$

Similarly, for the second part of the integral (we omit i for the ease of calculation),

$$\begin{aligned}\int_{-\pi}^{\pi} \sin((j-k)x)dx &= \frac{-\cos((j-k)x)}{i(j-k)}|_{-\pi}^{\pi} \\ &= \frac{1}{i(k-j)}(\cos((j-k)\pi) - \cos((j-k)(-\pi))) \\ &= \frac{1}{i(k-j)}(\cos((j-k)\pi) - \cos((j-k)\pi)) \\ &= 0.\end{aligned} \tag{1.18}$$

Combining (1.17) and (1.18) we conclude that when $j \neq k$

$$\langle \varphi_k, \varphi_j \rangle = 0. \tag{1.19}$$

For the remaining case, when $k = j$,

$$\begin{aligned}\langle \varphi_k, \varphi_k \rangle &= \int_{-\pi}^{\pi} \overline{\varphi_k}(x)\varphi_k(x)dx \\ &= \int_{-\pi}^{\pi} e^{-ikx}e^{ikx}dx \\ &= \int_{-\pi}^{\pi} 1dx \\ &= 2\pi.\end{aligned} \tag{1.20}$$

From the equations (1.19) and (1.20), we obtain $E = \{\varphi_n\}_{n\in\mathbb{Z}} \subset \mathbf{V}$ as a set of orthogonal functions with respect to the inner product $\langle , \rangle$ of $\mathbf{V}$. Actually, the set E is a basis for $\mathbf{C}[a, b]$ and for any $f \in C[a, b]$, we can write

$$f(x) = \sum_{k=-\infty}^{\infty} c_k\varphi_k = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \tag{1.21}$$

where $c_k \in \mathbb{C}$.

We note a remark that stating a countable basis for $\mathbf{C}[-\pi, \pi]$ does not violate the fact a Hilbert space is separable if and only if it has a countable basis, since we already explained that $\mathbf{C}[-\pi, \pi]$ is not a Hilbert space.

1.3 Fourier Series

In this section we investigate Equation (1.21) and determine the coefficients $c_k \in \mathbb{C}$. However, before that we introduce an equivalent representation of functions in $\mathbf{V} = \mathbf{C}([a, b])$. Equation (1.21) is stated as a infinite sum in both ends (from $-\infty$ to ∞). Actually, using some trigonometric identities, we can convert it into an infinite sum in one end (from 0 to ∞). Our main reference in this section is [7]

First of all, we write $c_k = a_k + ib_k$ with $a, b \in \mathbb{R}$ for all k. When $k = 0$, the term $c_k e^{ikx}$ is only $c_0 = a_0 + ib_0$ since $e^0 = 1$. Therefore the $k = 0$ term can be separated from the infinite sum. Each summand in the sum can be stated as

$$\begin{aligned} c_k e^{ikx} &= (a_k + ib_k)(\cos kx + i \sin kx) \\ &= a_k \cos kx + ia_k \sin kx + ib_k \cos kx + i^2 b_k \sin kx \\ &= a_k \cos kx + ia_k \sin kx + ib_k \cos kx - b_k \sin kx. \end{aligned} \tag{1.22}$$

Since $\cos(x)$ is even and $\sin(x)$ is odd $(\cos(-x) = \cos(x), \sin(-x) = -\sin(x))$, for each positive integer k, the sum $c_k e^{ikx} + c_{-k} e^{-ikx}$ can be formulated as (using Equation (1.22))

$$\begin{aligned} c_k e^{ikx} + c_{-k} e^{-ikx} &= a_k \cos kx + ia_k \sin kx + ib_k \sin kx - b_k \sin kx \\ &\quad + a_{-k} \cos(-kx) + ia_{-k} \sin(-kx) + ib_{-k} \cos(-kx) + b_{-k} \sin(-kx) \\ &= \big((a_k + a_{-k}) \cos kx + (-b_k + b_{-k}) \sin kx\big) \\ &\quad + i\big((b_k + b_{-k}) \cos kx + (a_k - a_{-k}) \sin kx\big). \end{aligned} \tag{1.23}$$

Therefore, Equation (1.21) can be restated as

$$\begin{aligned} f(x) &= \sum_{k=-\infty}^{\infty} c_k e^{ikx} \\ &= (a_0 + ib_0) + \sum_{k=1}^{\infty} (a_k + a_{-k}) \cos kx + (-b_k + b_{-k}) \sin kx \\ &\quad + i \sum_{k=1}^{\infty} (b_k + b_{-k}) \cos kx + (a_k - a_{-k}) \sin kx. \end{aligned} \tag{1.24}$$

From Equation 1.24, we observe that f is a real-valued function on $[-\pi, \pi]$ if and

only if $b_k = -b_{-k}$ and $a_k = a_{-k}$ for all k.

We now move back to Equation (1.21) and determine the coefficients c_k by employing the equations we have derived in Section 1.1 and Section 1.2. From Equation (1.15) we get,

$$c_k = \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle} = \frac{\int_{-\pi}^{\pi} f(x) e^{-ikx} dx}{2\pi}. \tag{1.25}$$

Once more Equation (1.21) can be restated as

$$\begin{aligned} f(x) &= \sum_{k=-\infty}^{\infty} c_k e^{ikx} \\ &= \sum_{k=-\infty}^{\infty} \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle} e^{ikx} \\ &= \sum_{k=-\infty}^{\infty} \frac{\int_{-\pi}^{\pi} f(x) e^{-ikx} dx}{2\pi} e^{ikx}. \end{aligned} \tag{1.26}$$

Definition 1.3.1. *For any $f \in C[-\pi, \pi]$, the statement in Equation (1.26) is called the Fourier series representation of f.*

1.4 Fourier Transform

By carefully employing parameter change transformations, the Fourier series presentations for vector spaces rather than $\mathbf{C}([-\pi, \pi])$ can actually be obtained. For any $L \in \mathbb{R}$, we start with the vector space $\mathbf{V} = \mathbf{C}([-L, L))$. By applying a linear change of parameters, this time we take the base set $E = \{\varphi_k\}_{k \in \mathbf{Z}}$ as a set of functions where

$$\begin{aligned} \varphi_k : &[-L, L] \longrightarrow \mathbb{C} \\ &x \to e^{\frac{ik\pi x}{L}} = \cos(\frac{k\pi x}{L}) + i \sin(\frac{k\pi x}{L}). \end{aligned} \tag{1.27}$$

Then, when $j \neq k$

$$\langle \varphi_j, \varphi_k \rangle = \int_{-L}^{L} e^{\frac{i(k-j)\pi x}{L}} dx = 0 \tag{1.28}$$

and when $j = k$,

$$\langle \varphi_k, \varphi_k \rangle = \int_{-L}^{L} 1 dx = 2L. \tag{1.29}$$

Under this setting, the Fourier series of any $f \in C[-L, L]$ can be stated as

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{ik\pi x}{L}}. \tag{1.30}$$

As we did before, the coefficients c_k can be determined using Equation (1.14):

$$c_k = \frac{\langle \varphi_k, f \rangle}{\langle \varphi_k, \varphi_k \rangle} = \frac{\int_{-L}^{L} f(x) e^{-\frac{ik\pi x}{L}} dx}{2L}. \tag{1.31}$$

By substiting Equation (1.31) into the Equation (1.30), the Fourier series representation of a function f in $\mathbf{C}([-L, L])$ can be formulated as

$$f(x) = \sum_{k=-\infty}^{\infty} \frac{1}{2L} e^{\frac{ik\pi x}{L}} \int_{-L}^{L} f(x) e^{-\frac{ik\pi x}{L}} dx. \tag{1.32}$$

We note that the Fourier series representations we have introduced so far works for functions which are defined on closed intervals with boundaries symmetric to 0. A direct way of extending the descriptions to the whole real line $\mathbb{R}$ is to let $L \longrightarrow \infty$. When this is the case, we would also have $\Delta\alpha = \frac{\pi}{L} \longrightarrow 0$. We set $\Delta\alpha_k = k\Delta\alpha = \frac{k\pi}{L}$.

We already have

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{ik\pi x}{L}} = \sum_{k=-\infty}^{\infty} \frac{1}{2L} \left(\int_{-L}^{L} f(x) e^{\frac{-ik\pi x}{L}} dx \right) e^{\frac{ik\pi x}{L}}. \tag{1.33}$$

Since $\Delta\alpha = \dfrac{\pi}{L}$, we have $\dfrac{1}{2L} = \dfrac{\Delta\alpha}{2\pi}$ and $\dfrac{k\pi x}{L} = k\Delta\alpha x$. Letting $L \longrightarrow \infty$

$$\begin{aligned} f(x) &= \lim_{\Delta\alpha \longrightarrow 0} \sum_{k=-\infty}^{\infty} \frac{\Delta\alpha}{2\pi} \left(\int_{\frac{-\pi}{\Delta\alpha}}^{\frac{\pi}{\Delta\alpha}} f(y) e^{-ik\Delta\alpha y} dy \right) e^{ik\Delta\alpha x} \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi} \left(\int_{-\infty}^{\infty} f(y) e^{-i\alpha y} dy \right) e^{i\alpha x} d\alpha. \end{aligned} \tag{1.34}$$

In (1.34), the part of the equation written is parenthesis is called the Fourier transform of the function f and it is denoted by $\mathcal{F}(f(x))$. More explicitly,

$$\mathcal{F}(f(x)) = \int_{-\infty}^{\infty} f(x)e^{-i\alpha x}dx = \hat{f}(y)$$

with an inverse

$$\mathcal{F}^{-1}(\hat{f}(y)) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(y)e^{i\alpha y}dy = f(x).$$

We conclude this section by showing that the Fourier Transform is a linear operator since,

$$\begin{aligned}\mathcal{F}((f+g)(x)) &= \int_{-\infty}^{\infty}(f(x)+g(x))e^{-i\alpha x}dx \\ &= \int_{-\infty}^{\infty} f(x)e^{-i\alpha x}dx + \int_{-\infty}^{\infty} g(x)e^{-i\alpha x}dx \\ &= \mathcal{F}(f(x)) + \mathcal{F}(g(x))\end{aligned} \tag{1.35}$$

and

$$\begin{aligned}\mathcal{F}(cf(x)) &= \int_{-\infty}^{\infty} cf(x)e^{-i\alpha x}dx \\ &= c\int_{-\infty}^{\infty} f(x)e^{-i\alpha x}dx \\ &= c\mathcal{F}(f(x)).\end{aligned} \tag{1.36}$$

1.5 Fourier Transforms in Real World: Discrete Fourier Transform (DFT) and Fast Fourier Transform(FFT)

As we noted with Equations (1.35) and (1.36), the Fourier Transform is linear but we do not have matrix representation for it since it acts on the vector space of continuous functions but not discrete valued vector spaces. Moreover, in application, data are represented as vectors and the machinery built for processing the data employs mathematical techniques which are in accordance with the representation of the data. In digital signal processing for example, an analog signal which is continuous in the time domain is digitalized by listing its values in discrete times as a vector. As signal turns into a finite length vector, Fourier analysis actually provides a change of basis which makes it easier to analyse and manipulate the data.

In order to introduce the discrete version of Fourier Transform, called Discrete Fourier Transform, we take $\mathbf{V} = \mathbb{C}^n$. Given a vector $F = (f_1, ..., f_n) \in \mathbf{V}$, the output of the discrete Fourier transform is another vector $\hat{F} = (\hat{f}_1, ..., \hat{f}_n)$ such that

$$\hat{f}_k = \sum_{j=0}^{n-1} f_j \xi_n^{-i2\pi jk} \tag{1.37}$$

with inverse

$$f_k = \frac{1}{n} \sum_{j=0}^{n-1} \hat{f}_j \xi_n^{i2\pi jk} \tag{1.38}$$

where $\xi_n = e^{\frac{2\pi i}{n}}$ is the n-th root of unity. As a linear map between vectors, the discrete Fourier transform (DFT) has a matrix formulation as

$$\mathcal{F} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \xi_n & \xi_n^2 & \dots & \xi_n^{n-1} \\ 1 & \xi_n^2 & \xi_n^4 & \dots & \xi_n^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_n^{n-1} & \xi_n^{2(n-1)} & \dots & \xi_n^{(n-1)(n-1)}. \end{bmatrix} \tag{1.39}$$

The matrix representation of DFT given in Equation (1.39) has n^2 entries which implies a computational complexity of $\mathcal{O}(n^2)$ at each DFT calculation needed. However, there is an efficient method introduced in [9]which can replace the complexity with $\mathcal{O}(nlogn)$. The method is called Fast-Fourier Transform (FFT) and it actually is a divide-and-conquer algorithm applied on the big matrix given in Equation (1.39).

In order to describe the Fast Fourier Transform as an divide-and-conquer algorithm, we first let $N = 2^m$ and take a polynomial $p(x)$ as

$$\begin{aligned} p(x) &= \sum_{i=0}^{N-1} a_i x_i \\ &= a_0 + a_1 x + \cdots + a_i x^i + \cdots + a_{N-1} x^{N-1}. \end{aligned} \tag{1.40}$$

$P(x)$ can be written as a sum of an even polynomial $p_e(x) = a_0 + a_2 x^2 + \dots a_{N-2} x^{N-2}$ and an odd polynomial $p_o(x) = a_1 x + a_3 x^3 + \cdots + a_{N-1} x^{N-1}$. Therefore, we restate

Equation (1.38) as

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_i x^i + \ldots a_{N-1} x \\ &= (a_0 + a_2 x^2 + \cdots + a_{N-2} x^{N-2}) + (a_1 x + a_3 x^3 + \cdots + a_{N-1} x^{N-1}) \\ &= (a_0 + a_2 x^2 + \cdots + a_{N-2} x^{N-2}) + x(a_1 + a_3 x^2 + \cdots + a_{N-1} x^{N-2}). \end{aligned} \tag{1.41}$$

Setting $p_1(t) = a_0 + a_2 t + \cdots + a_{N-2} t^{\frac{N}{2}-1}$ and $a_1 + a_3 t + \cdots + a_{N-1} t^{\frac{N}{2}-1}$, we compose $p(x)$ as

$$p(x) = p_1(x^2) + x p_2(x^2). \tag{1.42}$$

Taking $\alpha = e^{-\frac{2\pi i}{N}}$, which is again an n-th root of unity, we recall Equation (1.37)

$$\begin{aligned} \hat{f}_k &= \sum_{j=0}^{n-1} f_j \xi_n^{-i2\pi jk} \\ &= \sum_{j=0}^{n-1} f_j (\xi_n^{-i2\pi k})^j. \end{aligned} \tag{1.43}$$

By using the technique as in Equation (1.37), for each $0 \leq j \leq \frac{N}{2} - 1$, we write

$$\begin{aligned} \hat{f}_k &= p_1(\alpha^{2k}) + \alpha^k p_2(\alpha^{2k}) \\ \hat{f}_{\frac{N}{2}+k} &= p_1(\alpha^{N+2k}) + \alpha^{\frac{N}{2}+k} p_2(\alpha^{N+2k}). \end{aligned} \tag{1.44}$$

Since α^2 is an $\frac{N}{2}$-th root of unity, Equation (1.44) actually consists of seperation of a Discrete Fourier Transform of n-tuple vectors into two Discrete Fourier Transform of two $\frac{n}{2}$ tuples. By iteration of the same steps, the length of DFT will decrease and the transformation can be completed with a cost of computational complexity of $\mathcal{O}(n\, logn)$. This method is called Fast Fourier Transform (FFT) and it is ubiquitously employed in modern communication and information processing systems.

The quantum version of Fourier Transform also exists and it is one of the most efficient tools of quantum information theory. In order to set the basics for Quantum Fourier Transform(QFT), we will introduce fundamentals of the quantum theory in the next chapter. But for now, at least we can list the most well-known quantum algorithms in which Quantum Fourier Transform is performed:

- **Grover's Algorithm:** Grover's algorithm is a search algorithm developed by Lov Grover in 1996. It is actually a probalistic algorithm but still can output the correct value with very high probability when the correct implementation of the iterations are provided.

  Given a binary function

  $$f : \Sigma^n \to \Sigma$$

  where $\Sigma = \{0, 1\}$ with at most one $x \in \Sigma^n$ such that $f(x) = 1$, Grover's algorithm outputs the correct x value if it exists. If such an x does not exist, Grover's algorithm returns no solution. Classically, one can try all possible inputs in Σ^n and find x with a complexity of $\mathcal{O}(2^n)$. However, the complexity of Grover's algorithm is $\mathcal{O}(\sqrt{n})$. As an application of being able to solving binary equations, query searchs on unstructed databases can be performed by using Grover's algorithm. [11]

- **Shor's Algorithm:** Shor's algorithm is the most crucial quantum algorithm with a promising capability of breaking the most of current cryptographic protocols. Shor's algorithm was introduced by Peter Shor in 1994 and it can be generalized to a solution of "Hidden Subgroup Problem for Abelian Groups". On the other hand, integer factorization problem can be reduced to finding the order of element modulo $N = pq$ which can be solved by Shor's algorithm. It has a computational complexity of $\mathcal{O}((\log(N))^2 \log(\log(N))$ it is much more efficient than its classical counterparts. [25]

CHAPTER 2

QUANTUM THEORY

In this chapter, we will introduce quantum mechanical facts and statements needed to comprehend quantum information tasks. Due to the advantages based upon quantum theory, quantum information processing promises classically unreachable capabilities in two distinct but closely related fields, secure communication and computing.

Comparing to the quantum computing's current achievement, quantum communication is already on the field with products available on the market. Using the quantum mechanical facts like no-cloning and measurement collapse, quantum communication protocols offer more secure channels where the information is encoded in the polarization of the photons. For the computing side, there are various quantum mechanical models employed to encode and process qubits. The quantum error correction methods are of great importance both for communication and computing, since the implementation of qubits on physical systems require very sensitive particles and they are always prone to errors due to the effects of the environment.

2.1 Introduction

Quantum mechanics as a fundamental but indeterministic theory which provides the framework for quantum information theory was not developed in order to replace classical mechanical explanations of nature for no reason. The existence of some physical phomenia which could not be explained with the principals of classical me-

chanics then could be understood thanks to peculiar aspects of quantum mechanics. For example, blackbody radiation is one of the quantum mechanical terms which explains the correspondence between an object's temperature and the wavelength of the radiation emitted by the object. Before the quantum mechanical approach, classically it was explained by Rayleigh-Jeans model[26]. Based on the estimations of the Rayleigh-Jeans model, at short wavelengths, the amount of the radiation being emitted due to black-body radiation would keep increasing with to finite limit amount of radiation. However, the experimental data did not agree with the model, by showing a sharp drop around shorter wavelengths. Actually, in the Rayleigh-Jeans model the emitted radiation was formulated as

$$B(\lambda, T) = \frac{2ck_bT}{\lambda^4} \tag{2.1}$$

where λ is the wavelength, T is the temperature, k_B is the Boltzmann constant and c is the speed of light. What really matters in Equation (2.1) is the denominator λ^4, causing the amount of emitted radiation to diverge in short wavelengths and it has been disclassified by a quantum mechanical approach.

In 1900, Max Planck made a key assumption that light could be emitted in discrete chunks with a constant energy proportional to the frequency [21] . Introducing a constant $h \approx 6.62 \times 10^{-34} Js$, Planck formulated the energy of light as

$$E = h\nu \tag{2.2}$$

where ν denotes the frequency of the light in terms of Hertz. Based on Planck's quantum mechanical principles, the behavior of black-body radiation is formulated with Planck's Law as,

$$B(\lambda, T) = \frac{2hc}{\lambda^5} \frac{1}{e^{\frac{hc}{k_BT}} - 1} \tag{2.3}$$

which is coherent with experimental data.

Another observation that could not be totally understood with the framework of classical mechanics was the photoelectric effect. In 1887, Heinrich Hertz realized that ap-

plying ultraviolet light on a metal object causes an electric flow of negative charge[15]. In 1902, Philipp Lenard observed that the voltage needed to stop the flow generated by shining light on metal depends not on the amplitude of light but on its frequency[19]. Therefore, shining brighter light did not cause electrons to have more kinetic energy, and the electrons were being emitted owing the same kinetic energy. Moreover, the kinetic energy of the electrons were determined by the wavelength of the light. On this peculiar observation, Einstein's proposal and Planck's earlier proposal were based on exactly the same idea, stating that electrons have energy of $E = h\nu$. But Einstein went further and he also claimed that light itself was made of particles and called them photons. Einstein also declared that for the emission, the energy of the photon must exceed a work function [10]. Einstein was awarded the 1922 Nobel Prize in Physics for his ideas on the photoelectric effect.

Since light was proven to be a wave by electromagnetic theory, it was an inevitable choice to consider light in wave-particle duality. In 1924, Louis de Broglie stated a hypothesis about wave-like behaviors of particles with mass[6]. Based on his proposal, the wavelength of a particle with mass could be determined by the formula

$$\lambda = \frac{h}{\rho} \tag{2.4}$$

where ρ is the momentum of the particle and classically formulated as

$$\rho = mv = \sqrt{2mE_K} \tag{2.5}$$

where m, v and E_K denote mass, velocity and kinetic energy respectively.

2.2 Schrödinger's Wave Equation

In 1925, Werner Heisenberg constructed a formulation of quantum mechanics based on matrices and linear operators[13]. Also, in 1926 Schrödinger proposed a differential equation which does not rely on matrices but still correspond to linearity and eigenstates [23]. In order to introduce Schrödinger's wave equation, we start with a

classical Helmholtz wave equation of the form

$$\frac{d^2\psi}{dz^2} = -k^2\psi \tag{2.6}$$

where

$$k = \frac{2\pi}{\lambda}. \tag{2.7}$$

Classical Helmholtz equations can be obtained from the behaviors of classical oscillations and it is used to describe classical monochromatic waves. Some of its solutions are

$$\psi_1(z) = e^{ikz} \tag{2.8}$$

$$\psi_2(z) = e^{-ikz} \tag{2.9}$$

$$\psi_3(z) = \cos(kz) \tag{2.10}$$

$$\psi_4(z) = \sin(kz) \tag{2.11}$$

Using de Broglie's hypothesis ,we restate Equation (2.7) as

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{\frac{\lambda}{\rho}} = \frac{2\pi\rho}{h} = \frac{\rho}{\hbar} \tag{2.12}$$

where $\hbar = \dfrac{h}{2\pi}$.

Substituting Equation (2.12) into the Equation (2.6), we have

$$\nabla^2\psi = -\frac{\rho^2}{\hbar^2}\psi \tag{2.13}$$

which can be written as

$$-\hbar^2\nabla\psi = \rho^2\psi. \tag{2.14}$$

Dividing both sides by $2m$, m being mass, we pass to

$$\begin{aligned} \frac{-\hbar^2}{2m}\nabla^2\psi &= \frac{\rho^2}{2m}\psi \\ &= (E_K)\psi \end{aligned} \tag{2.15}$$

since the kinetic energy, E_K, equals $\dfrac{\rho^2}{2m}$. For any mechanical system, the total energy E can be written as a sum of kinetic energy, E_K and the potential energy $V(r)$.

Since the potential energy depends on the location vector, we represent it by $V(r)$. Therefore,

$$E = E_K + V(r) \tag{2.16}$$

which implies

$$E_K = E - V(r). \tag{2.17}$$

Plugging Equation(2.17) into (2.15), we get

$$\begin{aligned} \frac{-\hbar^2}{2m}\nabla^2\psi &= (E_K)\psi \\ &= (E - V(r))\psi. \end{aligned} \tag{2.18}$$

Or, equivalently,

$$\left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)\psi = E\psi. \tag{2.19}$$

Equation (2.19) is called time-independent Schrödinger's equation and its solutions correspond to quantum states. Although we stated time-independent Schrödinger's equation, actually there is no way to derive it without starting with making first principal assumptions. Therefore, it can only be postulated and our calculation began with assuming the classical wave equation somehow could be transferred into quantum theory.

Although we have already stated that quantum theory is probabilistic, our descriptions so far did not emphasize its probabilistic nature. The probabilistic interpretation of quantum mechanics dates back to 1926, when Max Born proposed the idea that the modulus square of a wave function at some point equals the probability that the particle would be observed at that point after the measurement [4]. Therefore, the values $||\psi(r)||^2$ can be considered as probability densities relating to the quantum mechanical amplitudes. Here, we intentionally use the word "relating" since it is not directly equal to quantum mechanical amplitudes.

Probabilistic spaces need normalization, as the total proability equals 1. Therefore, in order to directly get the quantum mechanical amplitudes as probabilistic values, we need to normalize vawe functions. For this purpose, we first note that Schrödinger's Equation is linear since

$$\left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)(\psi_1 + \psi_2) = E(\psi_1 + \psi_2) \tag{2.20}$$

implies

$$\left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)\psi_1 + \left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)\psi_2 = E\psi_1 + E\psi_2 \tag{2.21}$$

and

$$\left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)(c\psi) = E(c\psi) \tag{2.22}$$

implies

$$c\left(\frac{-\hbar^2}{2m}\nabla^2 + V(r)\right)(\psi) = c(E(\psi)) \tag{2.23}$$

Now, let $P(r)$ be the probability that a particle is found at location r after the measurement. Therefore, around a small neighborhood d^3r of r in $\mathbb{R}^3$, we must have

$$\int P(r)d^3r = 1 \tag{2.24}$$

since all the possibilities must sum up to 1. If a solution ψ were to be a probability density directly, then our expectation would be

$$\int ||\psi(r)||^2 d^3r = 1. \tag{2.25}$$

However, this is not true in general, and a solution to Schrödinger's equation does not have to be a probability density. Therefore, there exists an additional constraint that the solutions have to be normalized by a factor.

Now assume that ψ is normalized wavefunction. ψ can be written as a linear combination of orthonormal functions $\{\varphi_k\}$ as

$$\psi(r,t) = \sum_{k=0}^{\infty} c_k(t)\varphi_k(r) \tag{2.26}$$

where the coefficients c_k are the coefficients depending on time .

Since $\psi(r,t)$ is normalized, we must have $\int_{-\infty}^{\infty}|\psi(r,t)|^2 d^3r = 1$. Therefore,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} |\psi(r,t)|^2 d^3r \\ &= \int_{-\infty}^{\infty} \overline{\psi}(r,t)\psi(r,t) d^3r \\ &= \int_{-\infty}^{\infty} \left(\sum_n \overline{c_n}(t)\overline{\varphi_n}(r)\right)\left(\sum_m c_m(t)\varphi_m(r)\right) d^3r. \end{aligned} \tag{2.27}$$

By the orthonogality of the spatial functions $\varphi_k(r)$, all the terms with $m \neq n$ vanish in Equation (2.27). By using orthonormality as well, we restate Equation (2.27) as

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \left(\sum_n \overline{c_n}(t)\overline{\varphi_n}(r)\right)\left(\sum_m c_m(t)\varphi_m(r)\right) d^3r \\ &= \int_{-\infty}^{\infty} \sum_n c_n(t)\overline{c_n}(t)\overline{\varphi_n}(r)\varphi_n(r) d^3r \\ &= \sum_n c_n(t)\overline{c_n}(t) \int_{-\infty}^{\infty} \overline{\varphi_n}(r)\varphi_n(r) d^3r \\ &= \sum_n |c_n|^2 \end{aligned} \tag{2.28}$$

Therefore, we conclude that modulus squares of the quantum amplitudes sum up to 1.

2.3 Matrix Formulation for Quantum Theory

The quantum theory we have introduced in this chapter so far is mostly based on Schrödinger's wave equation. In quantum computing however, Heisenberg's matrix formulation is dominantly used. Therefore, we do not dive deeper into Schrödinger's equation based approach and move our motivation through matrix formulation.

We start introducing the formulation of quantum theory in terms matrices by making a comparison between classical mechanical experiments and quantum mechanical experiments. The comparison is based on a very fundamental distinction between classical mechanics and quantum mechanics which leads to some advantages of quantum computing in terms of efficiency and speed-up. In a classical experiment, if the initial conditions and the dynamics of the system are perfectly known, then the outputs

of the measurement can be calculated in a deterministic manner using mathematical tools and models. Even in the flipping a coin experiment, if all the conditions that the coin will be exposed to are known, the output can be calculated. What makes a coin-flip is actually lack of information about the forces that will be in effect during the experiment.

On the other hand, in quantum mechanics, an experiment is analyzed in terms of 4 axioms of quantum theory as state, dynamics, measurement and observables. An experiment can be realized as a flow of 3 terms as

$$\text{state} \longrightarrow \text{dynamics} \longrightarrow \text{measurement.}$$

The terms observables is not included in the flow, since it actually corresponds to the interpretation of the experiment.

In a quantum experiment, even if the initial conditions and dynamics are perfectly known, the outputs can not be precisely determined since quantum theory is intrinsically probabilistic. Therefore, repeating an experiment in perfectly same conditions will yield the different outputs. However, an expectation value can be proposed, presented by the formula

$$\langle f \rangle_p = \sum f_i p_i \tag{2.29}$$

where f_i, p_i denote the distinct possible outputs and the probability of distinct possible outputs respectively.

Equation (2.29) is obviously is not very self-contained, but we will have a better explanation of it once we introduce the terms state, dynamics and measurement as components in Hilbert spaces. But for now we just state that the expected value of the experiment results are obtained by interpretation on the experiment and it corresponds to observables.

Implementations of the axioms states, dynamics, measurement and observables are performed via a calculation framework called "bra-ket notation" or "Dirac's nota-

tion". Within this notation, the inner product we have introduced via the symbol $\langle\ \rangle$ is substituted by $\langle\ |\ \rangle$.

Instead of considering quantum states as wave functions as we did before, we consider them as a vector $|\psi\rangle$ in a Hilbert Space $\mathbf{H}$. We also require $|\psi\rangle$ to be a unit vector, such that

$$||\psi\rangle| = 1 \implies \langle\psi|\psi\rangle = 1 \tag{2.30}$$

For the dynamics axiom, during an experiment, an initial state $|\psi_0\rangle$ is expected to change as it progresses through the dynamical components of the system. Therefore, in most basic terms , what happens in an experiment is the change of the initial state $|\psi_0\rangle$ to another state $|\psi_t\rangle$ via an operator U_t. We state this transition as

$$|\psi_t\rangle = U_t\,|\psi_0\rangle \tag{2.31}$$

We emphasize the condition that, the final state $|\psi_t\rangle$ is still a quantum state. Therefore it must have unit norm as well:

$$||\psi_t\rangle| = 1 \tag{2.32}$$

Transition of quantum states requires norm to be preserved and a geometrical approach would suggest that this could be done bu using rotations. For this reason, we place a condition on the U_t by stating that it has to be unitary. Therefore, U_t satisfies:

$$U_t^\dagger U_t = U_t U_t^\dagger = \mathbf{I} \tag{2.33}$$

where $\mathbf{I}$ is the identity matrix of convenient size.

For the measurement axiom, we consider a collection of operators M_i called POVM (Positive Operator-Valued Measure) where

- $M_i \geq 0$
- $\sum_i M_i = \mathbb{I}$.

Here, the condition is called the positivity condition and it actually means the probability value defined as

$$p_i = \langle\psi|M_i|\psi\rangle \geq 0.$$

Physically, each M_i in POVM corresponds to a detector being implemented in a quantum mechanical experiment with a probability of clicking p_i.
Under this setting of POVM, the general probabilistic condition $\sum p_i = 1$ is satisfied since

$$\begin{aligned}\sum_i p_i &= \sum_i \langle\psi|M_i|\psi\rangle \\ &= \langle\psi|\sum_i M_i|\psi\rangle \\ &= \langle\psi|\mathbf{I}|\psi\rangle \\ &= \langle\psi|\psi\rangle \\ &= 1.\end{aligned} \tag{2.34}$$

For single qubits, we set

$$M_0 = |0\rangle\langle 0| \tag{2.35}$$

and

$$M_1 = |1\rangle\langle 1|. \tag{2.36}$$

Observables are represented as Hermitian operators A satisfying $A = A^\dagger$. Since they are self-adjoint, they can be stated as a linear combination of M_i'ss with real coefficients as

$$A = \sum_i a_i M_i \tag{2.37}$$

where $a_i \in \mathbb{R}$ and M_i are the different projection operators corresponding to the different eigenvalues of A. Moreover, the expectation value of an observable A is

calculated as $\langle\psi| A |\psi\rangle$.

1-qubit states are stated as linear combinations of computational basis (B_c) states where $B_c = \{|0\rangle, |1\rangle\}$. For a 1-qubit state $|\psi\rangle$, we write

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \tag{2.38}$$

where $\alpha, \beta \in \mathbb{R}$ with $|\alpha|^2 + |\beta|^2 = 1$.

Equation (2.38) relies on superposition principle of quantum mechanics due to which a state can be a linear combination of basis states. Superposition principle is one of the main advantages of quantum computing comparing to the classical information processing systems. Another major advantage of quantum computing is entanglement and we will discuss it with many-qubit systems.

Using the matrix notation, Equation (2.38) can be restated as

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.39}$$

where

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since $\alpha, \beta \in S^3 \subset \mathbb{C}^2$, where S^3 is the unit sphere of topological dimension 3, they can be formulated as

$$\alpha = e^{ia} \cos\frac{\theta}{2} \tag{2.40}$$

and

$$\beta = e^{ib} \sin \frac{\theta}{2} \tag{2.41}$$

for $a, b \in \mathbb{R}$ and $0 \leq \theta \leq 2\pi$.

More explicitly, for 1-qubit state $|\psi\rangle$, we write

$$\begin{aligned} |\psi\rangle &= \alpha |0\rangle + \beta |1\rangle \\ &= e^{ia} \cos \frac{\theta}{2} |0\rangle + e^{ib} \sin \frac{\theta}{2} |1\rangle \\ &= e^{ia} \left(\cos \frac{\theta}{2} |0\rangle + e^{i(b-a)} \sin \frac{\theta}{2} |1\rangle \right) \\ &= e^{ia} |\psi_1\rangle \end{aligned} \tag{2.42}$$

where

$$|\psi_1\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i(b-a)} \sin \frac{\theta}{2} |1\rangle .$$

Here, we note that it is imposible to distinguish between $|\psi\rangle$ and $|\psi_1\rangle$ since the only difference between them is a phase shift:

$$\begin{aligned} \langle \psi_1 | M | \psi_1 \rangle &= \langle e^{ia} \psi | M | e^{ia} \psi \rangle \\ &= e^{-ia} e^{ia} \langle \psi | M | \psi \rangle \\ &= \langle \psi | M | \psi \rangle. \end{aligned} \tag{2.43}$$

2.3.1 Pauli Matrices and Some Other Quantum Gates

In this section we introduce Pauli matrices, which are unitary matrices corresponding to the dynamics changing the qubit states. They are also matrix representations of fundamental quantum computing gates, and one of the main problems in quantum computing is the realization of quantum mechanical circuits and systems through which quantum gates act on the qubits.

The Pauli matrices $\mathbb{I}, X, Y$ and Z have matrix representations,

$$\mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.44}$$

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{2.45}$$

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{2.46}$$

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.47}$$

As quantum computing gates, their actions on the computational basis states $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are as stated in the following.

$$\mathbb{I}\,|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle \tag{2.48}$$

$$\mathbb{I}\,|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle\,. \tag{2.49}$$

Obviously, $\mathbb{I}$ is the identity gate.

$$X\,|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle \tag{2.50}$$

$$X\,|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle\,. \tag{2.51}$$

Since X interchanges the computational basis states, it is also called quantum NOT-gate.

$$Y\,|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i\,|1\rangle \tag{2.52}$$

$$Y|1\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix} = -i|0\rangle \tag{2.53}$$

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle \tag{2.54}$$

$$Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -|1\rangle . \tag{2.55}$$

Pauli gates are especially important for quantum error correction schemes since they are used as basis functions both in representation and correction of the errors.

Another widely used quantum gate is the Hadamard gate. Hadamard gate provides a change of basis from computational basis to another basis called Hadamard basis which we denote as $B_{\pm} = \{|+\rangle, |-\rangle\}$. Also, the Hadamard gate is the Quantum Fourier Transform (QFT) for one-qubit systems.

The matrix formulation of the Hadamard gate is given as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.56}$$

The action of the Hadamard gate on the computational basis states $|0\rangle$ and $|1\rangle$ are given as

$$H|0\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = |+\rangle \tag{2.57}$$

$$H|1\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = |-\rangle \tag{2.58}$$

There is also a parametric gate, called phase gate,which only changes the phase of the qubit state. The matrix presentation of the phase gate is

$$P(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix} \tag{2.59}$$

with

$$P(\alpha)\,|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle \tag{2.60}$$

$$P(\alpha)\,|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ e^{i\alpha} \end{bmatrix} = e^{i\alpha}\,|1\rangle\,. \tag{2.61}$$

As we introduced the phase gate $P(\alpha)$, it is now a good time to underline a strange behavior of phase gates in order to prevent a misunderstanding on the effect of phases on the measurement outputs. As we already noted, two states are indistinguishable by measurement if they differ with phases only. However, due to interference, phase changes applied before some other gates has effects on the measurement results. Consider the following quantum algorithm desribed as in the following steps.

- **Step1:** Initialize the qubit state as $|0\rangle$.
- **Step 2:** Apply Hadamard gate.
- **Step 3:** Apply phase gate $P(\alpha)$.
- **Step 4:** Apply Hadamard gate again.
- **Step 5:** Do the measurement.

We know that if the phase gate is ignored, since $H^2 = \mathbf{I}$, we know that the the final state would expected to be $|0\rangle$ resulting in the probabilities $p_0 = 1$ and $p_1 = 0$. However, we would have different probability distributions for the circuit we have described above.

After Step 2, we will have $\frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle$. Applying the phase gate in Step 3 would take this state to $\frac{1}{\sqrt{2}}\,|0\rangle + e^{i\alpha}\frac{1}{\sqrt{2}}\,|1\rangle$.

After Step 4, we obtain,

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ e^{i\alpha}\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1+e^{i\alpha}}{2} \\ \frac{1-e^{i\alpha}}{2} \end{bmatrix} \tag{2.62}$$

Therefore, we would have

$$\begin{aligned} p_0 &= \left(\frac{1+e^{-i\alpha}}{2}\langle 0| + \frac{1-e^{-i\alpha}}{2}\langle 1|\right)|0\rangle\langle 0|\left(\frac{1+e^{i\alpha}}{2}|0\rangle + \frac{1-e^{i\alpha}}{2}|1\rangle\right) \\ &= \left(\frac{1+e^{-i\alpha}}{2}\right)\langle 0|\,|0\rangle\,\langle 0|\,|0\rangle\left(\frac{1+e^{i\alpha}}{2}\right) \\ &= \left(\frac{1+e^{-i\alpha}}{2}\right)\left(\frac{1+e^{i\alpha}}{2}\right) \\ &= \frac{1+\cos\alpha}{2} \end{aligned} \tag{2.63}$$

With a similar calculation (or even using $p_0 + p_1 = 1$), one can conclude

$$p_1 = \frac{1-\cos\alpha}{2}$$

$$p_2 = \frac{1+\cos\alpha}{2}$$

which are not the same probabilities with the former probabilities.

2.4 Many-qubit Systems and Entanglement

Many qubit systems in which information is encoded in a sequence of quantum single states is provided using tensor products. In classical information theory, actually, cartesian products are used to describe many-bit sytems. For example, a 2-bit sequence like 01 is just an information theoretic presentation of the cartesian product 0×1. an important thing to note about classical systems is when a new bit is added to the n-bit, then the number of different bit strings doubles since $2^{n+1} = 2^n \times 2$.

On the quantum computing scenario, tensor products are used to extend the systems.

In 2-qubit systems for example, we have

$$|00\rangle = |0\rangle \otimes |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$|01\rangle = |0\rangle \otimes |1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$|10\rangle = |1\rangle \otimes |0\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$|11\rangle = |1\rangle \otimes |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore, for two 1-qubit states $|\psi_1\rangle = \alpha_1 |0\rangle + \beta_1 |1\rangle$ and $|\psi_2\rangle = \alpha_2 |1\rangle + \beta_2 |1\rangle$, we have

$$\begin{aligned} |\psi_1\rangle \otimes |\psi_2\rangle &= (\alpha_1 |0\rangle + \beta_1 |1\rangle) \otimes (\alpha_2 |0\rangle + \beta_2 |1\rangle) \\ &= \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 |01\rangle + \beta_1\alpha_2 |10\rangle + \beta_1\beta_2 |11\rangle \end{aligned}$$

Using the matrix notation, the same tensor product of the single qubits, $|\psi_1\rangle \otimes |\psi_2\rangle$, can be formulated as:

$$|\psi_1\rangle \otimes |\psi_2\rangle = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \otimes \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_1\alpha_2 \\ \alpha_2\beta_2 \\ \beta_1\alpha_2 \\ \beta_1\beta_2 \end{bmatrix}$$

The idea of using tensor products for extending the systems are also applicable in the presentations of the quantum operators. For example, consider a 2-qubit state and

assume Hadamard gate and X gate are to be applied on the first and second qubits, respectively. Then, as a whole, this operation can be represented as

$$H \otimes X = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

Although the tensor product of two 1-qubit states results in a new 2-qubit state, in general, it is not true that all 2-states can be restated as tensor product of single products. This mathematical fact actually has a very important consequence in quantum mechanics, which could be explained in terms of entanglement. For example, consider the state $|\Psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ and assume it can be written as $|\varphi_1\rangle \otimes |\varphi_2\rangle$ where $|\varphi_1\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$ and $|\varphi_2\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$.

The assumption $|\Psi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$ would imply

$$\frac{1}{\sqrt{2}} = \alpha_1\alpha_2$$
$$0 = \alpha_1\beta_2$$
$$0 = \beta_1\alpha_2$$
$$\frac{1}{\sqrt{2}} = \beta_1\beta_2$$

which is impossible.

There exists an interesting 2-qubit gate whose action on one specific qubit depends on the other qubit. It is called CNOT gate and its actions on the basis states are as in the following.

$$CNOT|00\rangle = |00\rangle$$
$$CNOT|01\rangle = |01\rangle$$
$$CNOT|10\rangle = |11\rangle$$
$$CNOT|11\rangle = |10\rangle$$

According to the setting in the equations above for the 2-qubit states in the computational basis, the first qubit is the control qubit and the CNOT gate acts as an identity operator on it. However, this is not the case for the second qubit which is called as the target qubit. If the first qubit is $|0\rangle$, then the CNOT gate behaves like an identity operator on the second qubit as well. Otherwise, if the second qubit is $|1\rangle$, then it acts on the second qubit as X operator and changes its value.

A very important note about the CNOT gate is the fact that it does not involve a measurement on the first qubit. Any physical implementation of CNOT gate on a quantum computing device must handle the challenging process of interconnecting qubits in a way that the action needed to be taken on the target qubit is determined without a measurement on the first qubit. This is due to the quantum mechanical assumption that the quantum state collapses once it is measured. For example, assume we have a superposition qubit stated as

$$|\psi\rangle = \alpha\,|00\rangle + \beta\,|10\rangle\,.$$

If we attempt to perform a CNOT gate on$|\psi\rangle$, then as the the output state we would expect to have:

$$\begin{aligned} CNOT|\psi\rangle &= CNOT(\alpha\,|00\rangle + \beta\,|10\rangle) \\ &= CNOT(\alpha\,|00\rangle) + CNOT(\beta\,|10\rangle) \\ &= \alpha CNOT(|00\rangle) + \beta CNOT(|10\rangle) \\ &= \alpha\,|00\rangle + \beta\,|11\rangle \end{aligned}$$

which is again a superposition state. However, if we would perform a measurement on the first qubit of $|\psi\rangle = \alpha\,|00\rangle + \beta\,|10\rangle$ and then decide on the action on the second then, with probability $|\alpha|^2$ we would have the output state $|00\rangle$ and with probability $|\beta|^2$ we would have $|11\rangle$, but not a superposition of those basis states.

The CNOT gate, which fixes the states $|00\rangle$ and $|01\rangle$, but interchanges the states $|10\rangle$

and $|11\rangle$ has the following matrix representation.

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Any two-qubit level quantum computing system which can initialize the system in the state $|00\rangle$ and can perform Hadamard and CNOT gates can generate entangled qubit states. More explicitly, assume a quantum circuit which starts with the state $|00\rangle$. We first perform a Hadamard gate in the first qubit and get:

$$\begin{aligned}(H \otimes \mathbb{I})\,|0\rangle &= H(|0\rangle)\,|0\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,|0\rangle \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle).\end{aligned}$$

If we now perform a CNOT gate on the state, where the first qubit is the control qubit and the second qubit is the targe qubit, we get

$$\begin{aligned}CNOT(\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)) &= \frac{1}{\sqrt{2}}(CNOT(|00\rangle) + CNOT(|10\rangle)) \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).\end{aligned}$$

Here we have two notes about generation of entangled qubit states. For the entanglement generation, photonic systems have an advantage since SPDC (spontaneous parametric down conversion) type photon sources generate photon pairs which are entangled to each other. As a second note, the circuit mentioned above creates entangled pairs for the initial state $|00\rangle$ but not any generic two qubit state. In order to see this, assume that we have two qubit state described in the form of $(\alpha_1\,|0\rangle + \beta_1\,|1\rangle) \otimes (\alpha_2\,|0\rangle + \beta_2\,|1\rangle))$. If we apply a Hadamard gate on the first qubit,

(omitting the $\otimes$ signs) we have

$$
\begin{aligned}
&H(\alpha_1 |0\rangle + \beta_1 |1\rangle)(\alpha_2 |0\rangle + \beta_2 |1\rangle) \\
&= (\frac{\alpha_1}{\sqrt{2}} |0\rangle + \frac{\alpha_1}{\sqrt{2}} |1\rangle + \frac{\beta_1}{\sqrt{2}} |0\rangle \frac{\beta_1}{\sqrt{2}} |0\rangle)(\alpha_2 |0\rangle + \beta_2 |1\rangle) \\
&= (\frac{\alpha_1 + \beta_1}{\sqrt{2}} |0\rangle + \frac{\alpha_1 - \beta_1}{\sqrt{2}} |1\rangle)(\alpha_2 |0\rangle + \beta_2 |1\rangle) \\
&= \frac{(\alpha_1 + \beta_1)\alpha_2}{\sqrt{2}} |00\rangle + \frac{(\alpha_1 + \beta_1)\beta_2}{\sqrt{2}} |01\rangle \\
&+ \frac{(\alpha_1 - \beta_1)\alpha_2}{\sqrt{2}} |10\rangle + \frac{(\alpha_1 - \beta_1)\beta_2}{\sqrt{2}} |11\rangle .
\end{aligned}
$$

After applying a CNOT operation, we obtain

$$
\begin{aligned}
&CNOT\Big(\frac{(\alpha_1 + \beta_1)\alpha_2}{\sqrt{2}} |00\rangle\Big) + CNOT\Big(\frac{(\alpha_1 + \beta_1)\beta_2}{\sqrt{2}} |01\rangle\Big) + \\
&CNOT\Big(\frac{(\alpha_1 - \beta_1)\alpha_2}{\sqrt{2}} |10\rangle\Big) + CNOT\Big(\frac{(\alpha_1 - \beta_1)\beta_2}{\sqrt{2}} |11\rangle\Big) \\
&= \frac{(\alpha_1 + \beta_1)\alpha_2}{\sqrt{2}} |00\rangle + \frac{(\alpha_1 + \beta_1)\beta_2}{\sqrt{2}} |01\rangle + \frac{(\alpha_1 - \beta_1)\alpha_2}{\sqrt{2}} |11\rangle \\
&+ \frac{(\alpha_1 - \beta_1)\beta_2}{\sqrt{2}} |10\rangle .
\end{aligned}
$$

The resulting state is entangled in almost all cases but not in all cases. As a more specific example, consider applying Hadamard followed by a CNOT circuit to the qubit $|\varphi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)$.

With the Hadamard gate on the first qubit,

$$
\begin{aligned}
H \otimes \mathbb{I}(\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)) &= \frac{1}{2}(|00\rangle + |10\rangle) + \frac{1}{2}(|00\rangle - |10\rangle) \\
&= \frac{1}{2} |00\rangle + \frac{1}{2} |00\rangle \\
&= |00\rangle .
\end{aligned}
$$

After performing the CNOT gate, the resulting output state is obviously not entangled,

since,

$$\begin{aligned} CNOT\,|00\rangle &= |00\rangle \\ &= |0\rangle \otimes |0\rangle\,. \end{aligned}$$

2.5 Quantum State Discrimination

In a communication scheme based on quantum theory, the sender side, Alice, needs to prepare a set quantum states $|\Phi_i\rangle$ in which the intended message is encoded. The role of the quantum channel is to keep the states Φ_i unchanged or at least in a recoverable way. Assuming no error occurred during the preparation and transmission, the crucial part of the scheme is the measurement performed by the receiver side, Bob. Since the states collapses, the quantum nature of the system will vanish once the measurements are performed.

In a quantum communication scheme, Bob is not allowed to copy and store the qubits and perform repeated measurements on the qubit state. This is a due to a very fundamental fact of quantum mechanics known as no-cloning theorem.

Theorem 2.5.1. *[27] Unknown quantum states can not be cloned.*

Proof. Assume C is the cloning operator, i.e for any n-qubit quantum state $|\psi\rangle$ and a standard pure state $|e\rangle$,

$$C(|\psi\rangle\,|e\rangle) = |\psi\rangle\,|\psi\rangle\,.$$

Then for any two different n-qubit states $|\psi_i\rangle$ and $|\psi_j\rangle$, under the cloning operator assumption, we would have

$$\begin{aligned} C(|\psi_i\rangle\,|e\rangle) &= |\psi_i\rangle\,|\psi_i\rangle \\ C(|\psi_j\rangle\,|e\rangle) &= |\psi_j\rangle\,|\psi_j\rangle\,. \end{aligned}$$

Taking the inner product of the terms on the left hand side of the equations above, we would have

$$\begin{aligned}\langle e|\langle\psi_i|C^\dagger C|\psi_j\rangle|e\rangle &= \langle e|\langle\psi_i|\mathbb{I}|\psi_j\rangle|e\rangle \\ &= \langle e|\langle\psi_i|\psi_j\rangle|e\rangle \\ &= \langle e|e\rangle\langle\psi_i|\psi_j\rangle \\ &= \langle\psi_i|\psi_j\rangle\,.\end{aligned} \tag{2.64}$$

Similarly, taking the inner product of the terms on the right hand side of the equations, we would have

$$\begin{aligned}\langle\psi_i|\langle\psi_i|\psi_j\rangle|\psi_j\rangle &= \langle\psi_i|\psi_j\rangle\langle\psi_i|\psi_j\rangle \\ &= (\langle\psi_i|\psi_j\rangle)^2.\end{aligned} \tag{2.65}$$

Since the Equation (2.65) and Equation (2.66) must be equal, we have

$$\langle\psi_i|\psi_j\rangle = (\langle\psi_i|\psi_j\rangle)^2$$

which implies

$$\langle\psi_i|\psi_j\rangle = 0$$

or

$$\langle\psi_i|\psi_j\rangle = 1$$

The first case $\langle\psi_i|\psi_j\rangle = 0$ is the case when the two states $|\psi_i\rangle$ and $|\psi_j\rangle$ are orthogonal. The second case however, implies our cloning assumption is not true, the equations can only be obtained when $|\psi_i\rangle = |\psi_j\rangle$.
Therefore, there does not exist a unitary transform that acts on the all qubit states as a cloning map. Instead, if the state is known, an unitary transformation can be found copying only the known state.

Consequently, only orthogonal states can be cloned.

□

When a qubit state is prepared as superposition of basis states, the measurement outcome will definitely depend on the measurement operator applied. This is an important fact especially for quantum communication schemes. In the most general sense,

in quantum communication schemes , Alice sends quantum states to Bob and Bob is supposed to measure the states genuinely in order to achieve the best decoding of information sent by Alice. In order to increase the efficiency, the states Alice is allowed to prepare must have some restrictions.

Assume that Alice chooses her states from a finite set $\mathcal{K} = \{|\psi_n\rangle\}$ consisting of orthogonal quantum states. Then, for each $|\psi_i\rangle \in \mathcal{K}$, Bob can make measurements guided by the operators

$$M_i = |\psi_i\rangle \langle\psi_i| .$$

For each M_i, there exists another operator

$$D_i = \mathbb{I} - M_i = \mathbb{I} - |\psi_i\rangle \langle\psi_i|$$

which corresponds to no click for event i.

For the measurements, as we discussed in Section 2.3., Bob will apply a measurement which is a POVM. For M_i, we have,

$$\langle\psi_i| M_i |\psi_i\rangle = \langle\psi_i| |\psi_i\rangle \langle\psi_i| |\psi_i\rangle = 1.$$

When $i \neq j$, for the other POVM's M_j, we have

$$\langle\psi_i| M_j |\psi_i\rangle = \langle\psi_i| |\psi_j\rangle \langle\psi_j| |\psi_i\rangle = 0.$$

Therefore, when Alice prepares her states from a finite set of orthogonal states, then Bob will be able to discriminate the quantum states he received.

However, if the states prepared by Alice are not orthogonal and if they are chosen from an arbitrary set $S = \{|\varphi_n\rangle\}$, then we will have the following for the measurements:

$$\langle\varphi_i| M_j |\varphi_i\rangle = \langle\varphi_i| |\varphi_j\rangle \langle\varphi_j| |\varphi_i\rangle \neq 0$$

which makes it impossible to perfectly distinguish between the states. Therefore, the states prepared by Alice must be orthogonal.

So far, we have the following two priority observations on orthogonal quantum states.

- Only orthogonal states can be cloned.
- Only orthogonal states can be distinguished by measurements.

The first observation, the fact that only orthogonal states can be cloned results in some drawbacks for quantum computing devices. As the cloning is not possible in general, repetition code like approaches can not be performed in quantum error correction schemes. Moreover, the signal losses can not be solved by amplification, therefore, quantum repeaters and quantum data storage devices must handle with losses with different methods than classical ones. On the other hand, the first observation is useful in quantum key distribution since it restricts the attacker. The second observation is important for quantum communication protocols as we have already mentioned. However, the two observations of orthogonal states are quite related to each other. Actually, existence of one observation implies the other one.

To see the correspondence between the two observations, assume that Bob has received two states $|\varphi_0\rangle$ and $|\varphi_1\rangle$ that can not be perfectly distinguished where the bits 0 and 1 are encoded respectively. We also assume that these are the only options Alice have and she prepares $|\varphi_0\rangle$ with probability p_0 and $|\varphi_1\rangle$ with probability $p_1 = 1 - p_0$. For the measurement then, the probability of error, P_e is calculated as, ($P(x|y)$ being the conditional probability that the state $|x\rangle$ is the output of the measurement on $|y\rangle$.)

$$\begin{aligned} P_e &= P(\varphi_0)P(1|\varphi_0) + P(\varphi_1)P(0|\varphi_1) \\ &= p_0 \langle\varphi_0| M_1 |\varphi_0\rangle + p_1 \langle\varphi_1| M_0 |\varphi_1\rangle \\ &= p_0 \langle\varphi_0| M_1 |\varphi_0\rangle + p_1 \langle\varphi_1| \mathbb{I} - M_1 |\varphi_1\rangle \\ &= p_0 - Tr((p_0 |\varphi_0\rangle \langle\varphi_0| - p_1 |\varphi_1\rangle \langle\varphi_1|)M_0). \end{aligned}$$

Taking

$$|\varphi_0\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$$

and

$$|\varphi_1\rangle = \cos\theta\,|0\rangle - \sin\theta\,|1\rangle$$

for a convenient θ, the eigenvalues of the operator

$$(p_0\,|\varphi_0\rangle\,\langle\varphi_0| - p_1\,|\varphi_1\rangle\,\langle\varphi_1|)$$

turn out to be

$$\lambda_1 = \frac{1}{2}(p_0 - p_1 + \sqrt{1 - 4p_0p_1\cos^2 2\theta})$$
$$\lambda_2 = \frac{1}{2}(p_0 - p_1 - \sqrt{1 - 4p_0p_1\cos^2 2\theta})$$

implying,

$$P_e = \frac{1}{2}(1 - \sqrt{1 - 4p_0p_1|\langle\varphi_0|\varphi_1\rangle|^2}) \tag{2.66}$$

The statetement in Equation (2.67) is known as Helmstrom bound [14][3].

For n-qubit quantum systems where $p_0 = p_1 = \frac{1}{2}$, in [2] the Helmstrom bound is stated as

$$P_e = \frac{1}{2}|\langle\varphi_0|\varphi_1\rangle|^{2n}$$

Under this setting, we now introduce two impossible scenarios known as no-go theorems. In first scenario, assume that a perfect cloning transformation exists. Therefore, for any non-orthogonal pair $|\varphi_0\rangle\,|\varphi_1\rangle$ we can generate $(|\varphi_0\rangle\,|\varphi_1\rangle)^{\otimes n}$.Therefore, if arbitrarily large number of copies of $|\varphi_0\rangle\,|\varphi_1\rangle$ is sent instead of a single copy , then the probability of error in quantum state discrimination would approach 0 since

$$\lim_{n\to\infty}\frac{1}{2}|\langle\varphi_0|\varphi_1\rangle|^{2n} = 0$$

Therefore, if quantum cloning were possible, then quantum state discrimination would be possible as well.

As the second impossible scenario, assume that perfect quantum state discrimination

is possible with a perfect discriminator machine, D. Since D can distinguish between the state and decide whether it is Ψ_0 or Ψ_1, one could use D for discriminating between the states and then create arbitrarily many copies of the known state.
Hence, as we discussed above, no-cloning and indistinguishability actually implies each other. In the next section, we will describe a quantum key distribution protocol which directly emphasize the importance of no-cloning and indistinguishability in application.

2.6 Quantum Key Distribution (QKD)

Despite the fact that most of the recent developments and achievements in modern cryptography are in the field of asymmetric cryptography, theoretically speaking the most secure cryptographic protocol ever is a symmetric protocol called One-Time Pad (OTP). In OTP protocol, the sender and the receiver parties must agree on a key beforehand, and the length of key must be equal to the length of the message being transmitted. This is naturally a big drawback and the security of the protocol decreases if the same key is used more than once. If the key is used twice, then an attacker would be able to get the sum of the messages sent in the binary form and perform a succesful attack analysing the patterns of the language being used or the communication routines.

In 1984, a quantum mechanics based data transmission protocol was proposed in order to overcome the key sharing problem by Charles Bennett and Gilles Brassard [5]. The protocol named BB84 uses the polarizations of the photons in order to encode the classical bits 0 and 1. As the sender, Alice needs to prepare and send the photons to the receiver side, Bob. Bob is in the charge of receiving the photons and measure them. Once Bob completes the measurement process, they also need a classical communication channel in order to agree on the key as a part of the information being sent by Alice through the quantum channel. The classical channel used after quantum communication need not be encrypted, but it has to be authentic.

Before we dive into the steps of the BB84 protocol, we first study two different bases $\mathcal{B}_1$ and $\mathcal{B}_2$ for 1-qubit systems. The bases are connected to each other with Equations (2.57) and (2.58) and they are defined as

$$\mathcal{B}_1 = \{|0\rangle, |1\rangle\} \tag{2.67}$$

and

$$\mathcal{B}_2 = \{|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\} \tag{2.68}$$

For the states in $\mathcal{B}_1$ we have,

$$\begin{aligned}
\langle 0|0\rangle &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \\
\langle 0|1\rangle &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \\
\langle 1|0\rangle &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \\
\langle 1|1\rangle &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1.
\end{aligned} \tag{2.69}$$

From the equations listed in (2.69) we understand that the states $|0\rangle$ and $|1\rangle$ are orthogonal to each other. Therefore, they can be cloned and moreover they can also be perfectly distinguished from each other by applying convenient measurement operators. Since they are distinguishable, it is a good idea to use them for quantum communication protocols.

Similary, for the two states in $\mathcal{B}_2$, we have

$$
\begin{aligned}
\langle +|+\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 1 \\
\langle +|-\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = 0 \\
\langle -|+\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 0 \\
\langle -|-\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = 1.
\end{aligned}
\tag{2.70}
$$

Therefore, as seen in Equation (2.70), the states $|+\rangle$ and $|-\rangle$ are clonable and distinguishable as well. That makes the states in $\mathcal{B}_2$ a good choice for quantum communication protocols like the states in $\mathcal{B}_1$.

Moreover, if we consider the inner products of the vectors from different bases $\mathcal{B}_1$ and $\mathcal{B}_2$ we observe pattern as a result of calculations:

$$
\begin{aligned}
\langle +|0\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \\
\langle +|1\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \\
\langle -|0\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \\
\langle -|1\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{1}{\sqrt{2}}.
\end{aligned}
\tag{2.71}
$$

Therefore, for two different quantum states $|u\rangle$ and $|v\rangle$ are taken from different sets

$\mathcal{B}_1$ and $\mathcal{B}_2$, we always have

$$|\langle u|v\rangle| = \frac{1}{\sqrt{2}}. \tag{2.72}$$

The pattern in the Equation (2.72) is actually very useful. First, we were already convinced that using the state vectors in $\mathcal{B}_1$ and $\mathcal{B}_2$ for a quantum communication scheme was a good idea because they were clonable and distinguishable within each other. With the Equations (2.72) we understand that even if Alice and Bob employs their preparation and measurement process with different choices, the probabilty distributions of measurement outputs will be the same. Therefore, neither of the bases choices offer an advantage to Alice and Bob. For example, if Alice decides to send $|+\rangle$ to transmit the information 0, and if Bob decides to measure it with respect to the measurement operator corresponding to the states in $\mathcal{B}_1$, the probability that he will measure 0 or 1 is one-half.

In BB84 quantum key distribution protocol, we assume that there exist a quantum and classical channels between Alice and Bob. Also Alice is able to prepare photons and Bob is able to detect them. When Alice prepares photons, she has to pick between the computational basis $\mathcal{B}_1$ and the diagonal basis $\mathcal{B}_2$. For each bit of information being sent, there are 4 scenarios:

- Alice wants to send 0 and choses $\mathcal{B}_1$, then she prepares the photon in the state $|0\rangle$.
- Alice wants to send 0 and choses $\mathcal{B}_2$, then she prepares the photon in the state $|+\rangle$.
- Alice wants to send 1 and choses $\mathcal{B}_1$, then she prepares the photon in the state $|1\rangle$.
- Alice wants to send 1 and choses $\mathcal{B}_2$, then she prepares the photon in the state $|-\rangle$.

After the preparation, Alice sends the photons to Bob via the quantum channel. Since Bob does not know about Alice's basis choice, he has to make a random guess. due

to Equation (2.72), neither of the choices gives Bob an advantage. For Bob, we have two scenarios:

- Bob makes the same basis choice with Alice. Then he will get the correct output with probability 1.
- Bob makes the different basis choice with Alice. Then he will get the correct output with probability $\frac{1}{2}$.

After Bob's measurements on the transmitted photons, Alice and Bob use the authenticated classical channel in order to reveal the basis they used for each photon in the correct order. After receiving this information, they can determine the orders of the photons they encoded using the same bases which is expected to be half of the cases. On their measurement results, they discard the bits where they used distinct basis and keep the parts where their selections were the same. If N photons were sent from Alice to Bob, they now have $\frac{N}{2}$ remaining bits.

In order to make sure that their communication was not eavesdropped, they compare some part of the remaining $\frac{N}{2}$ bits. When the transmission is eavesdropped by a malicious actor Eve, she receives the photon send by Alice, meaasures it with respect to a basis of her own choice. After the measurement, Eve sends a new photon to Bob encoded with Eve's measurement result. But in half of the cases, Eve will use a different basis rather than the basis Alice and Bob agreed on. For this reason, there will some bits where Alice and Bob use the same basis but still obtain distinct measurement results.

As an example, assume that both Alice and Bob use the computational basis for the bit 0. Then Alice needs to prepare the state $|0\rangle$ and send it to Bob via the quantum channel eavesdropped by Eve. If Eve selects the diagonal basis and measures the state $|0\rangle$, with probability $\frac{1}{2}$ she will obtain the result 0 and with probability $\frac{1}{2}$, she will get 1. Therefore, she will prepere either $|+\rangle$ or $|-\rangle$ based on her measurement result. In any case, with probability $\frac{1}{2}$ Bob will get 0 as the result of his measurement and

Alice will remain unnoticed. However, Bob will obtain 1 with probability $\frac{1}{2}$ which is not the value intended to be transmitted by Alice.

Therefore, when Eve chooses the bases which Alice and Bob use, she remains unnoticed for sure. Therefore, at least in half of the bits she leaves no noticable symptoms. For the other half, i.e. when her basis is different, with probability $\frac{1}{2}$, she again remains unnoticed, but leaves symptoms with probability $\frac{1}{2}$. Therefore, out of the remaining $\frac{N}{2}$ bits, $\frac{N}{2} \times \frac{1}{4} = \frac{N}{8}$ of them will differ due to existence of Eve. The different bits can be detected when Alice and Bob share some part of their $\frac{N}{2}$ bits and compare them.

In the case the attacker Eve places herself in the classical channel, where the base selections are transmitted after Bob's measurements, again she will have no advantages since the quantum part of the transmission is already over and she already made her choices without knowing the selections of Alice and Bob. Therefore, there is no need to consider encrpyting the classical channel in the real-world implementations of BB84 protocol as long as the authenticity of the channel is provided.

2.7 Quantum Fourier Transform

In this section, we will describe Quantum Fourier Transformation (QFT) which requires implementation of powers of roots of unity as coeffients to basis state vectors. QFT is analogous to classical type Fourier Transforms and it is of critical importance for many quantum algorithms.

Let $|j\rangle$ be a basis state in a n-qubit level quantum system. The quantum Fourier transform (QFT) is analogous to inverse discrete Fourier transform and it is defined as

$$QFT(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \xi_N^{jk} |k\rangle$$

where $N = 2^n$ is the dimension of the complex vector space and ξ_N is a primitive N-th root of unity.

An alternative way of stating the formulation of the Quantum Fourier Transform would be

$$QFT = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} \xi_N^{jk} \left|k\right\rangle \left\langle j\right| .$$

On an arbitrary basis state $\left|\alpha\right\rangle = \left|\alpha_1 \alpha_2 ... \alpha_n\right\rangle$ of an n-qubit quantum system, QFT acts as in the following:

$$\begin{aligned} QFT \left|\alpha\right\rangle &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \xi_N^{\alpha k} \left|k\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{2\pi i \alpha k}{2^n}} \left|k\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{2\pi i \alpha (\sum_{m=1}^{n} k_m 2^{n-m})}{2^n}} \left|k\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi \alpha i (\sum_{m=1}^{n} \frac{k_m}{2^m})} \left|k_1 k_2 ... k_n\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \prod_{m=1}^{n} e^{\frac{2\pi i \alpha k_m}{2^m}} \left|k_1 k_2 ... k_m\right\rangle \\ &= \frac{1}{\sqrt{N}} \bigotimes (\left|0\right\rangle + e^{\frac{2\pi i \alpha}{2^m}} \left|1\right\rangle . \end{aligned} \tag{2.73}$$

More explicitly,

$$QFT(\left|\alpha\right\rangle) = \frac{1}{\sqrt{N}} (\left|0\right\rangle + e^{\frac{2\pi i \alpha}{2}} \left|1\right\rangle) \otimes (\left|0\right\rangle + e^{\frac{2\pi i \alpha}{2^2}} \left|1\right\rangle) \otimes \cdots \otimes (\left|0\right\rangle + e^{\frac{2\pi i \alpha}{2^n}} \left|1\right\rangle). \tag{2.74}$$

In order to conclude that Quantum Fourier Transform can be implemented in a quantum algorithm as an operational component, we nend to show it is a unitary operator. Note that the implementation QFT is in the core of many well-known quantum algorithms like Shor's algorithm.

Lemma 2.7.1. *QFT is a unitary operator.*

Proof. We need to show $(QFT)(QFT)^\dagger = I$

$$\begin{aligned}
(QFT)(QFT)^\dagger &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} \xi_N^{jk} |k\rangle \langle j| \frac{1}{\sqrt{N}} \sum_{k'=0}^{N-1} \sum_{j'=0}^{N-1} \xi_N^{-j'k'} |k'\rangle \langle j'| \\
&= \frac{1}{N} \sum_{j,k,j',k'} \xi_N^{(jk-j'k')} |k\rangle \langle j | j'\rangle \langle k'| \\
&= \frac{1}{N} \sum_{j,k,j',k'} \xi_N^{(jk-j'k')} |k\rangle \, \delta_{j,j'} \langle k'| \\
&= \frac{1}{N} \sum_{k,j,k'} \xi_N^{(j(k-k'))} |k\rangle \langle k'| \\
&= \sum_{k,k'} \delta_{k,k'} |k\rangle \langle k'| \\
&= \sum_{k} |k\rangle \langle k| \\
&= 1.
\end{aligned}$$

□

In 1-qubit quantum systems, where $n = 1$, $N = 2^1 = 2$, the quantum Fourier transform and the Hadamard gate coincide with each other since their actions on the basis state vectors $|0\rangle$ and $|1\rangle$ are the same.

$$\begin{aligned}
QFT\,|0\rangle &= \frac{1}{\sqrt{2}} \sum_{k=0}^{1} \xi_2^0 |k\rangle \\
&= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\
&= |+\rangle \\
&= H(|0\rangle)
\end{aligned}$$

and

$$\begin{aligned}
QFT\,|1\rangle &= \frac{1}{\sqrt{2}} \sum_{k=0}^{1} \xi_2^k |k\rangle \\
&= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
&= |-\rangle \\
&= H(|1\rangle).
\end{aligned}$$

CHAPTER 3

MUTUALLY UNBIASED BASES AND SOME OF THEIR CONSTRUCTIONS

Mutually unbiased bases are a specific kind of bases of Hilbert spaces with critical importance for quantum information theory. In quantum information theory, data is represented as a linear combinations of basis vectors thanks to the superposition principle of quantum mechanics. The measurement operators, as we seen in the Chapter 2 are also defined with respect to the choice of basis states underlying the importance of the selection of the bases.

Mutually unbiased bases are family of bases which encode the information in a way that when the measurement is performed with respect to another basis from the family, the output probability distribution of each basis state is equal. In formal terms, mutually unbiased bases are defined as in the following.

Definition 3.0.1. *Let* $\mathbf{H}$ *be a Hilbert space of dimension* n *and let* $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ *and* $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$ *be two bases of* $\mathbf{H}$*. Then,* $\mathcal{E}$ *and* $\mathcal{F}$ *are called mutually unbiased bases (MUBs) if the following conditions are satisfied:*

- $\langle e_i|e_j\rangle = 0 = \langle f_i|f_j\rangle$ *when* $i \neq j$.
- $\langle e_i|e_i\rangle = 1 = \langle f_i|f_i\rangle$ *for all* $1 \leq i \leq n$.
- $|\langle e_i|f_j\rangle|^2 = \frac{1}{n}$ *for all* $1 \leq i \leq n$ *and* $1 \leq j \leq n$.

The very first introduction of mutually unbiased bases was provided by Schwinger in

1960 [24]. 20 years later, in 1980, Alltop described a way of constructing mutually unbiased bases without realizing the correspondence of the mutually unbiased bases with quantum mechanics and quantum information theory [1]. Ivanovic was the first one considering the applications of mutually unbiased bases in the quantum side[16]. In this chapter, we first introduce some algebraic concepts needed in order to a have good understanding of some constructions of mutually unbiased bases. Later, we will describe some methods to construct mutually unbiased bases.

3.1 Characters and Their Properties

In this section we introduce the characters of groups and some of their properties. Since fields are intrinsically made up of two groups, we will also describe the use of characters on field. Our main reference in this section is [20].

Definition 3.1.1. *Given a finite abelian group G, a character on G is an homomorphism* $\chi : G \longrightarrow S^1 \subset \mathbb{C}$.

We note that S^1 is the set of complex numbers whose norm is 1, basically it is the unit circle. Therefore, any point in S^1 can obviously be stated as $e^{i\theta}$. Since homomorphisms are defined between groups, we emphasize the group structure of S^1. Given any two points $e^{i\theta_1}$, $e^{i\theta_2}$ their ordinary product $e^{i\theta_1}e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$ is also on S^1. Also $e^0 = 1 \in S^1$ and for any θ, $e^{i\theta}e^{-i\theta} = 1$. Moreover, the multiplicative group S^1 is isomorphic to special orthogonal group $SO(2)$. Indeed, there is a lot to say about the group structure of S^1, like its Lie group structure and its topological aspect, but they are not directly related to this study.

For a finite cyclic group G, the characters on G can be described as in the following. Since G is assumed to be cyclic, there exists an element $g \in G$ such that $G = |g|$ and any element of G is of the form g^k for some integer k where $0 \leq k \leq n-1$ where n is the order of G. For any integer j such that $0 \leq j \leq n-1$, define

$$\chi_j(g^k) = e^{\frac{2\pi ijk}{n}}. \tag{3.1}$$

Since χ_j maps G into the unit circle S^1 and it satisfies

$$\chi_j(g^k g^l) = \chi(g^{(k+l)}) = e^{\frac{2\pi ij(k+l)}{n}} = e^{\frac{2\pi ijk}{n}} e^{\frac{2\pi ijl}{n}} = \chi_j(g^k)\chi_j(g^l)$$

it is a character on G. Moreover, for any other character χ on the cyclic group G, since

$$(\chi(g))^n = \chi(g^n) = \chi(e) = 1,$$

(where e is the identity element in G), $\chi(g)$ must be an n-th root of unity. Therefore, $\chi(g) = e^{\frac{2\pi ij}{n}}$ for some integer j.

Given a finite abelian group G, the set of characters on G is denetod by $G^\wedge$. $G^\wedge$ is actually another group with respect to the multiplication operation. Moreover, as $G^\wedge$ is itself a group, it has its own group of characters constructed in terms of characters of G as in the following.

Considering $G^\wedge$ as a group of characters of a finite abelian group G, define the homomorphism

$$\overline{g} : G^\wedge \longrightarrow S^1$$
$$\overline{g}(\chi) = \chi(g)$$

for any $g \in G$. Obviously, for any $\chi_1, \chi_2 \in G^\wedge$,

$$\overline{g}(\chi_1\chi_2) = (\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g) = \overline{g}(\chi_1)\overline{g}(\chi_2).$$

The group of characters on $G^\wedge$ is denoted by $G^{\wedge\wedge}$. Fixing an element of $G^\wedge$ and $G^{\wedge\wedge}$ respectively we get the following two lemmas.

Lemma 3.1.2. *For a fixed $\chi \in G^\wedge$ where χ is not identically 1, $\sum_{g\in G} \chi(g) = 0$*

Proof. Since χ is not identically 1, there exists an $a \in G$ such that $\chi(a) \neq 1$. Then

$$\chi(a)\sum_{g\in G}\chi(g) = \sum_{g\in G}\chi(a)\chi(g) = \sum_{g\in G}\chi(ag) = \sum_{g\in G}\chi(g)$$

which implies $\sum_{g\in G}\chi(g)=0$. □

We also have a similar lemma for $G^{\wedge\wedge}$ whose proof is very similar to the proof of Lemma 1. Therefore we state the lemma but omit the proof.

Lemma 3.1.3. *For $g\in G$ which not the identity element of G, $\sum_{\chi\in G^{\wedge}}\chi(g)=0$.*

In the following lemma, we state a correspondence between G and $G^{\wedge}$ in terms of their orders as groups.

Lemma 3.1.4. *For any finite abelian group G, $|G|=|G^{\wedge}|$.*

Proof. Consider the double summation

$$\sum_{g\in G}\sum_{\chi\in G^{\wedge}}\chi(g).$$

during the iteration on G, by Lemma 3.2.3, the interior summand $\sum\limits_{\chi\in G^{\wedge}}\chi(g)$ does not vanish only when g is the identity element of G. When g is the identity element, $\chi(g)=1$, therefore,

$$\sum_{g\in G}\sum_{\chi\in G^{\wedge}}\chi(g)=\sum_{\chi\in G^{\wedge}}\chi(g)=\sum_{\chi\in G^{\wedge}}1=|G^{\wedge}|$$

Similarly, we also have (taking χ_e as the trivial character),

$$\sum_{\chi\in G^{\wedge}}\sum_{g\in G}\chi(g)=\sum_{g\in G}\chi_e(g)=\sum_{g\in G}1=|G|.$$

Therefore, we conclude that $|G|=|G^{\wedge}|$.

□

The following theorem introduces the orthogonality conditions on the characters. We note that for a character χ, its complex conjugate is another character and it is denoted as $\overline{\chi}$.

Theorem 3.1.5. *Given two characters χ,φ on G,*

$$\frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\varphi(g)}=\begin{cases}1 & \text{if } \chi=\varphi\\ 0 & \text{if } \chi\neq\varphi\end{cases}$$

and given two elements $g, h \in G$

$$\frac{1}{|G|}\sum_{\chi\in G^{\wedge}}\chi(g)\overline{\chi(h)} = \begin{cases} 1 & \text{if } g = h \\ 0 & \text{if } g \neq h \end{cases}$$

Proof. We note that only first equation will be proved since the second equation can be proved in a quite similar way.

Given two characters χ and φ on G, noting that $\overline{\varphi}$ is again a homomorphism, the product $\chi\overline{\varphi}$ is another character on G since it still maps G into S^1 and the product of two homomorphisms is again a homomorphism. In the case $\chi = \varphi$, the product $\chi\overline{\varphi}$ is the trivial character identically equal to 1.

However, when $\chi = \varphi$,

$$\frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\varphi(g)} = \frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\chi(g)} = \frac{1}{|G|}\sum_{g\in G}1 = 1.$$

For the other case, i.e when $\chi \neq \varphi$, the product $\chi\varphi$ is a non-trivial character on G and by Lemma 3.2.2,

$$\sum_{g\in G}\chi(g)\varphi(g) = 0.$$

Therefore, when $\chi \neq \varphi$,

$$\frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\varphi(g)} = 0.$$

The proof for the second equation can be obtained quite similarly. □

We now extend the descriptions of the characters and their properties from groups to finite fields. Since any field already contains two group structures (additive and multiplicative groups) it is natural to describe two distinct character structures on fields. We start with the additive groups.

Let $\mathbb{F}_q$ be the finite field with $q = p^m$ for a prime integer p. All the elements of $\mathbb{F}_q$ form a group with respect to addition. A well-known way of moving from $\mathbb{F}_q$ to the prime field $\mathbb{F}_p$ is to employ trace function, defined as

$$\mathrm{Tr}(c) = c + c^q + \cdots + c^{q^{m-1}} \tag{3.2}$$

for any $c \in \mathbb{F}_q$.

Note that $Tr : \mathbb{F}_q :\longrightarrow \mathbb{F}_p$ is already a homomorphism between the additive groups of $\mathbb{F}_q$ and $\mathbb{F}_p$ since

$$Tr(c_1 + c_2) = Tr(c_1) + Tr(c_2)$$

for any $c_1, c_2 \in \mathbb{F}_q$. As a result of this fact, the canonical additive character χ_1 on the finite field $\mathbb{F}_q$ is defined as

$$\chi_1(c) = e^{2\pi i Tr(c)/p}$$

for all $c \in \mathbb{F}_q$.

Under this setting, for any $c_1, c_2 \in \mathbb{F}_q$, we have

$$\begin{aligned}
\chi_1(c_1 + c_2) &= e^{2\pi i Tr(c_1+c_2)/p} \\
&= e^{2\pi i (Tr(c_1)+Tr(c_2))/p} \\
&= e^{2\pi i Tr(c_1)/p} e^{2\pi i Tr(c_2)/p} \\
&= \chi_1(c_1)\chi_1(c_2).
\end{aligned}$$

Comparing to the descriptions of characters on additive groups on finite field, the descriptions of characters on multiplicative groups are more simple. For this, let $\mathbb{F}_q^*$ denote the multiplicative group of $\mathbb{F}_q$. Also, let g be a fixed primitive element of $\mathbb{F}_q$. Then the function φ_j defined as

$$\varphi_j(g^k) = e^{2\pi i jk/(q-1)}$$

where j is an integer with $0 \leq j \leq q-2$ is a character of $\mathbb{F}_q^*$.

We have the following orthogonality and summation conditions for additive characters χ_a , χ_b and multiplicative characters φ, τ on finite fields.

$$\sum_{c \in \mathbb{F}_q} \chi_a(c)\overline{\chi_b(c)} = \begin{cases} 0 & \text{if } a \neq b \\ q & \text{if } a = b \end{cases} \tag{3.3}$$

$$\sum_{c\in\mathbb{F}_q}\chi_a(c)=0 \text{ when } a\neq 0 \tag{3.4}$$

$$\sum_{b\in\mathbb{F}_q}\chi_b(c)\overline{\chi_b(d)}=\begin{cases}0 & \text{if } c\neq d\\ q & \text{if } c=d\end{cases} \tag{3.5}$$

$$\sum_{c\in\mathbb{F}_q^*}\varphi(c)\overline{\tau(c)}=\begin{cases}0 & \text{if } \varphi\neq\tau\\ q-1 & \text{if } \varphi=\tau\end{cases} \tag{3.6}$$

$$\sum_{c\in\mathbb{F}_q^*}\varphi(c)=0 \text{ when } \varphi\neq\varphi_0 \tag{3.7}$$

$$\sum_{\varphi}\varphi(c)\overline{\varphi(d)}=\begin{cases}0 & \text{if } c\neq d\\ q-1 & \text{if } c=d\end{cases} \tag{3.8}$$

Given a multiplicative character φ and an additive character χ on a finite field $\mathbb{F}_q$, the summation formulated as $G(\varphi,\chi)=\sum_{c\in\mathbb{F}_q^*}\varphi(c)\chi(c)$ is called a Gaussian sum. We have the following theorem for the Gaussian sums.

Theorem 3.1.6. *For a multiplicative character φ and an additive character χ on $\mathbb{F}_q$, the Gaussian sum $G(\varphi,\chi)$ satisfies*

$$G(\varphi,\chi)=\begin{cases}q-1 & \textit{if } \varphi=\varphi_0 \textit{ and } \chi=\chi_0\\ -1 & \textit{if } \varphi=\varphi_0 \textit{ and } \chi\neq\chi_0\\ 0 & \textit{if } \varphi\neq\varphi_0 \textit{ and } \chi=\chi_0\end{cases}$$

In the only remaining case, where $\varphi\neq\varphi_0$ and $\chi\neq\chi_0$, we have

$$|G(\varphi,\chi)|=q^{1/2}.$$

Proof. We only give a proof for the last case ($\varphi \neq \varphi_0$ and $\chi \neq \chi_0$) since the first three cases can easily be obtained from the orthogonality conditions.

For the only remaining case, note that for any complex number z, $|z|^2 = z\overline{z}$. Therefore

$$\begin{aligned}
|G(\varphi,\chi)|^2 &= G(\varphi,\chi)\overline{G(\varphi,\chi)} \\
&= \sum_{c\in\mathbb{F}_q^*}\sum_{c_1\in\mathbb{F}_q^*} \overline{\varphi(c)\chi(c)}\varphi(c_1)\chi(c_1) \\
&= \sum_{c\in\mathbb{F}_q^*}\sum_{c_1\in\mathbb{F}_q^*} \varphi(c^{-1}c_1)\chi(c_1-c).
\end{aligned}$$

Take $c^{-1}c_1 = d$. Then,

$$c_1 - c = c(c_1c^{-1} - 1) = c(d-1).$$

Therefore,

$$\begin{aligned}
|G(\varphi,\chi)|^2 &= \sum_{c\in\mathbb{F}_q^*}\sum_{d\in\mathbb{F}_q^*} \varphi(d)\chi(c(d-1)) \\
&= \sum_{d\in\mathbb{F}_q^*} \varphi(d)\Big(\sum_{c\in\mathbb{F}_q} \chi(c(d-1)) - \chi(0)\Big) \\
&= \sum_{d\in\mathbb{F}_q^*} \varphi(d)\Big(\sum_{c\in\mathbb{F}_q} \chi(c(d-1))\Big)
\end{aligned}$$

When $d = 1$, the sum $\sum_{c\in\mathbb{F}_q} \chi(c(d-1))$ turns out to be

$$\sum_{c\in\mathbb{F}_q} \chi(0) = \sum_{c\in\mathbb{F}_q} 1 = q.$$

On the other hand, when $d \neq 1$,

$$\sum_{c\in\mathbb{F}_q} \chi(c(d-1)) = 0$$

Therefore,

$$\begin{aligned}
|G(\varphi,\chi)|^2 &= \sum_{d\in\mathbb{F}_q^*} \varphi(d)\Big(\sum_{c\in\mathbb{F}_q} \chi(c(d-1))\Big) \\
&= \varphi(1)\Big(\sum_{c\in\mathbb{F}_q} \chi(0)\Big) \\
&= q
\end{aligned}$$

which implies

$$|G(\varphi, \chi)| = \sqrt{q}.$$

□

By using the orthogonality conditions of characters, characters can be written in terms of Gaussian sums. The following lemma offers a restatement of characters using Gaussian sums.

Lemma 3.1.7. *Let φ and χ be multiplicative and additive characters of the finite field $\mathbb{F}_q$. Then, for any $c \in \mathbb{F}_q^*$,*

- $\varphi(c) = \frac{1}{q} \sum_{\chi} G(\varphi, \overline{\chi})\chi(c)$

- $\chi(c) = \frac{1}{q-1} \sum_{\chi} G(\overline{\varphi}, \chi)\varphi(c)$.

Proof. For the multiplicative character φ, we claim that

$$\varphi(c) = \frac{1}{q} \sum_{d \in \mathbb{F}_q^*} \varphi(d) \sum_{b \in \mathbb{F}_q} \chi_b(c)\overline{\chi_b(d)}. \tag{3.9}$$

In Equation (3.9) above, we note that by the orthogonality condition (Equation 3.7), the summand $\sum_{b \in \mathbb{F}_q} \chi_b(c)\overline{\chi_b(d)}$ is equal to q when $c = d$ and vanishes for all other cases. Then we can write

$$\begin{aligned} \frac{1}{q} \sum_{d \in \mathbb{F}_q^*} \varphi(d) \sum_{b \in \mathbb{F}_q} \chi_b(c)\overline{\chi_b(d)} &= \frac{1}{q}\varphi(c) \sum_{b \in \mathbb{F}_q} \chi_b(c)\overline{\chi_b(c)} \\ &= \frac{1}{q}\varphi(c)q \\ &= \varphi(c). \end{aligned}$$

Therefore

$$\begin{aligned} \varphi(c) &= \frac{1}{q} \sum_{d \in \mathbb{F}_q^*} \varphi(d) \sum_{b \in \mathbb{F}_q} \chi_b(c)\overline{\chi_b(d)} \\ &= \frac{1}{c} \sum_{b \in \mathbb{F}_q} \chi_b(c) \sum_{d \in \mathbb{F}_q^*} \varphi(d)\overline{\chi_b}(d) \\ &= \frac{1}{q} \sum_{\chi} G(\varphi, \overline{\chi})\chi(c). \end{aligned}$$

Similarly, for the additive character χ, we write

$$\begin{aligned}\chi(c) &= \frac{1}{q-1}\sum_{d\in\mathbb{F}_q^*}\chi(d)\sum_{\varphi}\varphi(c)\overline{\varphi(d)}\\ &= \frac{1}{q-1}\sum_{\varphi}\varphi(c)\sum_{d\in\mathbb{F}_q^*}\overline{\varphi}(d)\chi(d)\\ &= \frac{1}{q-1}\sum_{\varphi}G(\overline{\varphi},\chi)\varphi(c).\end{aligned}$$

□

Theorem 3.1.8. *Let χ be a nontrivial additive character of $\mathbb{F}_q$, $n\in\mathbb{N}$ and λ a multiplicative character of $\mathbb{F}_q$ of order $d = gcd(n, q-1)$. Then*

$$\sum_{c\in\mathbb{F}_q}\chi(ac^n+b) = \chi(b)\sum_{j=1}^{d-1}\overline{\lambda^j}(a)G(\lambda^j,\chi)$$

for any $a, b\in\mathbb{F}_p$ with $a\neq 0$.

Proof. Defining a new non-trivial additive character τ where $\tau(c) = \chi(ac)$, the summation $\sum_{c\in\mathbb{F}_q}\chi(ac^n+b)$ can be reformulated as

$$\begin{aligned}\sum_{c\in\mathbb{F}_q}\chi(ac^n+b) &= \chi(b)\sum_{c\in\mathbb{F}_q}\chi(ac^n)\\ &= \chi(b)\sum_{c\in\mathbb{F}_q}\tau(c^n).\end{aligned} \tag{3.10}$$

Since τ is an additive character, by Lemma 3.2.7, we can state

$$\tau(c^n) = \frac{1}{q-1}\sum_{\varphi}G(\overline{\varphi},\tau)\tau(c^n)$$

Moreover,

$$\begin{aligned}\sum_{c\in\mathbb{F}_q}\tau(c^n) &= \sum_{c\in\{0\}\cup\mathbb{F}_q^*}\tau(c^n)\\ &= \tau(0)+\sum_{c\in\mathbb{F}_q^*}\tau(c^n)\\ &= 1+\frac{1}{q-1}\sum_{\varphi}G(\overline{\varphi},\tau)\sum_{c\in\mathbb{F}_q^*}\varphi^n(c).\end{aligned}$$

The very last part of the equation, $\sum_{c\in\mathbb{F}_q^*}\varphi^n(c)$ equals $q-1$ if and only if φ^n is trivial. Otherwise, it is zero. On the other hand, φ^n is trivial if and only if its order divides d.For that reason, we must have $\varphi=\overline{\lambda^j}$ where $j\in\{0,1,...,d-1\}$.

Therefore

$$\begin{aligned}\sum_{c\in\mathbb{F}_q}\tau(c^n) &= 1+\sum_{j=0}^{d-1}G(\lambda^j,\tau)\\ &= \sum_{j=1}^{d-1}G(\lambda^j,\tau).\end{aligned} \tag{3.11}$$

Plugging Equation (3.13) into Equation (3.12) we conclude,

$$\begin{aligned}\sum_{c\in\mathbb{F}_q}\chi(ac^n+b) &= \chi(b)\sum_{c\in\mathbb{F}_q}\chi(ac^n)\\ &= \chi(b)\sum_{c\in\mathbb{F}_q}\tau(c^n)\\ &= \chi(b)\sum_{j=1}^{d-1}G(\lambda^j,\tau)\\ &= \chi(b)\sum_{j=1}^{d-1}\overline{\lambda}^j(a)G(\lambda^j,\chi).\end{aligned} \tag{3.12}$$

□

Theorem 3.1.9. *Let χ be a non-trivial additive character of $\mathbb{F}_q$ where q is a power of an odd prime, and let $f(x)=a_2x^2+a_1x+a_0\in\mathbb{F}_q[x]$ be a polynomial of degree 2. Then*

$$\sum_{c\in\mathbb{F}_q}\chi(f(c))=\chi(a_0-a_1^2(4a_2)^{-1})\eta(a_2)G(\eta,\chi)$$

where η is the quadratic character.

Proof. First, we reformulate $f(x)=a_2x^2+a_1x+a_0\in\mathbb{F}_q[x]$ as

$$f(x)=a_2(x+a_1(2a_2)^{-1})^2+a_0-a_1^2(4a_2)^{-1}.$$

This reformulation is a correct one, since

$$\begin{aligned}
&a_2(x + a_1(2a_2)^{-1})^2 + a_0 - a_1^2(4a_2)^{-1} \\
&= a_2(x^2 + a_1^2(4a_2^2)^{-1} + 2xa_1(2a_2)^{-1}) + a_0 - a_1^2(4a_2^{-1}) \\
&= a_2x^2 + a_1^2 4a_2^{-1} + 4xa_1 + a_0 - a_1(4a_2^{-1}) \\
&= a_2x^2 + a_1x + a_0 \\
&= f(x).
\end{aligned}$$

Therefore, relying on this formulation, we can write

$$f(c) = a_2(c + a_1(2a_2)^{-1})^2 + a_0 - a_1^2(4a_2)^{-1}.$$

Now setting

$$\alpha = c + a_1(2a_2)^{-1}$$

and

$$\beta = a_0 - a_1^2(4a_2)^{-1}$$

and using Theorem 3.1.8, we obtain

$$\begin{aligned}
\sum_{c\in\mathbb{F}_q} \chi(f(c)) &= \sum_{\alpha\in\mathbb{F}_q} \chi(a_2\alpha^2 + \beta) \\
&= \chi(\beta)\eta(a_2)G(\eta,\chi) \\
&= \chi(a_0 - a_1^2(4a_2)^{-1})\eta(a_2)G(\eta,\chi).
\end{aligned}$$

□

Definition 3.1.10. *Let f be a polynomial in $\mathbb{F}_q[x]$. Then f is called a permutation polynomial if the mapping $c \longrightarrow f(c)$ permutes in $\mathbb{F}_q$, i.e $f(x) = a$ has a unique solution in $\mathbb{F}_q$ for each $a \in \mathbb{F}_q$.*

Lemma 3.1.11. *[22] Let $L(x) \in \mathbb{F}_q[x]$ defined as*

$$L(x) = \sum_{i=0}^{n-1} a_i x^{p^i}.$$

Then for any $c \in \mathbb{F}_q$,

$$Tr(cL(x)) = Tr(xL^*(c))$$

where $L^*(x) = \sum_{i=0}^{n-1} a_i^{p^{n-i}} x^{p^{n-i}}$.

Proof. We first note, for all $a \in \mathbb{F}_q$,

$$\begin{aligned} Tr(a) &= \sum_{i=0}^{n-1} a^{p^i} \\ &= \left(\sum_{i=0}^{n-1} a^{p^i} \right)^p \\ &= \sum_{i=0}^{n-1} \left(a^{p^i} \right)^p \\ &= \sum_{i=0}^{n-1} (a^p)^{p^i} \\ &= Tr(a^p). \end{aligned}$$

We calculate,

$$\begin{aligned} Tr(cL(x)) &= Tr(c \sum_{i=0}^{n-1} a_i x^{p^i}) \\ &= Tr(\sum_{i=0}^{n-1} c a_i x^{p^i}) \\ &= \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} c a_i x^{p^i} \right)^{p^j} \\ &= \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} c^{p^j} a_i^{p^j} x^{p^{i+j}} \right). \end{aligned}$$

On the other hand,

$$\begin{aligned} Tr(xL^*(c)) &= Tr(x \sum_{i=0}^{n-1} a_i^{p^{n-i}} c^{p^{n-i}}) \\ &= Tr(\sum_{i=0}^{n-1} x a_i^{p^{n-i}} c^{p^{n-i}}) \\ &= \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} x^{p^j} a_i^{p^{n-i+j}} c^{p^{n-i+j}} \right) \\ &= \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} x^{p^j} a_i^{p^{j-i}} c^{p^{j-i}} \right) \\ &= \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1} x^{p^{i+j}} a_i^{p^j} c^{p^j} \right) \\ &= Tr(cL(x)). \end{aligned}$$

□

3.2 Some Constructions of Mutually Unbiased Bases

The following theorem describes the first construction mutually unbiased bases included in our study.

Theorem 3.2.1. *(Theorem 1 in [18]) Let $\mathbb{F}_q$ be a finite field with characteristic $p \geq 5$. For any $\alpha, \lambda \in \mathbb{F}_q$, set the vectors*

$$b_{\lambda,\alpha} = \frac{1}{\sqrt{q}}(\xi_p^{Tr((k+\alpha)^3+\lambda(k+\alpha))})$$

where $k \in \mathbb{F}_q$ and ξ_p is the p-th complex root of unity. Then, under this setting, the sets $B_\alpha = \{b_{\lambda,\alpha} | \lambda \in \mathbb{F}_q\}$ with the standard basis form $q+1$ distinct mutually unbiased bases of $\mathbb{C}^q$.

Proof. Let $b_{\lambda_1,\alpha}, b_{\lambda_2,\alpha}$ be two vectors from the set B_α. Then, their inner product is

$$\begin{aligned}
\langle b_{\lambda_1,\alpha}, b_{\lambda_2,\alpha} \rangle &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr((k+\alpha)^3+\lambda_2(k+\alpha)-(k+\alpha)^3-\lambda_1(k+\alpha))} \\
&= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr(\lambda_2(k+\alpha)-\lambda_1(k+\alpha))} \\
&= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr((\lambda_2-\lambda_1)(k+\alpha))}
\end{aligned}$$

If the two vectors are the same, i.e $\lambda_2 = \lambda_1$, then

$$\begin{aligned}
\langle b_{\lambda_1,\alpha}, b_{\lambda_1,\alpha} \rangle &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr((\lambda_1-\lambda_1)(k+\alpha))} \\
&= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr(0)} \\
&= \frac{1}{q} \sum_{k \in \mathbb{F}_q} 1 \\
&= \frac{1}{q} q \\
&= 1.
\end{aligned}$$

If the two vectors are distinct, i.e $\lambda_2 \neq \lambda_1$ then,

$$\begin{aligned}\langle b_{\lambda_1,\alpha}, b_{\lambda_2,\alpha}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr((\lambda_2-\lambda_1)(k+\alpha))} \\ &= \frac{1}{q}0 \\ &= 0\end{aligned} \tag{3.13}$$

where we used

$$\sum_{k\in\mathbb{F}_q} \xi_p^{Tr((\lambda_2-\lambda_1)(k+\alpha))} = 0$$

since k ranges through all elements in the finite field $\mathbb{F}_q$ and $\lambda_2 - \lambda_1 \neq 0$.

When the vectors are picked from two distinct sets, $b_{\lambda_1,\alpha_1} \in B_{\alpha_1}$, $b_{\lambda_2,\alpha_2} \in B_{\alpha_2}$, we have the inner product

$$\begin{aligned}\langle b_{\lambda_1,\alpha_1}, b_{\lambda_2,\alpha_2}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr((k+\alpha_2)^3+\lambda_2(k+\alpha_2)-(k+\alpha_1)^3-\lambda_1(k+\alpha_1))} \\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr(k^3+3k^2a_2+3ka_2^2+a_2^3+\lambda_2k+\lambda_2a_2-k^3-3k^2a_1-3ka_1^2-a_1^3-\lambda_1k-\lambda_1a_1)} \\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr((3a_2-3a_1)k^2+(3a_2^2-3a_1^2+\lambda_2-\lambda_1)k+a_2^3-a_1^3+\lambda_2a_2-\lambda_1a_1).}\end{aligned}$$

Setting

$$\begin{aligned}A_2 &= 3a_2 - 3a_1 \\ A_1 &= 3a_2^2 - 3a_1^2 + \lambda_2 - \lambda_1 \\ A_0 &= a_2^3 - a_1^3 + \lambda_2a_2 - \lambda_1a_1\end{aligned}$$

we have,

$$\begin{aligned}\langle b_{\lambda_1,\alpha_1}, b_{\lambda_2,\alpha_2}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr((3a_2-3a_1)k^2+(3a_2^2-3a_1^2+\lambda_2-\lambda_1)k+a_2^3-a_1^3+\lambda_2a_2-\lambda_1a_1)} \\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \xi_p^{Tr(A_2k^2+A_1k+A_0)} \\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q} \chi(A_2k^2+A_1k+A_0)\end{aligned}$$

where we define the nontrivial character $\chi(A_2k^2 + A_1k + A_0)$ as

$$\chi(A_2k^2 + A_1k + A_0) = \xi_p^{Tr(A_2k^2+A_1k+A_0)}.$$

By Theorem 3.1.9 we have,

$$\begin{aligned}\langle b_{\lambda_1,\alpha_1}, b_{\lambda_2,\alpha_2}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\chi(A_2k^2 + A_1k + A_0)\\ &= \frac{1}{q}\chi(A_0 - A_1^2(4(A_2)^{-1})\eta(A_2)G(\eta,\chi).\end{aligned}$$

Therefore, using the fact that both the characters ν and χ are non-trivial and Theorem 3.1.6,

$$\begin{aligned}|\langle b_{\lambda_1,\alpha_1}, b_{\lambda_2,\alpha_2}\rangle| &= |\frac{1}{q}\chi(A_0 - A_1^2(4(A_2)^{-1})\eta(A_2)G(\eta,\chi)|\\ &= \frac{1}{q}|\chi(A_0 - A_1^2(4(A_2)^{-1})| \times |\eta(A_2)| \times |G(\eta,\chi)|]\\ &= \frac{1}{q}\sqrt{q}\\ &= \frac{1}{\sqrt{q}}.\end{aligned}$$

For the inner product with the standard base, since the root of unity has norm 1 and each base set has a coefficient $\frac{1}{\sqrt{q}}$, the norm of the inner product is $\frac{1}{\sqrt{q}}$.

Therefore, the sets described in the statement of the thereom builds $q+1$ mutually unbiased bases.

□

The following theorem describes a way of building mutually unbiased bases even when $p = 3$. We note that $p = 3$ case was not covered in Theorem 3.2.1.

Theorem 3.2.2. *(Theorem 2 in [18]) Let $\mathbb{F}_q$ be a finite field with odd characteristic p. For any $a, b \in \mathbb{F}_q$, set the vectors*

$$v_{a,b} = \frac{1}{\sqrt{q}}(\xi_p^{Tr(ak^2+bk))})$$

where $k \in \mathbb{F}_q$ and ξ_p is the $p-th$ complex root of unity. Then, under this setting, the sets $B_a = \{v_{a,b}|\lambda \in \mathbb{F}_q\}$ with the standard basis form $q+1$ distinct mutually unbiased bases of $\mathbb{C}^q$.

Proof. For any two vectors, say $v_{a,b}$ and $v_{c,d}$, we have the inner product

$$\begin{aligned}\langle v_{c,d}, v_{a,b}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{Tr(ak^2+bk-ck^2-dk)}\\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{Tr((a-c)k^2+(b-d)k)}.\end{aligned}$$

When the two vectors are the same, i.e, $a = c$ and $b = d$, we would have

$$\begin{aligned}\langle v_{a,b}, v_{a,b}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{Tr((a-a)k^2+(b-b)k)}\\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{0}\\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}1\\ &= \frac{1}{q}q\\ &= 1.\end{aligned}$$

If the two vectors are picked from the same B_a set as distinct vectors, i.e. $a = c$ but $b \neq d$, then

$$\begin{aligned}\langle v_{a,d}, v_{a,b}\rangle &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{Tr((a-a)k^2+(b-d)k)}\\ &= \frac{1}{q}\sum_{k\in\mathbb{F}_q}\xi_p^{Tr((b-d)k)}\\ &= \frac{1}{q}0\\ &= 0.\end{aligned}$$

Here we have

$$\sum_{k\in\mathbb{F}_q}\xi_p^{Tr((b-d)k)} = 0$$

since k ranges through all the field $\mathbb{F}_q$.

For the remaining case, $a \neq c$ and $b \neq d$, we set

$$A_2 = a - c$$
$$A_1 = b - d$$
$$A_0 = 0$$

Then,

$$\begin{aligned}\langle v_{c,d}, v_{a,b} \rangle &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr((a-c)k^2+(b-d)k)} \\ &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr(A_2k^2+A_1k)}\end{aligned}$$

By Theorem 3.1.9,

$$\begin{aligned}\langle v_{c,d}, v_{a,b} \rangle &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \xi_p^{Tr(A_2k^2+A_1k)} \\ &= \frac{1}{q} \sum_{k \in \mathbb{F}_q} \chi(A_2k^2 + A_1k) \\ &= \frac{1}{q} \chi(-A_1^2(4(A_2)^{-1})\eta(A_2)G(\eta, \chi).\end{aligned}$$

Therefore, as the norm of the inner product,

$$\begin{aligned}|\langle v_{c,d}, v_{a,b} \rangle| &= |\frac{1}{q} \chi(-A_1^2(4(A_2)^{-1})\eta(A_2)G(\eta, \chi)| \\ &= \frac{1}{q} |\chi(-A_1^2(4(A_2)^{-1})||\eta(A_2)||G(\eta, \chi)| \\ &= \frac{1}{q} \sqrt{q} \\ &= \frac{1}{\sqrt{q}}.\end{aligned}$$

The inner product with the vectors in the standard basis would again be $\frac{1}{\sqrt{q}}$ since each component has unit length and the base sets have a coefficient term $\frac{1}{\sqrt{q}}$. Therefore, the sets B_a with the standard basis build up $q+1$ mutually unbiased bases for $\mathbb{C}^q$. $\square$

In the following theorem, a way of constructing mutually unbiased bases using bent functions are described. However, we need to introduce some terminology beforehand.

Given a p-ary function from $\mathbb{F}_q$ to $\mathbb{F}_p$, where $q = p^m$, the Walsh transform of $f(x)$ at a point $\beta \in \mathbb{F}_q$ is defined as

$$W_f(\beta) = \sum_{x \in \mathbb{F}_q} \xi_p^{f(x) - \mathrm{Tr}(\beta \mathrm{x})}$$

If $|W_f(\beta)| = p^{\frac{n}{2}}$ for all $\beta \in \mathbb{F}_q$, then $f(x)$ is called a bent function. Moreover, under the same setting, if there exists a complex number c of unit norm satisfying $W_f(\beta) = cp^{\frac{n}{2}} \xi_p^{f^*(\beta)}$ for some function $f^*(x)$, f is called a weakly regular function. For a weakly regular function $f(x)$, the function $f^*(x)$ is called the dual of $f(x)$.

Theorem 3.2.3. *(Theorem 3.1 in [22]) Let $\mathbb{F}_q$ be a finite field with characteristic p, i.e $q = p^n$ for some n. Moreover, let $D = \{(x, y) \in \mathbf{F}_q \times \mathbf{F}_p : f(x) + y = 0\}$ where f is a p-ary weakly regular bent function. Set*

$$v_{a,b} = \frac{1}{\sqrt{q}} (\xi_p^{Tr(ax+by)})_{(x,y) \in D} = \frac{1}{\sqrt{q}} (\chi(ax + by))_{(x,y) \in D} \in V_b$$

where ξ_p is the $p-th$ complex root of unity. Set $B = \mathbf{F}_q / \sim$ where $\sim$ is an equivalence relation on $\mathbb{F}_q$ defined by

$$(b_1 \sim b_2 \iff Tr(b_1 - b_2) = 0).$$

The set

$$\mathcal{A} = (\bigcup_{b \in B} V_b) \cup \{\epsilon_q\}$$

where ϵ_q is the standard basis, forms a set of $p + 1$ mutually unbiased bases of $\mathbb{C}^q$.

Proof. First, we would like to note that in this theorem we construct only $p+1$ bases, not $q+1$. We need to see MUB conditions are satisfied with respect to the Hermitian inner product. For any $b_1, b_2 \in B$ and $a_1, a_2 \in \mathbb{F}_q$ we have

$$\begin{aligned} \langle v_{a_2,b_2}, v_{a_1,b_1} \rangle &= \frac{1}{q} \sum_{(x,y) \in D} \xi_p^{Tr(a_1 - a_2)x + (b_1 - b_2)y} \\ &= \frac{1}{q} \sum_{x \in \mathbb{F}_q} \sum_{y \in \mathbb{F}_p} \frac{1}{p} \sum_{z \in \mathbb{F}_p} \xi_p^{z(f(x)+y) + Tr((a_1 - a_2)x + (b_1 - b_2)y)} \end{aligned} \tag{3.14}$$

In Equation (3.14), x, y values initially taken from the set $D \subset \mathbb{F}_q \times \mathbb{F}_p$ are extended to all of $\mathbb{F}_q \times \mathbb{F}_p$ in a clever way. Note that when the pair (x, y) is in D, we have

$$\frac{1}{p}\sum_{z\in\mathbb{F}_p}\xi_p^{z(f(x)+y))+Tr((a_1-a_2)x+(b_1-b_2)y)}=\frac{1}{p}\sum_{z\in\mathbb{F}_p}\xi_p^{Tr((a_1-a_2)x+(b_1-b_2)y)}$$

since $f(x)+y=0$. Note that the summation is repeated p times $z\in\mathbb{F}_p$, therefore the coefficient $\frac{1}{p}$ cancels out.

We now expect the equation to vanish when $(x,y)\in\mathbb{F}_q\times\mathbb{F}_p\setminus D$ since we already have established the Equation (3.16) when $(x,y)\in D$. Note that $f(x)+y\neq 0$ when $(x,y)\in\mathbb{F}_q\times\mathbb{F}_p\setminus D$.

Assuming $(x,y)\in\mathbb{F}_q\times\mathbb{F}_p\setminus D$,

$$\frac{1}{p}\sum_{z\in\mathbb{F}_p}\xi_p^{z(f(x)+y))+Tr((a_1-a_2)x+(b_1-b_2)y)}=\frac{1}{p}\xi_p^{Tr((a_1-a_2)x+(b_1-b_2)y)}\sum_{z\in\mathbb{F}_p}\xi_p^{z(f(x)+y)}=0$$

since

$$\sum_{z\in\mathbb{F}_p}\xi_p^{z(f(x)+y)}=0.$$

We continue with

$$\begin{aligned}\langle v_{a_2,b_2},v_{a_1,b_1}\rangle=&\frac{1}{q}\sum_{x\in\mathbb{F}_q}\sum_{y\in\mathbb{F}_p}\frac{1}{p}\sum_{z\in\mathbb{F}_p}\xi_p^{z(f(x)+y)+Tr((a_1-a_2)x+(b_1-b_2)y)}\\=&\frac{1}{pq}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{Tr((b_1-b_2)y)}\\&+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2))y)}\Big).\end{aligned}\tag{3.15}$$

Equation (3.15) states the formulation of $\langle v_{a_2,b_2},v_{a_1,b_1}\rangle$ in the form of two summands. The first summand,

$$\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{Tr((b_1-b_2)y)}$$

corresponds to the case $z=0$. The second one,

$$\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2))y)})$$

corresponds to the remaining case, $z\in\mathbb{F}_p^*$.

Therefore, we finally have

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle = \frac{1}{pq}\Big(&\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{Tr((b_1-b_2)y)}\\ &+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2))y)}\Big).\end{aligned} \tag{3.16}$$

Now, in order to check MUB conditions, we pick two vectors v_{a_1,b_1}, v_{a_2,b_2} from the same basis set, i.e. $b_1 = b_2$. With this condition,

$$\begin{aligned}\langle v_{a_2,b_1}, v_{a_1,b_1}\rangle &= \frac{1}{pq}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{Tr(0)}\\ &+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr(0)}\Big)\\ &= \frac{1}{pq}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}1+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}\xi_p^{zy}\Big)\\ &= \frac{1}{pq}\Big(p\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}\sum_{y\in\mathbb{F}_p}(\xi_p^z)^y\Big)\\ &= \frac{1}{pq}\Big(p\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}+\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q}\xi_p^{zf(x)+Tr((a_1-a_2)x)}.0\Big)\\ &= \frac{1}{pq}\Big(p\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\Big)\\ &= \frac{1}{q}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_2)x)}\Big)\end{aligned} \tag{3.17}$$

When the vectors v_{a_1,b_1} andv_{a_2,b_1} are the same vectors, i.e, $a_1 = a_2$,

$$\begin{aligned}\langle v_{a_1,b_1}, v_{a_1,b_1}\rangle &= \frac{1}{q}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr((a_1-a_1)x)}\Big)\\ &= \frac{1}{q}\Big(\sum_{x\in\mathbb{F}_q}\xi_p^{Tr(0)}\Big)\\ &= \frac{1}{q}\Big(\sum_{x\in\mathbb{F}_q}1\Big)\\ &= \frac{1}{q}(q)\\ &= 1.\end{aligned} \tag{3.18}$$

Otherwise, if the vectors v_{a_1,b_1} andv_{a_2,b_1} are distinct vectors, i.e, $a_1 \neq a_2$,

$$\begin{aligned}\langle v_{a_2,b_1}, v_{a_1,b_1}\rangle &= \frac{1}{q}(\sum_{x\in\mathbb{F}_q} \xi_p^{Tr((a_1-a_2)x)}) \\ &= \frac{1}{q}(\sum_{x\in\mathbb{F}_q} (\xi_p^{Tr((a_1-a_2)x)}) \\ &= \frac{1}{q}(0) \\ &= 0.\end{aligned} \tag{3.19}$$

If we happen to pick two basis vectors v_{a_1,b_1}, v_{a_2,b_2} from two distinct base sets, i.e. $b_1 \neq b_2$,

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle &= \frac{1}{pq}(\sum_{x\in\mathbb{F}_q} \xi_p^{Tr((a_1-a_2)x)} \sum_{y\in\mathbb{F}_p} \xi_p^{Tr((b_1-b_2)y)} \\ &+ \sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)+Tr((a_1-a_2)x)} \sum_{y\in\mathbb{F}_p} \xi_p^{zy+Tr((b_1-b_2)y)}) \\ &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)+Tr((a_1-a_2)x)} \sum_{y\in\mathbb{F}_p} \xi_p^{zy+Tr((b_1-b_2)y)}.\end{aligned} \tag{3.20}$$

In Equation (3.20), the first part of the equation vanishes since

$$\sum_{y\in\mathbb{F}_p} \xi_p^{Tr((b_1-b_2)y)} = \sum_{y\in\mathbb{F}_p} (\xi_p^{Tr(b_1-b_2)})^y = 0.$$

We continue as

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)+Tr((a_1-a_2)x)} \sum_{y\in\mathbb{F}_p} \xi_p^{zy+Tr((b_1-b_2)y)} \\ &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{y\in\mathbb{F}_p} \xi_p^{zy+Tr((b_1-b_2)y)} \sigma_z(W_f(z^{-1}(a_2-a_1))).\end{aligned} \tag{3.21}$$

Here, we note that

$$\begin{aligned}\sigma_z(W_f(z^{-1}(a_2-a_1))) &= \sigma_z(\sum_{x\in\mathbb{F}_q} \xi_p^{f(x)-Tr(z^{-1}(a_2-a_1)x)}) \\ &= \sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)-zTr(z^{-1}(a_2-a_1)x)} \\ &= \sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)-Tr(zz^{-1}(a_2-a_1)x)} \\ &= \sum_{x\in\mathbb{F}_q} \xi_p^{zf(x)+Tr((a_1-a_2)x)}\end{aligned}$$

and we use

- σ_z is an automorphism.

- $zTr(\xi) = Tr(z\xi)$.

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2))y)}\sigma_z(W_f(z^{-1}(a_2-a_1)))\\ &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2)y)}\sigma_z\big(\epsilon_f\sqrt{p^*}^n\xi_p^{f^*(z^{-1}(a_2-a_1))}\big).\end{aligned} \tag{3.22}$$

In equation (3.22), we use

$$W_f(\beta) = \epsilon_f\sqrt{p^*}^n\xi_p^{f^*(\beta)}$$

where p^* denotes $\eta(-1)p = (-1)^{\frac{p-1}{2}}p$ and $\epsilon_f \in \{-1, 1\}$ is the sign of the Walsh transform of f.

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle &= \frac{1}{pq}\sum_{z\in\mathbb{F}_p^*}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2)y)}\sigma_z\big(\epsilon_f\sqrt{p^*}^n\xi_p^{f^*(z^{-1}(a_2-a_1))}\big)\\ &= \frac{1}{pq}\epsilon_f\sqrt{p^*}^n\sum_{z\in\mathbb{F}_p^*}\overline{\eta}^n(z)\xi_p^{zf^*(z^{-1}(a_2-a_1))}\big)\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr((b_1-b_2)y)}.\end{aligned} \tag{3.23}$$

In Equation (3.23) we use

$$\sigma_a(\sqrt{p^*}^n) = \overline{\eta}^n(a)(\sqrt{p^*}^n)$$

where σ_a is the automorphism of the p-th cyclotomic field $\mathbb{Q}(\xi_p)$ determined by $\sigma_a(\xi_p) = \xi_p^a$.

Now assume that n is even. Then

$$\begin{aligned}\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle &= \frac{1}{pq}\epsilon_f\sqrt{p^*}^n\sum_{z\in\mathbb{F}_p^*}\overline{\eta}^n(z)\xi_p^{zf^*(z^{-1}(a_2-a_1))}\sum_{y\in\mathbb{F}_p}\xi_p^{zy+Tr(b_1-b_2)y}\\ &= \frac{1}{pq}\epsilon_f\sqrt{p^*}^n\sum_{z\in\mathbb{F}_p^*}\xi_p^{zf^*(z^{-1}(a_2-a_1))}\sum_{y\in\mathbb{F}_p}\xi_p^{(z+Tr(b_1-b_2))y}.\end{aligned} \tag{3.24}$$

In this expression, the sum

$$\sum_{y\in\mathbb{F}_p}\xi_p^{(z+Tr((b_1-b_2)y)}$$

does not vanish only when

$$z + Tr(b_1 - b_2) = 0.$$

Therefore, we have

$$\begin{aligned}
\langle v_{a_2,b_2}, v_{a_1,b_1} \rangle &= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \xi_p^{z f^*(z^{-1}(a_2 - a_1))} \sum_{y \in \mathbb{F}_p} \xi_p^{(z + Tr((b_1 - b_2))y)} \\
&= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \xi_p^{Tr(b_2 - b_1) f^*((Tr(b_2 - b_1))^{-1}(a_2 - a_1))} \sum_{y \in \mathbb{F}_p} 1 \\
&= \frac{1}{q} \epsilon_f \sqrt{p^*}^n \xi_p^{Tr(b_2 - b_1) f^*((Tr(b_2 - b_1))^{-1}(a_2 - a_1))}
\end{aligned} \tag{3.25}$$

We check

$$|\langle v_{a_2,b_2}, v_{a_1,b_1} \rangle| = |\frac{1}{q} \epsilon_f \sqrt{p^*}^n \xi_p^{Tr(b_2 - b_1) f^*((Tr(b_2 - b_1))^{-1}(a_2 - a_1))}| = \frac{1}{\sqrt{q}}.$$

If n is odd,

$$\begin{aligned}
\langle v_{a_2,b_2}, v_{a_1,b_1} \rangle &= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \overline{\eta}^n(z) \xi_p^{z f^*(z^{-1}(a_2 - a_1))}) \sum_{y \in \mathbb{F}_p} \xi_p^{zy + Tr((b_1 - b_2)y)} \\
&= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \overline{\eta}(z) \xi_p^{z f^*(z^{-1}(a_2 - a_1))}) \sum_{y \in \mathbb{F}_p} \xi_p^{zy + Tr((b_1 - b_2)y)}.
\end{aligned} \tag{3.26}$$

Again, the only case satisfying

$$\sum_{y \in \mathbb{F}_p} \xi_p^{zy + Tr((b_1 - b_2)y)} \neq 0$$

is when

$$z = Tr(b_2 - b_1).$$

Therefore

$$\begin{aligned}
\langle v_{a_2,b_2}, v_{a_1,b_1} \rangle &= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \overline{\eta}(z) \xi_p^{z f^*(z^{-1}(a_2 - a_1))}) \sum_{y \in \mathbb{F}_p} \xi_p^{zy + Tr((b_1 - b_2)y)} \\
&= \frac{1}{pq} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \overline{\eta}(Tr(b_2 - b_1)) \xi_p^{Tr(b_2 - b_1) f^*(Tr(b_2 - b_1))^{-1}(a_2 - a_1))}) \sum_{y \in \mathbb{F}_p} 1 \\
&= \frac{1}{q} \epsilon_f \sqrt{p^*}^n \sum_{z \in \mathbb{F}_p^*} \overline{\eta}(Tr(b_2 - b_1)) \xi_p^{Tr(b_2 - b_1) f^*(Tr(b_2 - b_1))^{-1}(a_2 - a_1))})
\end{aligned} \tag{3.27}$$

Again,

$$|\langle v_{a_2,b_2}, v_{a_1,b_1}\rangle| = |\frac{1}{q}\epsilon_f\sqrt{p^*}^n \sum_{z\in\mathbb{F}_p^*} \overline{\eta}(Tr(b_2-b_1))\xi_p^{Tr(b_2-b_1)f^*(Tr(b_2-b_1))^{-1}(a_2-a_1))}|= \frac{1}{\sqrt{q}}.$$

If one of the vectors is from the standard base $\varepsilon_q = \{e_1, ..., e_q\}$, for any $v_{a,b} = \{x_1 + iy_1, ..., x_q + iy_q\}$ we have,

$$|\langle v_{a,b}, e_j\rangle |= \frac{1}{\sqrt{q}}|\xi_p^{Tr(ax_j+by_j)}|= \frac{1}{\sqrt{q}} \tag{3.28}$$

□

In the following theorem, we require the construction of p being bigger than 3 since the construction includes polynomials of degree 3.

Theorem 3.2.4. *[22] Let p be a prime number greater than 3 and $q = p^n$. Let $L(x) = \sum_{i=1}^{n-1} a_i x^{p^i}$ be a linearized permutational polynomial and χ be a nontrivial additive character over $\mathbb{F}_q$. Construct a set of basis vectors as $V_a = \{v_{a,b} : b \in \mathbb{F}_q\}$ where*

$$v_{a,b} = \frac{1}{\sqrt{q}}\{\chi((x+a)^3 + bL(x))\}_{x\in\mathbb{F}_q}$$

and $a \in \mathbb{F}_q$. Then, taking $\varepsilon_q = \{e_1, ..., e_q\}$ as the standard basis for $\mathbb{C}^q$, the set

$$A = (\bigcup_{a\in\mathbb{F}_{11}} V_a) \cup \{\varepsilon_q\}$$

forms a complete set of mutually unbiased bases of $\mathbb{C}^q$.

Proof. First of all, we would like to note that, this theorem constructs a complete set of mutually unbiased bases of $\mathbb{C}^q$, which means $q + 1$, the maximum number of possible mutually unbiased bases is reached. For any $v_{a_1,b_1} \in V_{a_1}$, $v_{a_2,b_2} \in V_{a_2}$, we have

$$\begin{aligned}
&\langle v_{a_1,b_1}, v_{a_2,b_2}\rangle = \frac{1}{q}\sum_{x\in\mathbb{F}_q} \chi((x+a_1)^3 + b_1L(x) - (x+a_2)^3 - b_2L(x)) \\
&= \frac{1}{q}\sum_{x\in\mathbb{F}_q} \chi(x^3 + 3x^2a_1 + 3xa_1^2 + a_1^3 + b_1L(x) - x^3 - 3x^2a_2 - 3xa_2^2 - a_2^3 - b_2L(x)) \\
&= \frac{1}{q}\sum_{x\in\mathbb{F}_q} \chi(3x^2(a_1-a_2) + 3x(a_1^2 - a_2^2) + a_1^3 - a_2^3 + (b_1-b_2)L(x)).
\end{aligned} \tag{3.29}$$

If we pick two vectors v_{a_1,b_1}, v_{a_2,b_2} from the same basis set, i.e $a_1 = a_2$,

$$\begin{aligned}\langle v_{a_1,b_1}, v_{a_1,b_2}\rangle &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi(3x^2(a_1-a_1)+3x(a_1^2-a_1^2)+a_1^3-a_1^3+(b_1-b_2)L(x)) \\ &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi((b_1-b_2)L(x))\end{aligned} \tag{3.30}$$

If these two vectors v_{a_1,b_1}, v_{a_1,b_2} are the same, i.e $b_1 = b_2$ as well,

$$\begin{aligned}\langle v_{a_1,b_1}, v_{a_1,b_1}\rangle &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi((b_1-b_1)L(x)) \\ &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi(0) \\ &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}1 \\ &= 1.\end{aligned} \tag{3.31}$$

Otherwise, if v_{a_1,b_1}, v_{a_1,b_2} are distinct, i.e. $b_1 \neq b_2$, we have

$$\begin{aligned}\langle v_{a_1,b_1}, v_{a_1,b_2}\rangle &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi((b_1-b_2)L(x)) \\ &= 0.\end{aligned} \tag{3.32}$$

In Equation 3.32, the identity

$$\frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi((b_1-b_2)L(x)) = 0$$

is obtained by the fact that $L(x)$ is a permutation polynomial.

If the two vectors v_{a_1,b_1}, v_{a_1,b_2} are picked from two different basis sets, i.e. $a_1 \neq a_2$, then,

$$\begin{aligned}\langle v_{a_1,b_1}, v_{a_2,b_2}\rangle &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi(3x^2(a_1-a_2)+3x(a_1^2-a_2^2)+a_1^3-a_2^3+(b_1-b_2)L(x)) \\ &= \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi(3x^2(a_1-a_2)+3x(a_1^2-a_2^2)+a_1^3-a_2^3+xL^*(b_1-b_2)).\end{aligned} \tag{3.33}$$

In Equation 3.33, $(b_1 - b_2)L(x)$ is replaced by $xL^*(b_1 - b_2)$ as an application of Lemma 3.1.11.

We continue with

$$\langle v_{a_1,b_1}, v_{a_2,b_2}\rangle = \frac{1}{q}\sum_{x\in\mathbb{F}_q}\chi(3x^2(a_1-a_2)+3x(a_1^2-a_2^2)+a_1^3-a_2^3+xL^*(b_1-b_2)$$
$$= \frac{1}{q}\chi(a_1^3-a_2^3-(3(a_1^2-a_2^2)+L^*(b_1-b_2))^2(12(a_1-a_2)^{-1})\eta(3(a_1-a_2))G(\eta,\chi). \tag{3.34}$$

Since $|G(\eta,\chi)|=\sqrt{q}$, we have

$$|\langle v_{a_1,b_1}, v_{a_2,b_2}\rangle = \frac{1}{q}\sqrt{q} = \frac{1}{\sqrt{q}}.$$

If one of the vectors is chosen from the standard basis set ε_q, say e_j, then

$$|\langle v_{a,b}, e_j\rangle |= \frac{1}{\sqrt{q}}|\chi((a+x_j)^3+bL(x_j))|= \frac{1}{\sqrt{q}}1 = \frac{1}{\sqrt{q}}.$$

Therefore, A forms a complete set of mutually unbiased bases for $\mathbb{C}^q$. □

In order to describe another way of building mutually unbiased bases, we need to introduce the perfect non-linear (PN) functions and some of their basic properties with characters.

Definition 3.2.5. *For two abelian groups G_1 and G_2 and a function $f : G_1 \longrightarrow G_2$, the map $\Delta_{a:f}$, defined as*

$$\Delta_{a:f} : G_1 \longrightarrow G_2$$
$$\Delta_{a:f}(x) = f(a+x) - f(x)$$

is called the difference operator of f at the point $a \in G_1$.

Definition 3.2.6. *A function f as described in Definition 3.2.5 is called a perfectly non-linear function (PN-function) if its difference operator at each point is a permutation polynomial.*

Theorem 3.2.7. *[12] Given any finite field $\mathbb{F}_q$, let $f : \mathbb{F}_q \longrightarrow \mathbb{F}_q$ be a perfectly non-linear function. Then for any $a, \in \mathbb{F}_q^*$, $b \in \mathbb{F}_q$ and a non-trivial additive character χ of $\mathbb{F}_q$,*

$$\left| \sum_{x \in \mathbb{F}_q} \chi(af(x) + bx) \right| = \sqrt{q}$$

Proof.

$$\begin{aligned} \left| \sum_{x \in \mathbb{F}_q} \chi(af(x) + bx) \right|^2 &= \left(\sum_{x \in \mathbb{F}_q} \chi(af(x) + bx) \right) \left(\sum_{y \in \mathbb{F}_q} \chi(-af(y) - by) \right) \\ &= \sum_{x \in \mathbb{F}_q} \chi(af(x)) \sum_{y \in \mathbb{F}_q} \chi(-af(y) - b(y - x)). \end{aligned}$$

By making a change of parameters, as $z = x - y$, we state

$$\begin{aligned} \left| \sum_{x \in \mathbb{F}_q} \chi(af(x) + bx) \right|^2 &= \sum_{x \in \mathbb{F}_q} \chi(af(x)) \sum_{y \in \mathbb{F}_q} \chi(-af(y) - b(y - x)) \\ &= \sum_{x \in \mathbb{F}_q} \chi(af(x)) \sum_{z \in \mathbb{F}_q} \chi(-af(x - z) + bz) \\ &= \sum_{z \in \mathbb{F}_q} \chi(bz) \sum_{x \in \mathbb{F}_q} \chi(a(f(x) - f(x - z))). \end{aligned} \tag{3.35}$$

Since f is a perfectly non-linear function, we have

$$\sum_{x \in \mathbb{F}_q} \chi(a(f(x) - f(x - z))) = 0$$

when $z \neq 0$. However, when $z = 0$,

$$\begin{aligned} \sum_{x \in \mathbb{F}_q} \chi(a(f(x) - f(x - z))) &= \sum_{x \in \mathbb{F}_q} \chi(a(f(x) - f(x))) \\ &= \sum_{x \in \mathbb{F}_q} \chi(0) \\ &= \sum_{x \in \mathbb{F}_q} 1 \\ &= q. \end{aligned}$$

We have already realized that $\left| \sum_{x \in \mathbb{F}_q} \chi(af(x) + bx) \right|^2$ does not vanish only when $z = 0$, therefore, the iteration $\sum_{z \in \mathbb{F}_q}$ turns into $z = 0$ in our calculation.
Therefore,

$$\left|\sum_{x\in\mathbb{F}_q}\chi(af(x)+bx)\right|^2 = \chi(0)q \qquad (3.36)$$
$$= q.$$

Therefore,

$$\left|\sum_{x\in\mathbb{F}_q}\chi(af(x)+bx)\right| = \sqrt{q}. \qquad (3.37)$$

□

The following theorem describes a way of constructing mutually unbiased bases using perfectly non-linear functions on finite fields.

Theorem 3.2.8. *[22] Let χ be a nontrivial additive character of the finite field $\mathbb{F}_q$ with a characteristic $p \geq 5$. Set*

$$v_{a,b,c} = \frac{1}{q}\left(\chi(x+b)^3 + ax + cy + bf(y))\right)_{x,y\in\mathbb{F}_q}$$

included in $V_b = \{v_{a,b,c} : a, c \in \mathbb{F}_q\}$ where f is a perfectly non-linear function from $\mathbb{F}_q$ to $\mathbb{F}_q$.
Then the set

$$\mathcal{A} = \left(\cup_{b\in\mathbb{F}_q} V_b\right) \cup \{\epsilon_{q^2}\}$$

where $\{\epsilon_{q^2}\}$ is the standard basis on $\mathbb{C}^{q^2}$ forms a set of $q+1$ mutually unbiased bases for $\mathbb{C}^{q^2}$.

Proof. Under the setting described in the statement of the theorem, for the most gen-

eral case we have the inner product

$$
\begin{aligned}
&\langle v_{a_1,b_1,c_1}, v_{a_2,b_2,c_2}\rangle \\
&= \frac{1}{q^2}\Big(\big(\sum_{x,y\in\mathbb{F}_q} \chi((x+b_2)^3 + a_2x + c_2y + b_2f(y) \\
&- ((x+b_1)^3 + a_1x - c_1y - b_1f(y))\Big) \\
=&\frac{1}{q^2}\Big(\sum_{x,y\in\mathbb{F}_q} \chi((x+b_2)^3 - (x+b_1)^3 + (a_2-a_1)x \\
&+ (c_2-c_1)y + (b_2-b_1)f(y))\Big) \\
=&\frac{1}{q^2}\Big(\sum_{x,y\in\mathbb{F}_q} \chi((x^3+3x^2b_2+3xb_2^2+b_2^3) - (x^3+3x^2b_1+3xb_1^2+b_2^3) \\
&+ (a_2-a_1)x + (c_2-c_1)y + (b_2-b_1)f(y))\Big) \\
=&\frac{1}{q^2}\Big(\sum_{x,y\in\mathbb{F}_q} \chi(3(b_2-b_1)x^2 + (3(b_2^2-b_1^2)+a_2-a_1)x + b_2^3 - b_1^3 \\
&+ (c_2-c_1)y + (b_2-b_1)f(y))\Big) \\
=&\frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q} \chi(3(b_2-b_1)x^2 + (3(b_2^2-b_1^2)+a_2-a_1)x + b_2^3 - b_1^3) \\
&\sum_{y\in\mathbb{F}_q} \chi((c_2-c_1)y + (b_2-b_1)f(y)))\Big).
\end{aligned}
$$

When two vectors are selected from the same set V_b, i.e. $b_1 = b_2$, we have the inner product as,

$$
\begin{aligned}
\langle v_{a_1,b_1,c_1}, v_{a_2,b_1,c_2}\rangle =&\frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q} \chi(3(b_1-b_1)x^2 + (3(b_1^2-b_1^2)+a_2-a_1)x + b_1^3 - b_1^3) \\
&\sum_{y\in\mathbb{F}_q} \chi((c_2-c_1)y + (b_1-b_1)f(y)))\Big) \\
=&\frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q} \chi((a_2-a_1)x)\sum_{y\in\mathbb{F}_q} \chi((c_2-c_1)y)\Big).
\end{aligned}
\tag{3.38}
$$

In the Equation (3.38) above, if $a_1 \neq a_2$ or $c_1 \neq c_2$, , i.e. the vectors are not the same

$$
\begin{aligned}
\langle v_{a_1,b_1,c_1}, v_{a_2,b_1,c_2}\rangle &= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q} \chi((a_2-a_1)x)\sum_{y\in\mathbb{F}_q} \chi((c_2-c_1)y)\Big) \\
&= 0.
\end{aligned}
$$

However, when the vectors are the same, i.e $a_1 = a_2$ and $c_1 = c_2$ as well

$$\begin{aligned}
\langle v_{a_1,b_1,c_1}, v_{a_2,b_1,c_2}\rangle &= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}\chi((a_2-a_1)x)\sum_{y\in\mathbb{F}_q}\chi((c_1-c_2)y)\Big)\\
&= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}\chi((a_1-a_1)x)\sum_{y\in\mathbb{F}_q}\chi((c_1-c_1)y)\Big)\\
&= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}\chi(0)\sum_{y\in\mathbb{F}_q}\chi(0)\Big)\\
&= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}1\sum_{y\in\mathbb{F}_q}1\Big)\\
&= \frac{1}{q^2}q.q\\
&= 1.
\end{aligned}$$

If the two vectors v_{a_1,b_1,c_1} and v_{a_2,b_2,c_2} are taken from different V_b sets, in other words if $b_1 \neq b_2$:

$$\begin{aligned}
\langle v_{a_1,b_1,c_1}, v_{a_2,b_2,c_2}\rangle =&\frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}\chi(3(b_2-b_1)x^2+(3(b_2^2-b_1^2)+a_2-a_1)x+b_2^3-b_1^3)\\
&\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big).
\end{aligned}$$

Taking

$$\begin{aligned}
A_2 &= 3(b_2-b_1)\\
A_1 &= 3(b_2^2-b_1^2)+a_2-a_1\\
A_0 &= b_2^3-b_1^3
\end{aligned}$$

and by Theorem 3.1.9

$$\begin{aligned}
&\langle v_{a_1,b_1,c_1}, v_{a_2,b_2,c_2}\rangle\\
&= \frac{1}{q^2}\Big(\sum_{x\in\mathbb{F}_q}\chi(A_2x^2+A_1x+A_0)\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big)\\
&= \frac{1}{q^2}\Big(\chi(A_0-A_1^2(4A_2)^{-1})\eta(A_2)G(\eta,\chi)\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big)\\
&= \frac{1}{q^2}\Big(\chi((b_2^3-b_1^3)-(3(b_2^2-b_1^2)+a_2-a_1)^2(4(3(b_2-b_1)))^{-1})\eta(3(b_2-b_1))\\
&G(\eta,\chi)\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big).
\end{aligned}$$

Therefore,

$$
\begin{aligned}
&|\langle v_{a_1,b_1,c_1}, v_{a_2,b_2,c_2}\rangle| \\
=&\Big|\frac{1}{q^2}\Big(\chi((b_2^3-b_1^3)-(3(b_2^2-b_1^2)+a_2-a_1)^2(4(3(b_2-b_1)))^{-1}) \\
&\eta(3(b_2-b_1))G(\eta,\chi)\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big)\Big| \\
=&\frac{1}{q^2}\Big|\chi((b_2^3-b_1^3)-(3(b_2^2-b_1^2)+a_2-a_1)^2(4(3(b_2-b_1)))^{-1})\Big| \\
&\Big|\eta(3(b_2-b_1))\Big|\Big|G(\eta,\chi)\Big|\Big|\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big| \\
=&\frac{1}{q^2}1.1.\sqrt{q}\Big|\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big| \\
=&\frac{\sqrt{q}}{q^2}\Big|\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big|.
\end{aligned}
$$

Since f is assumed to be a perfectly non-linear function from $\mathbb{F}_q$ to $\mathbb{F}_q$, by Theorem 3.2.7, we have

$$\Big|\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big| = \sqrt{q}$$

which implies

$$
\begin{aligned}
|\langle v_{a_1,b_1,c_1}, v_{a_2,b_2,c_2}\rangle| &= \frac{\sqrt{q}}{q^2}\Big|\sum_{y\in\mathbb{F}_q}\chi((c_2-c_1)y+(b_2-b_1)f(y))\Big| \\
&= \frac{\sqrt{q}}{q^2}\sqrt{q} \\
&= \frac{q}{q^2} \\
&= \frac{1}{q}.
\end{aligned}
$$

□

3.3 Inequivalence of Mutually Unbiased Bases

In Section 3.2, we studied some ways of constructing mutually unbiased bases on complex space $\mathbb{C}^q$. It turns out, however, with a different perspective, it is possible to construct an equivalence relation between different mutually unbiased bases[17][8] .

Definition 3.3.1. *A family $F = \{l_i\}$ of 1-dimensional subspaces of $\mathbb{C}^q$, where $1 \leq i \leq q$, is called an orthoframe for $\mathbb{C}^q$ if the elements of F are pairwise orthogonal.*

Definition 3.3.2. *Let F_1 and F_2 be two distinct orthoframes for $\mathbb{C}^q$. Then F_1 and F_2 are called mutually unbiased if $|\langle u, v\rangle| = \frac{1}{\sqrt{q}}$ for two unit vectors $u \in F_1$ and $v \in F_2$.*

We note that one can move from mutually unbiased orthoframes to mutually unbiased bases by picking one unit vector from each mutually unbiased orthoframes.

Let $\mathbb{V} = \mathbb{Z}_p^n$ be the vector space in order to determine the indices on the complex space $\mathbb{C}^q$, where $q = p^n$. For any $a \in \mathbb{V}$ set

$$X_a(e_v) = e_{v+a} \tag{3.39}$$

and

$$Z_a(e_v) = (-\xi_p)^{a.v} e_v \tag{3.40}$$

where e_w denotes the standard basis vector of $\mathbb{C}^q$ corresponding to the index vector $w \in \mathbb{V}$.

Here, we note that the description of X and Z operators in Equations (3.39) and (3.40) are in accordance with the X and Z operators we already defined in Equations (2.46) and (2.48) for 1-qubit quantum systems, since

$$X_1 |0\rangle = |0+1\rangle = |1\rangle = X |0\rangle$$
$$X_1 |1\rangle = |1+1\rangle = |0\rangle = X |1\rangle$$

and

$$Z_1 |0\rangle = (-1)^{1.0} |0\rangle = |0\rangle = Z |0\rangle$$
$$Z_1 |1\rangle = (-1)^{1.1} |1\rangle = - |1\rangle = Z |1\rangle .$$

Using the operators X and Z, we define two new groups of order $q = p^n$ as

$$X(\mathbb{V}) = \{X(u) | u \in \mathbb{V}\}$$

and

$$Z(\mathbb{V}) = \{Z(u)|u \in \mathbb{V}\}.$$

Under this setting, we define another group in terms of the product

$$\Sigma = X(\mathbb{V})Z(\mathbb{V})\{\xi_p^k \mathbb{I}\big|0 \leq k \leq p-1\}.$$

The group Σ is called the extraspecial p-group. Its center is

$$Z(\Sigma) = \{\xi_p^k \mathbb{I}\big|0 \leq k \leq p-1\}.$$

Also note that

$$|\Sigma| = q.q.p = p^n.p^n.p = p^{2n+1}$$

We also have the natural homomorphism

$$\phi : \Sigma \longrightarrow \Sigma/Z(\Sigma) \tag{3.41}$$

where $\Sigma/Z(E) \cong \mathbb{V} \times \mathbb{V}$.

The quotient space $\Sigma/Z(E)$ is a symplectic space with respect to:

$$ab' - a'b := (X(a)Z(b))^{-1}(X(a')Z(b'))^{-1}(X(a)Z(b))(X(a')Z(b')). \tag{3.42}$$

Definition 3.3.3. *Taking ϕ and Σ as in Equation (3.41), a symplectic spread of $\phi(\Sigma)$ is a family $\{A_i\}$ of q+1 totally isotropic n-spaces of $\phi(\Sigma)$ where $A_i \cap A_j = \emptyset$ when $i \neq j$.*

With this setting we have the following theorem:

Theorem 3.3.4. *[8] For each symplectic spread Π of $\psi(P)$, there exists a complete set $\mathcal{F}(\Pi) = \{\mathcal{F}(\psi(A))\big|\psi(A) \in \Pi\}$ of q+1 mutually unbiased basis in $\mathbb{C}^q$ such that each $\mathcal{F}(\psi(A))$ is invariant under Σ.*

Moreover, if Π' is another symplectic spread of $\psi(P)$, then Π and Π' are said to be equivalent if and only if $\mathcal{F}(\Pi)$ and $\mathcal{F}(\Pi')$ are equivalent under a unitary transformation on $\mathbb{C}^q$.

CHAPTER 4

CONCLUSION

In this thesis, we studied mutually unbiased bases (MUBs) which have both potential and current applications for quantum communication protocols. We also showed some methods to construct mutually unbiased bases for Hilbert spaces using some auxiliary concepts like weakly-regular bent functions, linearized permutation polynomials and perfectly non-linear functions.

Since mutually unbiased bases are convenient for quantum communication protocols, we also introduced fundamental quantum mechanics and quantum computing principles. We studied the no-go theorems and explained the famous BB84 quantum key distribution protocol as a showcase where mutually unbiased bases are effectively employed.

REFERENCES

[1] W. Alltop, Complex sequences with low periodic correlations (corresp.), IEEE Transactions on Information Theory, 26(3), pp. 350–354, 1980.

[2] J. Bae and L.-C. Kwek, Quantum state discrimination and its applications, Journal of Physics A: Mathematical and Theoretical, 48(8), p. 083001, 2015.

[3] S. M. Barnett and S. Croke, Quantum state discrimination, Advances in Optics and Photonics, 1(2), pp. 238–278, 2009.

[4] M. Born, Quantenmechanik der stoßvorgänge, Zeitschrift für Physik, 38(11-12), pp. 803–827, 1926.

[5] G. Brassard and C. H. Bennett, Quantum cryptography: Public key distribution and coin tossing, in *International Conference on Computers, Systems and Signal Processing*, pp. 175–179, 1984.

[6] L. d. Broglie, XXXV. A tentative theory of light quanta, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 47(278), pp. 446–458, 1924.

[7] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*, Cambridge University Press, 2022.

[8] A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel, Z_4-kerdock codes, orthogonal spreads, and extremal euclidean line-sets, Proceedings of the London Mathematical Society, 75(2), pp. 436–480, 1997.

[9] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex fourier series, Mathematics of Computation, 19(90), pp. 297–301, 1965.

[10] A. Einstein, Über einem die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt, Annalen der physik, 4, 1905.

[11] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pp. 212–219, 1996.

[12] J. Hall, *Mutually unbiased bases and related structures*, Ph.D. thesis, RMIT University, 2011.

[13] W. Heisenberg, Quantum-theoretical re-interpretation of kinematic and mechanical relations, Z. Phys, 33, pp. 879–893, 1925.

[14] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics, 1, pp. 231–252, 1969.

[15] H. Hertz, Ueber sehr schnelle electrische Schwingungen, Annalen der Physik, 267(7), pp. 421–448, 1887.

[16] I. Ivonovic, Geometrical description of quantal state determination, Journal of Physics A: Mathematical and General, 14(12), p. 3241, 1981.

[17] W. Kantor, Mubs inequivalence and affine planes, Journal of mathematical physics, 53(3), 2012.

[18] A. Klappenecker and M. Rötteler, Constructions of mutually unbiased bases, in *Finite Fields and Applications: 7th International Conference, Fq7, Toulouse, France, May 5-9, 2003. Revised Papers*, pp. 137–144, Springer, 2004.

[19] P. Lenard, Über den elektrischen Bogen und die Spektren der Metalle, Annalen der Physik, 316(7), pp. 636–650, 1903.

[20] R. Lidl and H. Niederreiter, *Introduction to finite fields and their applications*, Cambridge University Press, 1994.

[21] M. Planck, On the law of distribution of energy in the normal spectrum, Annalen der Physik, 4(553), p. 1, 1901.

[22] L. Qian and X. Cao, Several new constructions of mutually unbiased bases derived from functions over finite fields, Quantum Information Processing, 21(8), p. 296, 2022.

[23] E. Schrödinger, An undulatory theory of the mechanics of atoms and molecules, Physical Review, 28(6), p. 1049, 1926.

[24] J. Schwinger, Unitary operator bases, Proceedings of the National Academy of Sciences, 46(4), pp. 570–579, 1960.

[25] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th Annual symposium on Foundations of Computer Science*, pp. 124–134, Ieee, 1994.

[26] M. Vázquez and A. Hanslmeier, *Ultraviolet radiation in the solar system*, volume 331, Springer Science & Business Media, 2005.

[27] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature, 299(5886), pp. 802–803, 1982.