Good Additive Cyclic Quantum Error-Correcting Codes

Ai Luo

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at

Concordia University

Montreal, Quebec, Canada

August 2004

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-94706-8
Our file Notre référence
ISBN: 0-612-94706-8

ABSTRACT

Good Additive Cyclic Quantum Error-Correcting Codes

Ai Luo

This thesis is about the quantum error-correcting codes. Although there are many methods to construct them, finding a good quantum code is still a complicated and difficult task, for we do not know which methods to use. To solve this problem, we did a thorough research on a special class of quantum codes – additive cyclic quantum codes. A new search algorithm has been designed and a lot of good additive cyclic quantum codes found by this algorithm are presented in this thesis. By showing the success of this algorithm and great value of additive cyclic quantum codes, we have greatly reduced the complexity of finding good quantum codes.

Quantum error-correction theory is key part of quantum information theory. As quantum information theory is a quite new field which has been developing fast during the last decade, we spent much time explaining the primary concepts of it. We introduced linear algebra, quantum mechanics, quantum operations, quantum error-correction theory and quantum error-correcting codes.

In a word, this thesis serves two functions: 1) an exploration of quantum error-correcting codes and 2) an introduction to quantum information theory.

Dedicated to my wife Lorna ……

ACKNOWLEDGEMENTS

Exploring a new field is an arduous task, and writing a thesis is also a long and difficult journey. Fortunately, I have a wonderful teacher and friend beside me, my supervisor Dr. M. R. Soleymani. He was always patient to help me to solve problems and gave me directions when I lost my way during the research. He never grudges giving me his time and encouragement. Without his help, I could not imagine how I could finish these two undertakings. Therefore, I am grateful to him for his kindness, patience and intelligence. I also thank him for giving me this precious opportunity to explore a new and interesting field.

There are many people who give my help indirectly. They are the authors of the papers I referred to in my thesis. Without their work and accomplishments, I could do nothing. Therefore, I also thank them.

TABLE OF CONTENTS

LISTS OF TABLES

LISTS OF FIGURES

LISTS OF ABBREVIATIONS AND SYMBOLS

GF	Galois Fields
CSS	Calderbank-Shor-Steane codes
$\vert \ \rangle$	Dirac notation
$\langle \ \vert \ \rangle$	Inner product
$\Vert \vert v\rangle \Vert$	The norm of a vector $\vert v\rangle$
$\vert \ \rangle\langle \ \vert$	Outer product
$\otimes$	Tensor product
A^T	Transpose operation of a matrix A
A^*	Conjugate operation of a matrix A
$A^\dagger$	Hermitian operation of a matrix A
$\oplus$	Modulo-2

Pauli matrices:

$$\sigma_0 \equiv I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \sigma_1 \equiv \sigma_X \equiv X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\sigma_2 \equiv \sigma_Y \equiv Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 \equiv \sigma_Z \equiv Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Hadamard operator: $H \equiv \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Controlled-Not gate: $U_C \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Chapter 1

Introduction

Quantum information theory is a new fast developing theory during the last decade. Despite the fact that not many practical applications of quantum information theory have been developed so far, it still appeals us intensely, for it is possible that the further study in this area may bring about important consequences to our society. The quantum information theory can enable us to develop applications which are impossible in classical world. The technologies of telecommunication and computers have already changed our lives greatly. If, one day, we could realize the quantum communications and quantum computers, then we would not be surprised to find that our lives would be changed radically again.

Quantum information theory is a rich subject. It involves many topics, such as quantum error-correction, data compression, network information theory, cryptography, and so on. However, in this thesis we will limit ourselves to quantum error-correction theory. There are many papers on how to find and construct good quantum error-correcting codes. Based on these accomplishments, we did a thorough research on a special kind of quantum error-correcting codes. A new search algorithm is presented and many good resulted quantum error-correcting codes are given in this thesis.

As quantum information theory is a new field and it is quite strange to most engineers, we try to draw a clear picture from the beginning so that it is easy for readers to follow. Therefore, this thesis serves two functions: introduction to a new field and our research results.

1.1 Development of quantum mechanics

The modern theory of quantum mechanics is a theory explaining the behaviors of the micro-particles, such as electrons and atoms. It provides us with a series of axioms for the world of micro-world. Since its creation in 1920s, the modern theory of quantum mechanics has been applied successfully in many areas, such as the structure of the atoms, nuclear fusion in stars, superconductors, etc.

Quantum mechanics is developed for a general use, not for some particular problems. It is like a skeleton and we need to add flesh and blood to make it fulfill a certain function. For example, there is an application of quantum mechanics called *quantum electrodynamics*, a theory explaining the interaction between the light and the atoms. In this theory, there are physical models and concepts which are not included in quantum mechanics, but they all obey the rules of quantum mechanics.

The development of applications of quantum mechanical systems has experienced two phases. The first phase was from 1920s to 1970s. During that time, scientists did much research on the control over the mass quantum systems, and we prefer to call the technologies during that period the "macro-way". A famous example is the superb quantum mechanical explanation for the *superconductivity,* which is a truly great discovery in the physical science during the course of the twenties century made by the Dutch physicist Heike Kammerlingh Onnes in 1911. Onnes found that the sensitivity of a sample of Mercury becomes zero beneath a certain critical temperature. This discovery brought about radical changes to our understanding of the electrical properties of materials, and Heike Kammerlingh Onnes received Nobel Prize in Physics 1913 for his great work on the properties of matter at low temperature. This remarkable phenomenon is a fundamentally quantum mechanics phenomenon which is observed on a macroscopic scale, and can only be explained within a quantum mechanic framework [1].

However, the “macro-way” provides us with a high level control over a huge number of quantum systems, leaving each single quantum mechanical system inaccessible. Because of this limitation, this way only allows us to probe a few aspects of quantum mechanical systems.

The second phase began in 1970s, when scientist shifted their interests to the control over single quantum systems, such as an atom, an electron, or a photon. We call the technologies during this period the “micro-way”. So far, many techniques of the “micro-way” have been developed. A typical example of this technique is the “atom trap”, which was first developed by Steven Chu, Claude Cohen-Tannoudji and William D. Phillips [2]. This technique enables us to isolate a single atom from others. It has great significance, for it opens the way of a deeper understanding of a single quantum system, and may lead to the appearance of more valuable applications, like the more precise atomic clocks for use in, e.g., space navigation and accurate determination of position. Because of this reason, the 1997 Nobel Prize in Physics was awarded jointly to Steven Chu, Claude Cohen-Tannoudji and William D. Phillips.

The techniques developed during the second phase are extremely important for us to explore quantum mechanics, for they enable us to understand and control this field from inside, not just observing from outside. In the history, great discoveries in science often accompany the advent of new methods for probing new regimes, and there is no doubt that as more powerful techniques of controlling single quantum systems are developed, we can make some remarkable discoveries in quantum world and even learn a great deal about the nature of the universe itself.

1.2 Development of quantum information theory

Quantum information theory is an application of quantum mechanics which is developed fast during last decade. This theory is attempting to solve almost the same problem which classical information theorem solves: transmitting information from one point to another point exactly or approximately. The only difference between the two theories is the carrier - in quantum information theory, we transmit information over quantum systems and we call whatever can be transmitted by using quantum systems *quantum information*. Therefore, it is not surprising that the primary questions of quantum information theory are as same as those of classical information theory: how much quantum information is carried by a given quantum system or a quantum channel, how can such information be encoded and decoded efficiently.

In classical information theory, such questions were solved by Claude Shannon [3]. His two famous theorems - noiseless channel coding theorem and noisy channel coding theorem - have laid the foundations for classical information theory. In the noiseless channel coding theorem, he solved the problem of how to measure the information produced from an information source. In the noisy channel coding theorem, he told us how much information can be transmitted reliably through a noisy channel. Meanwhile, he pointed out that error-correcting codes could be used to protect information against the noise in communication channels. Ever since the appearance of Shannon's theorems, there have been countless researchers bent on constructing good error-correcting codes. These efforts, at last, resulted in a sophisticated theory of error-correcting codes which is being widely used in every area of communication and computer science.

By comparison, quantum information theory is still in its developing phase. Even its basic questions have not been completely answered. In 1995, Ben Schumacher [4] developed a quantum theorem analogous to Shannon's noiseless channel coding theorem, and, in the process, he defined the "quantum bit" or *qubit* as the physical resource to transmit information. But the analogue quantum theorem to Shannon's noisy channel coding theorem has not yet been found. However, it does not prevent quantum error-correction theory from developing.

As we transmit information over quantum systems or store information in some sort of "quantum memories", the quantum systems carrying information will inevitably interact with their environment, for it is impossible to isolate our quantum systems absolutely from other quantum systems. Thus, the information will be lost to a less or greater degree. Similar problem happens in classical information theory, and we have already developed a sophisticated error-correction theory to circumvent such problems. But this theory can not be transferred to quantum regime. In classical information theory, we protect information against noise by adding redundancy, and a most simple method is the *repetition code*: instead of transmitting one bit we transmit several copies and decide during readout by a majority vote which bit we send. But quantum mechanics prevents the same procedure happening for quantum systems due to the not cloning theory discovered by W. K. Wooters and W. H. Zurek [5] in 1982, in which they pointed out that a single quantum system can not be cloned. What is more, in classical information theory, it is always possible to measure the value of a bit, but in quantum information theory, measuring an unknown quantum system will quite possibly collapse its state, a phenomenon we will explain later. Yet, despite these difficulties, it has been proved that quantum error-correction still can be done, of course not in classical way, but in quantum mechanics' own way. In 1995, P. W. Shor [6] found the first quantum error-correcting code. Since then, quantum error-correction theory has been developing rapidly.

The research shows that, despite the great differences between classical error-correction theory and quantum error-correction theory, these two fields are strongly connected. Indeed, classic coding theory has greatly improved the development and the understanding of the quantum error-correction. In 1996, two groups, Robert Calderbank and Peter Shor [7], and Andrew Steane [8], discovered an important class of quantum error-correcting codes independently. It is now known as *CSS* codes representing the initials of its inventors. A little later, a more powerful class of quantum codes, *stabilizer codes*, was discovered independently by Robert Calderbank, Peter Shor, Eric Rains, Neil Sloane [9], and Daniel Gottesman [10], and

CSS codes turned out to be a special case of stabilizer codes. Both CSS codes and stabilizer codes are strongly connected with classical error-correcting codes.

Quantum information theory is a new branch of information theory, for it involves something amazing called *entanglement* which only exists in quantum mechanics. Entanglement is the correlation between quantum systems. This correlation has very striking properties impossible in any classical world, and is ubiquitous in quantum mechanics. If a quantum system is entangled with the other quantum system, then each of these two quantum systems does not have its own quantum state, and only the two quantum systems together give a well-defined quantum state. What is more, the entanglement has nothing to do with local position. No matter how away these two quantum systems are separated, it always exists and does not wane. In quantum information theory, entanglement is an essential resource needed to perform otherwise impossible information processing or computation. A famous operation is *entanglement enhanced teleportation* which was first illustrated by Bennett et al [11]. Teleportation means the process of transmitting quantum information on a classical channel. It is found that teleportation is impossible only by purely classical means, but if we make use of a pair of entanglement quantum systems and put one of them at transmitter and the other one at receiver, then the classical channel can be upgraded to transmit quantum information. This entanglement enhanced teleportation is described in [11], and first experimental realization also exist [12]. There are also other famous applications of entanglement, such as *quantum cryptography* [13], and super-dense coding [14]. As the research is being carried out, scientists realize that entanglement is a fundamental resource in Nature, just as other fundamental resources like energy, entropy, information, etc. It is believed that the study of entanglement will give us more insights into the mysterious quantum world and lead us to new applications of quantum information theory.

1.3 Outline of the thesis

This thesis consists of seven chapters. In Chapter 2, we introduce the basic concepts of linear algebra and Galois Fields, which are necessary to explain quantum mechanics and quantum error-correction theory. In Chapter 3, we present the fundamental postulates and primary ideas of quantum mechanics. In Chapter 4, we introduce open-system dynamics of quantum systems, which is mainly about quantum channels and quantum noise. In Chapter 5, we give the theorem of quantum error-correction and an important class of quantum error-correcting code, *stabilizer* codes. In Chapter 6, we introduce an important conclusion which connects classical codes over $GF(4)$ with stabilizer codes, and we then go on to discuss our research based on this conclusion. Chapter 7 is the conclusion of this thesis.

Summary

The purpose of this chapter is to give readers a simple and general introduction to quantum mechanics and quantum information theory. Meanwhile, we outlined the structure of this thesis. As quantum information theory is a quite new field to most readers, in this thesis we will go to great lengths to explain it, and then present our research results: a new algorithm for searching a special kind of quantum error-correcting codes and the search results.

Due to the limited space, our descriptions of quantum mechanics and quantum information theory are very simple and short. In order to let readers get more information, we suggest following readings. For the history of quantum mechanics, we strongly recommend Pais's great work [15,16,17]; for the development of technologies based upon quantum mechanics, we recommend Milburn's work [18,19]; for classical information theory, we suggest MacWilliams and Sloane's great book on error-correction theory [20], Cover and Thomas's excellent text on information theory [21], the huge collection of classical information papers edited by Sloane and Wyner [22] and the collection edited by Slepian [23].

Chapter 2

Linear algebra and Galois fields

In this chapter, we will review necessary mathematical knowledge for the study of quantum mechanics and quantum error-correction theory: linear algebra and Galois fields. This chapter is very important, for the results and notations mentioned in this chapter will be frequently used in the whole thesis.

2.1 Linear algebra

Linear algebra is the main mathematical language for quantum mechanics. In this Chapter, we will introduce the necessary linear algebra knowledge. In order to make text easy to read and follow, we only list the conclusions and results, and omit the proofs, for they could be found in most text books and we do not need to repeat them here. We do the same way when we introduce Galois Fields.

2.1.1 Vector space

In quantum mechanics, we will limit ourselves to C^n vector space, which is the set of all column vectors $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, with entries z_i from the set of complex numbers C.

There are two operations defined on C^n: *addition* of two vectors, giving $z + z'$, and *scalar multiplication* of a vector z with a complex number λ, giving λz. They are defined as:

$$z + z' \equiv \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} + \begin{bmatrix} z'_1 \\ \vdots \\ z'_n \end{bmatrix} \equiv \begin{bmatrix} z_1 + z'_1 \\ \vdots \\ z_n + z'_n \end{bmatrix} \tag{2.1}$$

$$\lambda z \equiv \lambda \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \equiv \begin{bmatrix} \lambda z_1 \\ \vdots \\ \lambda z_n \end{bmatrix} \qquad (2.2)$$

The *zero vector* 0 is defined to be the vector with its entries being all zeros.

A subspace of C^n is defined to be a subset $W \subseteq C^n$ which contains the zero vector 0 and is closed under the operations of addition and scalar multiplication. That is, for any $x, y \in W$ and a complex number λ, $(x+y)$ and λx (and λy) must be in W.

In quantum mechanics, a vector z is usually represented by the notation $|\ \rangle$, called *Dirac notation,* $|z\rangle = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$.

2.1.2 Linear independence and bases

In vector space C^n, we say a set of vectors $\{|\Psi_1\rangle, |\Psi_2\rangle, \cdots, |\Psi_m\rangle\}$ is *linearly independent* if the equation $\sum_{i=1}^{m} a_i |\Psi_i\rangle = 0$ implies that $a_i = 0,\ i = 1, \ldots, m$

Otherwise, this set of vectors is *linearly dependent.* For example, the two vectors $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are linearly independent in C^2, while the vectors $u_1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ are linearly dependent for $u_1 - 2u_2 = 0$.

Any vector in a vector space can be represented as a linear combination of a certain set of vectors. Such a set of vectors is called a *spanning set* for the vector space. A *basis* of a vector space is a spanning set whose elements are linearly independent. In another word, a basis is one of the "smallest" spanning sets. For examples, the set of vectors $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\end{bmatrix}$ is a basis for C^2. The *dimension* of a vector space is the number of elements in the basis of this vector space. C^2 is a 2-dimensional vector space.

2.1.3 Matrices and linear operators

An $n \times m$ matrix is an array of numbers of the form:

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \tag{2.3}$$

Sometimes, this matrix will be denoted as $\left[a_{ij}\right]$, where a_{ij} is an arbitrary element in the matrix, and the i of a_{ij} represents the index of the row and the j represents the index of the column. If $n = m$, we call such a matrix a *square matrix.* If we say a matrix M is over C, we mean that the entries of M are all from C.

Two operations are defined for matrices: addition and multiplication. If and only if two matrices, $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$, are both $n \times m$ matrices, they can be added and the result $C = A + B$ is also an $n \times m$ matrices, with its entries $c_{ij} = a_{ij} + b_{ij}$.

There are two multiplications for matrices. The first one is the *scalar multiplication,* which is defined as: given a matrix $A = \left[a_{ij}\right]$, and a number $\lambda \in C$, we have $\lambda A = \left[\lambda a_{ij}\right]$. The second one is the multiplication between matrices. If and only if two matrices, $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$, have the size $n \times m$ and $m \times k$ respectively for some n, m, k, we have $C = AB = \left[c_{ij}\right]$, where $c_{ij} = \sum_{r=1}^{m} a_{ir} b_{rj}$. Note that the matrix multiplication is not in general commutative, which means that the existence of AB does not guarantee the existence of BA, and even if it exists, they are not likely to be equal, $AB \neq BA$. There are two special matrices, called *zero matrix* and *identity matrix*. The $n \times m$ zero matrix is a matrix with its entries being all zeros and is denoted as 0. The $n \times n$ identity matrix is a square matrix with diagonal entries being all 1 and all of the other being 0. It is denoted as I_n. The matrix multiplication obeys following laws, as long as A, B, C have the correct size:

$$(AB)C = A(BC) \tag{2.4}$$

$$A(B+C) = AB + AC \tag{2.5}$$

$$(A+B)C = AC + BC \tag{2.6}$$

$$0A = A0 = 0 \tag{2.7}$$

$$IA = AI = A \tag{2.8}$$

On its own, an $n \times m$ matrix A is just an array of numbers, but in the vector space, it represents a linear operator mapping a vector $|\psi\rangle$ of C^m to a vector $|\varphi\rangle$ of C^n. To see this, note that $|\psi\rangle$ is an $m \times 1$ vector, and the result of $A|\psi\rangle$ is an $n \times 1$ vector of C^n. Later, if we say that a linear operator A is defined on a vector space V, we mean that A is an $n \times n$ square matrix, where n is the dimension of V, for A maps a vector of V to another vector of V. The Equation (2.5) also shows that A is a linear operator, for $A\left(\sum \lambda_i |\psi_i\rangle\right) = \sum \lambda_i A|\psi_i\rangle$, where λ_i are complex numbers.

2.1.4 Pauli matrices

The four extremely important matrices in quantum mechanics are Pauli matrices, which are frequently used in quantum information theory. Here, we list these four matrices:

$$\sigma_0 \equiv I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \sigma_1 \equiv \sigma_X \equiv X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\sigma_2 \equiv \sigma_Y \equiv Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 \equiv \sigma_Z \equiv Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

2.1.5 Inner products and outer products

Inner product is a map taking two vectors to a scalar and in this thesis, we only care about the inner product over C. In what follows we write a^* for the complex conjugate of a complex number a.

Definition: *If* V *is a vector space over* C*, then the map* $\langle \, | \, \rangle$ *taking any two vectors in* V *to a complex number in* C *is an inner product if the following are true for all* $|u\rangle, |v\rangle, |w\rangle \in V$*, and* $\lambda, \mu \in C$*:*

1) (*Linearity*) $\langle u | \lambda v + \mu w \rangle = \lambda \langle u | v \rangle + \mu \langle u | w \rangle$;
2) (*Conjugate-symmetry*) $\langle u | v \rangle = \langle v | u \rangle^*$;
3) (*Positive definiteness*) $\langle v | v \rangle$ is a real number and $\langle v | v \rangle \geq 0$, the equality holds if and only if $|v\rangle = 0$.

A vector space equipped with an inner product function is called an *inner product space*. In the study of quantum mechanics, we will often use the term "*Hilbert space*". For finite dimensional case, a Hilbert space is exactly the same as an inner product space.

Let $V = C^n$, and let $|v\rangle = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $|u\rangle = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ be any two vectors in V, we define the map:

$$\langle v | u \rangle = \sum_{i=1}^{n} v_i^* u_i \tag{2.9}$$

It is easy to prove that (2.9) defines an inner product. It is called the *standard inner* *product* on C^n. This inner product is what we will always use in the thesis.

The *norm* of a vector $|v\rangle$ is defined as: $\||v\rangle\| \equiv \sqrt{\langle v|v\rangle}$, and we *normalize* a vector $|v\rangle$ by dividing $|v\rangle$ by its norm $|v\rangle / \||v\rangle\|$. If $\||v\rangle\| = 1$, we say $|v\rangle$ is a *unit vector* or a *normalized vector*. Two vectors are defined to be *orthogonal* if their inner product is zero. An *orthonormal basis* is defined to be such a basis that all of its elements are unit vectors and orthogonal to each other. In another word, the basis $\{|v_1\rangle, \cdots, |v_d\rangle\}$ is orthonormal if and only if $\langle v_i \mid v_j\rangle = \delta_{ij}$ where δ_{ij} is the *Dirac delta* function defined as:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2.10}$$

For example, the basis $\left\{|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ is an orthonormal basis for C^2, so is the basis $\{|+\rangle, |-\rangle\} = \left\{\frac{|0\rangle + |1\rangle}{\sqrt{2}}, \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right\}$.

The orthonormal basis is very useful for C^n, and it is necessary to mention an important means of transforming an arbitrary basis to an orthonormal basis. This procedure is called the *Gram-Schmidt algorithm*. Suppose that $\{|w_1\rangle, \cdots, |w_d\rangle\}$ is a basis for the space V. Define $|v_1\rangle = \frac{|w_1\rangle}{\||w_1\rangle\|}$. Then define $|v_{k+1}\rangle$, for $1 \leq k \leq d$, recursively by:

$$|v_{k+1}\rangle = \frac{|w_{k+1}\rangle - \sum_{i=1}^{i=k} \langle v_i | w_{k+1}\rangle |v_i\rangle}{\left\| |w_{k+1}\rangle - \sum_{i=1}^{i=k} \langle v_i | w_{k+1}\rangle |v_i\rangle \right\|} \tag{2.11}$$

It is easy to prove the basis $\{|v_i\rangle\}$ is an orthonormal basis.

Outer product is a useful and convenient way for expressing a linear operator. Together with inner product, it greatly simplifies our calculation notations. Suppose $|v\rangle = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ is a vector in space V, and $|w\rangle = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$ is a vector in space W, the outer product $|w\rangle\langle v|$ is defined to be a linear operator $A = |w\rangle\langle v| = \left[a_{ij}\right]$, where $a_{ij} = w_i v_j^*$. This linear operator maps the vectors in V to the vectors in W. Why? Let us see what happens when we apply this outer product to a vector $|v'\rangle$ in V:

$$(|w\rangle\langle v|)|v'\rangle \equiv |w\rangle\langle v|v'\rangle = \langle v|v'\rangle|w\rangle \qquad (2.12)$$

The right hand side of the Equation (2.12) is a vector in space W. Therefore, the outer product $|w\rangle\langle v|$ is a linear operator mapping vectors in V to vectors in W. If V is an n-dimensional space and W is an m-dimensional space, the outer product $|w\rangle\langle v|$ is an $m \times n$ matrix. It is easy to see that the linear combination of outer products is also an outer product.

With the help of the outer product, we can prove an important result, which is known as the *completeness relation* for the orthonormal basis. Let $\{|i\rangle\}$ be an orthonormal basis for the vector space C^n. A vector $|v\rangle$ in C^n can be expressed as: $|v\rangle = \sum_i v_i|i\rangle$, where $v_i = \langle i|v\rangle$. We apply the outer product $\sum_i |i\rangle\langle i|$ to this vector to get:

$$\left(\sum_i |i\rangle\langle i|\right)|v\rangle = \sum_i |i\rangle\langle i|v\rangle = \sum_i \langle i|v\rangle|i\rangle = \sum_i v_i|i\rangle = |v\rangle \qquad (2.13)$$

Because Equation (2.13) is true for any vector in C^n, we conclude that

$$\sum_i |i\rangle\langle i| = I \qquad (2.14)$$

The Equation (2.14) is known as the completeness relation. This equation is very useful in quantum mechanics.

2.1.6 Some important operators

In this section, we introduce several important operators that we will use frequently.

At first, we introduce the *inverse* of a matrix. If A is a square matrix, it may be possible that there is a square matrix B, which satisfies $AB = I$. B is called the inverse of A, and is denoted as $B = A^{-1}$. If the inverse of a square matrix exists, we say this matrix is *invertible*. A square matrix A and its inverse A^{-1} have the properties: $AA^{-1} = A^{-1}A = I$, or $\left(A^{-1}\right)^{-1} = A$, and this inverse is unique.

The *transpose* operation converts a matrix $A = \left[a_{ij}\right]$ to a matrix $B = \left[a_{ji}\right]$, and B is denoted as A^T. Thus, if A is an $n \times m$ matrix, A^T is an $m \times n$ matrix. For square matrices A, B, the transpose operation has the following properties:

$$(AB)^T = B^T A^T \tag{2.15}$$

$$\text{if } A \text{ is invertible, } \left(A^T\right)^{-1} = \left(A^{-1}\right)^T \tag{2.16}$$

$$\left(A^T\right)^T = A \tag{2.17}$$

Trace operation of a square matrix A, is defined as the sum of its diagonal elements: $Tr(A) = \sum_{i=1}^{n} a_{ii}$. For any two same size square matrices A and B, trace operation has the following important properties:

1) *cyclic property*: $Tr(AB) = Tr(BA)$,

2) *linear property*: $Tr(\lambda A + \mu B) = \lambda Tr(A) + \mu Tr(B)$, $\lambda, \mu \in C$

3) $Tr(A) = \sum_i \langle i|A|i\rangle$, $\{|i\rangle\}$ is an orthonormal basis.

4) $Tr(A|\psi\rangle\langle\psi|) = \langle\psi|A|\psi\rangle$, $|\psi\rangle$ is a vector

Hermitian operation on a matrix A, denoted as $A^\dagger$, is defined as: $A^\dagger = (A^*)^T$, where A^* means turning each entry of A to its conjugate. For example,

$$\begin{bmatrix} 1+3i & 2-i \\ 4+5i & 7 \end{bmatrix}^\dagger = \begin{bmatrix} 1-3i & 4-5i \\ 2+i & 7 \end{bmatrix} \tag{2.18}$$

Hermitian operation is also known as *adjoint* operation. For square matrices A, B, it has the following two properties:

$$(AB)^\dagger = B^\dagger A^\dagger \tag{2.19}$$

$$(A^\dagger)^\dagger = A \tag{2.20}$$

For the outer products like $|v\rangle\langle v|$, it is easy to prove that $(|v\rangle\langle v|)^\dagger = |v\rangle\langle v|$.

A linear operator A is called a *Hermitian* or *self-adjoint* operator if $A = A^\dagger$. An important class of the Hermitian operators is *projectors*. Suppose that W is a k-dimensional subspace of the d-dimensional space V. With Gram-Schmidt procedure, it is always possible to construct an orthonormal basis $\{|i\rangle\}$, whose first k elements: $|i\rangle, 1 \le i \le k$ is a basis for W. The projector P onto the subspace W is defined as:

$$P = \sum_{i=1}^{k} |i\rangle\langle i| \tag{2.21}$$

It is easy to check that P is a Hermitian operator. The projector is very important in quantum measurement and quantum error-correction. It functions as a 'filter'. To see this, let the vector $|v\rangle = \sum_i v_i |i\rangle$ be an arbitrary vector in V, and we apply the projector P to it:

$$P|v\rangle = \sum_{i=1}^{k} |i\rangle\langle i| \left[\sum_{j=1}^{k} v_j |j\rangle + \sum_{j=k+1}^{d} v_j |j\rangle \right] = \sum_{i=1}^{k} v_i |i\rangle \tag{2.22}$$

The part of the vector $|v\rangle$ that is not in the subspace W has been cut off, and only the part belonging to W remains. The vector $|v\rangle$ has been "filtered". Though P seems to consist of the orthonormal basis $\{|i\rangle\}$, it is only determined by subspace W, not by which orthonormal basis we use to span W. Later, we will use the phrase "vector space" P to represent the space whose projector is P. In this example, the vector space P is the subspace W. The *orthogonal complement* of P is defined as: $Q \equiv I - P$. Recalling the completeness relation (2.14), it is obvious that $Q = \sum_{i=k+1}^{d} |i\rangle\langle i|$.

Another important and special class of Hermitian operators is *positive* operators. An operator A is called a positive operator if the inner product $\langle v|A|v\rangle$ is a real and non-negative number for any non-zero vector $|v\rangle$. If the inner product $\langle v|A|v\rangle$ is always greater than 0 for any $|v\rangle \neq 0$, the operator A is said to be *positive definite.*

Hermitian operators are a subclass of a more general kind of operators, *normal operators*. A *normal operato*r A is an operator satisfying $AA^\dagger = A^\dagger A$. It is apparent that when $A = A^\dagger$, the equation holds.

There is another important subclass of normal operators, *unitary operators*. A square operator U is a *unitary operator* if it satisfies $UU^\dagger = I$. The point that makes unitary operator important is that it preserves the inner product of two vectors. To see this, we apply U to any two vectors $|v\rangle$ and $|w\rangle$ in V and get two new vectors $|v'\rangle = U|v\rangle, |w'\rangle = U|w\rangle$. The inner product of two new vectors is as same as the inner product of two original vectors:

$$\langle v'|w'\rangle = \langle v|U^{\dagger}U|w\rangle = \langle v|w\rangle \qquad (2.23)$$

Thus, unitary operators do not change the geometric relationship of a space. If two vectors in this space are orthogonal, after being applied by a unitary operator, these two vectors are still orthogonal. Conversely, if two spaces have exactly the same geometric relationship, there must be a unitary operator capable of mapping one space to the other. This conclusion is very useful and we will use it later.

2.1.7 Eigenvalues and eigenvectors

An *eigenvalue* of a square matrix A, is a non-zero complex value λ such that: $A|\psi\rangle = \lambda|\psi\rangle$, where $|\psi\rangle$ is a non-zero vector. The vector $|\psi\rangle$ is called an *eigenvector* of the linear operator A associated with the eigenvalue λ. Eigenvalues and eigenvectors are the basic characteristics of a linear operator. 'Eigen' is the German for 'characteristic of' or 'peculiar to', and sometimes eigenvalues and eigenvectors are also called the characteristic values and vectors.

The equation $c(\lambda) \equiv \det(A - \lambda I) = |A - \lambda I|$ is called the *characteristic equation,* where 'det' denotes the determinant of a matrix. According to the polynomial theory, any polynomial has at least one complex root. The roots of the characteristic equation are the eigenvalues of the linear operator A. For example, the characteristic equation of the matrix $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is:

$$c(\lambda) = |X - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1) \qquad (2.24)$$

The roots are -1 and 1. When $\lambda = 1$, from $X|\psi\rangle = |\psi\rangle$, we can find the corresponding eigenvector: $|\psi\rangle = a\begin{bmatrix} 1 \\ 1 \end{bmatrix} = a|0\rangle + a|1\rangle$, where a is any non-zero complex number. In the same way, we can find the eigenvector for the eigenvalue -1 as $|\psi\rangle = a\begin{bmatrix} 1 \\ -1 \end{bmatrix} = a|0\rangle - a|1\rangle$. The eigenvales of three Pauli matrices, X, Y, Z, are all 1 and -1.

An *eigenspace* of a linear operator A associated with an eigenvalue λ is the set of vectors: $\{|\psi\rangle : A|\psi\rangle = \lambda|\psi\rangle\} \cup \{0\}$. An eigenspace is a subspace of V. An eigenspace is *degenerate* when its dimension is greater than one. The eigenspace of the linear operator X with respective to 1 or -1 is not degenerate, for the dimension of each eigenspace is one. The eigenspace of a linear operator is the ensemble of the eigenspaces of all its eigenvalues.

A square linear operator A is *diagonalizable* if it has a *diagonal representation*:

$$A = \sum \lambda_i |i\rangle\langle i| \tag{2.25}$$

where $\{\lambda_i\}$ is the set of eigenvalues of A, and $\{|i\rangle\}$ is the set of corresponding orthonormal eigenvectors. For example, the operator X has the diagonal representation:

$$X = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| + \langle 1|}{\sqrt{2}}\right) - \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| - \langle 1|}{\sqrt{2}}\right) \tag{2.26}$$

Diagonal representations are sometimes also referred to as *orthonormal decompositions*. Not all the linear operators have diagonal representations. It has been proven that an operator is diagonalizable if and only if it is a normal operator. Therefore, all of the special types of operators we spoke of in the last section, i.e., Hermitian operators, positive operators, projectors, unitary operators, are diagonalizable.

2.1.8 Tensor product

Tensor product is very important in the study of multi-particle systems, for it merges small vector spaces into a larger space. Suppose *V* and *W* are two Hilbert spaces of dimension *n* and *m*, respectively. Then $V \otimes W$ is a Hilbert space of dimension *nm*. The notation $V \otimes W$ reads "*V* tensor *W*". If $|v\rangle$ is a vector of *V* and $|w\rangle$ is a vector of *W*, then $|v\rangle \otimes |w\rangle$ is a vector of $V \otimes W$. For convenience the notation $|v\rangle \otimes |w\rangle$ is sometimes represented as $|v, w\rangle$, $|vw\rangle$ or $|v\rangle|w\rangle$. The tensor product of two matrices A and B is denoted as $A \otimes B$, and is defined as:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & \cdots \\ a_{21}B & a_{22}B & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{2.27}$$

Tensor product satisfies the following basic properties:

1) For any scalar *z*, any element $|v\rangle$ of *V* and any element $|w\rangle$ of *W*:

$$z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle) \tag{2.28}$$

2) For any $|v_1\rangle$ and $|v_2\rangle$ in *V* and any $|w\rangle$ in *W*:

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle \tag{2.29}$$

3) For any $|v\rangle$ in *V* and any $|w_1\rangle$, $|w_2\rangle$ in *W*:

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle \tag{2.30}$$

4) Let A and B be two arbitrary linear operators defined on V and W, respectively. Let $|v_i\rangle$ and $|w_i\rangle$ be sets of vectors in *V* and *W*, respectively. Then,

$$(A \otimes B)\left(\sum a_i |v_i\rangle \otimes |w_i\rangle\right) = \sum a_i (A|v_i\rangle) \otimes (B|w_i\rangle) \tag{2.31}$$

The inner product in the space $V \otimes W$ can be defined by the inner products in the *V* and *W*. Suppose $\sum a_i |v_i w_i\rangle$ and $\sum b_j |v_j' w_j'\rangle$ are two vectors of $V \otimes W$, where $|v_i\rangle$, $|v_j'\rangle$ are elements of *V*, and $|w_i\rangle$, $|w_j'\rangle$ are elements of *W*. The inner product of these two vectors is defined as:

$$\left\langle \sum a_i \langle v_i w_i | \middle| \sum b_j | v_j' w_j' \rangle \right\rangle \equiv \sum_{i,j} a_i^* b_j \langle v_i | v_j' \rangle \langle w_i | w_j' \rangle \tag{2.32}$$

With this definition, we can solve the problem of finding a basis for the space $V \otimes W$ from the bases of V and W. Suppose that $\{|v_i\rangle\}$ is an orthonormal basis for V, and $\{|w_k\rangle\}$ is an orthonormal basis for W, then we can prove that $\{|v_i w_k\rangle\}$ is an orthonormal basis for $V \otimes W$. The inner product of any two vectors $|v_i w_m\rangle$ and $|v_j w_n\rangle$ is:

$$\langle v_i w_m | v_j w_n \rangle = \langle v_i | v_j \rangle \langle w_m | w_n \rangle = \delta_{i,j} \delta_{m,n} \tag{2.33}$$

The value is equal to 1 if and only if $i = j, n = m$, and for all other cases, the value is equal to zero. So $\{v_i w_k\rangle\}$ is an orthonormal basis for the space $V \otimes W$.

Suppose A and B are linear operators defined on spaces V and W, respectively. They have the following properties:

$$(A \otimes B)^* = A^* \otimes B^* \tag{2.34}$$

$$(A \otimes B)^T = A^T \otimes B^T \tag{2.35}$$

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger \tag{2.36}$$

At last, we mention a useful notation $a^{\otimes k}$. 'a' could be a linear operator or a state. This notation means that a linear operator or a state tensor product itself k times. For example, $A^{\otimes 3} = A \otimes A \otimes A$.

2.1.9 Commute and anti-commute

The commutator of two same size square matrices A and B is defined as:

$$[A, B] \equiv AB - BA \tag{2.37}$$

If $[A, B]=0$, we say A *commutes* with B. The anti-communtator of two matrices of A and B is defined as:

$$\{A, B\} \equiv AB+BA \tag{2.38}$$

If $\{A, B\}=0$, we say A *anti-commutes* with B.

Commutators and anti-commutators play an important role in quantum error-correction theory which we will introduce later. Here, we list the commutators and anti-commutators of Pauli matrices:

$$[X, Y]=2iZ; \quad [Y, Z]=2iX; \quad [Z, X]=2iY$$

$\{\sigma_i, \sigma_j\}=0$, where $i \neq j$ are both chosen from the set *1, 2, 3*.

The most important result of the commutators may be the *simultaneous diagonalization theorem.*

Theorem 2.1: (Simultaneous diagonalization theorem): Consider two Hermitian operators A and B. They satisfy $[A, B]=0$ if and only if there exists an orthonormal basis with respect to which both operators are diagonal. In such a case, we say that A and B are simultaneously diagonalizable.

For example, if the Hermitian operator A has the diagonal representation: $A=\sum_i a_i |i\rangle\langle i|$, and the Hermitian operator B has the diagonal representation: $B=\sum_i b_i |i\rangle\langle i|$ with respect to the same orthonormal basis $|i\rangle$, then A and B are simultaneously diagonal, and A must commute with B.

2.1.10 Singular value decomposition

In this section, we introduce a useful decomposition of square matrices.

Theorem 2.2: (Singular value decomposition) For any square matrix A, there are unitary matrices U and V, and a diagonal matrix D whose diagonal elements are all non-negative, such that $A = UDV$. The diagonal elements of D are called the *singular values* of A.

2.2 Galois fields

Group theory and Galois fields not only are the foundation for classical error-correction theory, but are proved great value in quantum error-correction theory as well. In this section, we introduce the primary definitions and the most important conclusions which we will use in later chapters.

2.2.1 Basic concepts of groups

A *group* is defined as a set of elements, G, and an operation $*$, satisfying:

1) G is closed under $*$, that is, for any $g,h \in G$, we have $g*h \in G$;

2) $*$ is associative: $(a*b)*c = a*(b*c)$;

3) There exists an identity element, e, such that for any $g \in G$, $g*e = e*g = g$;

4) For any $g \in G$, there exists another element g' such that $g*g' = g'*g = e$.

We call g' the inverse of g.

If a group G also satisfies: for any $g,h \in G$, $g*h = h*g$, then G is called a *commutative group* or *abelian group*. The groups we will refer to in this thesis are all abelian groups.

For example, $G = (Z,+)$ is a group, where Z is the set of integers. The identity element is 0 and the inverse of an arbitrary integer a is $-a$.

The groups with finite elements are called *finite groups*. The number of the elements of a finite group G is called the *order* of this group, and is denoted as $|G|$. For example, the group $G = (\{0,1\},\oplus)$, where $\oplus$ is the modulo 2 operation, has the order $|G| = 2$.

The *subgroup* H of a group G is a group which is contained in G and has the same operation as G. Associated with the subgroup H, there is an important concept called *Coset*. A Coset of H in G is formed by the following process. We take any $g \in G$, and form the set $gH = \{gh_i : h_i \in H\}$. This set is a Coset of H in G. Clearly, there may be many Cosets of H. If the element g belongs to H, the Coset gH is exactly H itself. We denote this Coset of H as H_0, and for other Cosets as H_i, where $i \geq 1$. For a finite group, its subgroups have finite Cosets, and from now on, what we talk about is all based on finite groups. There are several important results about a subgroup H and its Cosets H_i.

1) $|H_0| = |H_1| = \cdots$;

2) The elements of two Cosets are either totally identical or totally different. For example, if there is an element appearing in both Cosets, then these two Cosets are identical.

3) $|G|$ is a multiple of $|H_0|$;

4) H and its Cosets H_i partition G: $G = \bigcup_{i=0}^{N} H_0$, where $N = {|G|}/{|H_0|} - 1$.

The forth result is called the *Coset decomposition* of G induced by H. This result is very useful and will be used later.

2.2.2 Basic concepts of Galois field

A *field* F is defined to be a set of elements, together with two operators, addition and multiplication, satisfying the following properties:

1) $(F, +)$ is an abelian group and the identity element in this group is denoted as 0, and the inverse of an element α in this group is denoted as $-\alpha$;

2) $(F-\{0\},\cdot)$ is an abelian group and the identity element is called *unit* and denoted as 1, and the inverse of an element α in this group is denoted as α^{-1};

3) Distributive laws for any α,β,γ in the field: $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$;

4) For any element α in the field, $0\alpha=0$.

A *finite* field is defined to be field containing a finite number of elements, and this number is called the *order* of this field. Finite fields are called *Galois fields.* If the order of a Galois field is q, we denote this field as $GF(q)$. $GF(q)$ has many interesting properties, most of which can be derived from the definition of finite field. But we do not prove these properties, just list the conclusions directly.

For a $GF(q)$, q must be a prime number or a power of a prime number. We denote $q=p^m$, where p is a prime number and $m\geq 1$. p is called the *characteristic* of a field. The two operations in $GF(p^m)$ are modulo p addition and modulo p multiplication. The simplest finite field is $GF(p)$, the elements of which are $0,1,\cdots,p-1$. For example, $GF(2)=\{0,1\}$.

In $GF(p^m)$, there are m elements which are linear independent over $GF(p)$, $\beta_0,\beta_1,\cdots,\beta_{m-1}$ such that any element α in $GF(p^m)$ can be expressed as a linear combination of these m elements:

$$\alpha=a_0\beta_0+\cdots a_{m-1}\beta_{m-1} \tag{2.39}$$

where $a_i\in GF(p)$. Thus, $GF(p^m)$ is an m-dimensional space over $GF(p)$.

For any non-zero element α in $GF(p^m)$, consider the sequence $1,\alpha,\alpha^2,\cdots$, because $GF(p^m)$ is a finite field, there must be a number j such that $\alpha^j=1$, this number is called the *order* of the element α in $GF\left(p^m\right)$. The subgroup $1,\alpha,\alpha^2,\cdots,\alpha^{j-1}$ is called *cyclic multiplicative subgroup.* The element α is called the *generator* of this cyclic multiplicative subgroup. Any non-zero element in $GF\left(p^m\right)$ has an order.

Every element α in $GF\left(p^m\right)$ satisfies

$$\alpha^{p^m}=\alpha \qquad (2.40)$$

or

$$x^{p^m}-x=\prod_{\alpha\in GF\left(p^m\right)}(x-\alpha) \qquad (2.41)$$

A little change to the equation (2.41) draws the conclusion that all of the non-zero elements are roots of the equation

$$x^{p^m-1}-1=\prod_{\alpha\in GF\left(p^m\right),\alpha\neq 0}(x-\alpha) \qquad (2.42)$$

In $GF\left(p^m\right)$, there exists a *primitive element* whose order is p^m-1. In another word, this primitive element generates all of the non-zero elements in $GF\left(p^m\right)$.

If a field $GF\left(p^m\right)$ is contained in a field $GF\left(p^n\right)$, then $GF\left(p^m\right)$ is called a *subfield* of $GF\left(p^n\right)$, and $GF\left(p^n\right)$ is called an *extension* field of $GF\left(p^m\right)$.

2.2.3 Galois Field $GF(4)$

$GF(4)$ is one of the most important extension fields of $GF(2)$ and it is already widely used in classical communications. $GF(4)$ can be represented as $GF\left(2^2\right)$, so the characteristic is 2 and it is a 2-dimensional space over $GF(2)$. Traditionally,

$GF(4)$ is denoted as $\{0,1,w,w^2\}$, where w is the primitive element and its order is 3, $w^3=1$. The *conjugation* of an arbitrary element $\alpha \in GF(4)$ is defined as $\bar{\alpha} \equiv \alpha^2$. Thus, we have $\overline{0}=0$, $\overline{1}=1$, $\overline{w}=w^2$, $\overline{w^2}=w$. The elements $\{w,\overline{w}\}$ are linear independent, and all of the elements in $GF(4)$ can be expressed as:

$$0=0\cdot w+0\cdot\overline{w},\ \ 1=1\cdot w+1\cdot\overline{w},\ \ w=1\cdot w+0\cdot\overline{w},\ \ \overline{w}=0\cdot w+1\cdot\overline{w}$$

Summary

In this chapter, we introduced briefly the necessary mathematics knowledge we need for this thesis: linear algebra and Galois fields. Linear algebra is the main mathematic language for quantum mechanics, and Galois fields play an important role for connecting quantum error-correcting codes to classical error-correcting codes. As the contents in this chapter can be found in most linear algebra and Galois fields text books, we only listed conclusions. To find more details about these two fields, we suggest following readings. For linear algebra, there are many good books, and we suggest some of them: Horn and Johnson's two volumes [24,25], Halmos's book [26] and Strang's book [27]; for the Galois Fields, we suggest MacWilliams and Sloane's great book on error-correction theory [20] and Gallager's book on information theory [28].

Chapter 3

Fundamental of quantum mechanics

Quantum mechanics is the fundamental of quantum information theory. In this Chapter, we will review the carrier for quantum information – qubits, and some basic postulates and theorems of quantum mechanics. Though some principles seem counter-intuitive, they are simple themselves. They are rules in a totally different world from our classical world, and we just need to accept them as axioms, do not ask "why". Once we learn to think and understand what happens in quantum mechanics world in its own way, everything will turn out to be beautiful and manageable.

3.1 Quantum bits

Quantum bits [4], or *qubits*, are the mathematical models for quantum systems which are used to store quantum information. A qubit is the smallest unit to carry quantum information, just as a classical bit. However, unlike the bit, which has only two states, 1 and 0, the state of the qubit is a normalized vector over C^2. We define an orthonormal basis for C^2

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{3.1}$$

Thus, any state could be represented as:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{3.2}$$

where α, β are complex numbers and satisfy $|\alpha^2| + |\beta^2| = 1$. The state $|\psi\rangle$ is called *superposition*, and the special states $|0\rangle$ and $|1\rangle$ are called the *computational basis* *states*.

Why the state of a qubit must be a normalized vector? And why the state $|\psi\rangle$ is called superposition? The answers lie in the measurement of quantum mechanics. In classical information theory, we never have any difficulties in measuring the state of a bit and no one will consider the measurement as part of classical information theory. If a measurement gives us a result 0, the bit we are measuring is in the state 0. We can always trust the results. Therefore, the measurements act the role of an apparent interface between the states of bits and our observations. But things change in quantum mechanics. Not only can not we trust the measurement results, but measurements themselves are essential part of quantum information theory. It is one of the fundamental postulates in quantum mechanics. We will introduce its definition later in this chapter, and here we only illustrate the phenomenon.

The meaning of α, β is this: when we measure the state $|\psi\rangle$ on computational basis $\{|0\rangle, |1\rangle\}$, we may obtain either a measurement result corresponding to the state $|0\rangle$ with probability $|\alpha|^2$ or a measurement result corresponding to the state $|1\rangle$ with probability $|\beta|^2$. As the sum of probabilities should be equal to 1, we get the equation $|\alpha|^2 + |\beta|^2 = 1$. This is why the state of a qubit must be a normalized vector. The sign "+" in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ means "or": the state $|\psi\rangle$ is in the state $|0\rangle$ or in the state $|1\rangle$. This is why we call $|\psi\rangle$ superposition. For example, there is a state like:

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \tag{3.3}$$

After we measure this state, with probability of 0.5, it will collapse to the state $|0\rangle$ and give us a measurement result corresponding to this state, or with probability of 0.5, it will collapse to the state $|1\rangle$ and give us a corresponding result. This state, denoted as $|+\rangle$, together with the state $|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$, forms another frequently used orthonormal basis $\{|+\rangle, |-\rangle\}$, besides $\{|0\rangle, |1\rangle\}$.

Now, let's go back to the classical information theory. When we measure the state of a bit, we can get 1 or 0, and no matter what we get, the result is exactly what the state is. There is a direct connection between the measured result and the information itself. But for qubits, this connection disappears. Then, how do we extract the information from this carrier? Fortunately, ways have been found to measure the qubit states according to their properties. Therefore, despite the unavailability of direct access to qubits' states, they have the real, experimentally verifiable consequences, which are essential for the power of the quantum information theory and quantum computation.

The coefficients of the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ can be any value as long as they satisfy the equation $|\alpha|^2 + |\beta|^2 = 1$. This means that a qubit can be in any of infinite possible states. Therefore, in principle, we can store infinite amount of classical information in a qubit. However, the measurement only gives us two possible results - there is no way to gain the access to the infinite information that seems to be there, so it is incorrect to think that we can use one qubit to transmit a infinite amount of information. Then, how much information on earth can be stored in a qubit? Jozsa and Schumacher [29] and Schumacher [4] gave the answer: one two-state system's worth!

It seems that Nature is playing a game with us. She hides the most precious treasure and set the most difficult obstacles. She wants to test human's wisdom, diligence and patience. Maybe one day, our performance makes her satisfied, and she would willingly tell us how to get that treasure.

Now, we show several quantum systems which may serve as a qubit. In classical information theory, we can use many physical parameters to realize the states of a bit. For example, we can define a voltage of +5 volts to represent a "1", and a voltage of -5 volts to represent a "0". In quantum information theory, the frequently used quantum systems to represent a qubit are:

1) The ground and excited states of ions stored in a linear ion trap, with interactions between ions provided through a joint vibrational mode [30, 31].
2) Photons in either polarization, with interactions via cavity QED [32].
3) Nuclear spin states in polymers, with interactions provided by nuclear magnetic resonance techniques [33].

In 1), the 'ground' state or the 'excited' state corresponds to $|0\rangle$ or $|1\rangle$, respectively. Radiating the atom with a ray of light with an appropriate frequency, we can force the electron to 'jump' from the state $|0\rangle$ to the state $|1\rangle$. Quite interestingly, by reducing the duration of the radiation, the electron in the state $|0\rangle$ could move to a middle state between $|0\rangle$ and $|1\rangle$, which is $|+\rangle$.

One useful geometric model for a qubit is the following. As $|\alpha|^2 + |\beta|^2 = 1$, we can write the state this way:

$$|\psi\rangle = e^{i\gamma}(\cos(\frac{\theta}{2})|0\rangle + e^{i\varphi}\sin(\frac{\theta}{2})|1\rangle) \qquad (3.4)$$

Because the factor e^{ir} has no observable effect, we omit it.

$$|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\varphi}\sin(\frac{\theta}{2})|1\rangle \qquad (3.5)$$

The parameters θ and φ define a point on a three-dimensional sphere, which is called Bloch Sphere [34], as shown in Figure 3.1. This model provides us with a convenient visualization tool. But it is not useful for a general quantum mechanical system. It is only useful for a single qubit.

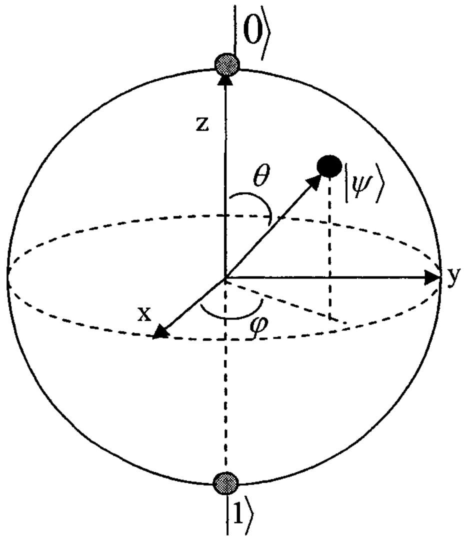

Figure 3.1 Bloch Sphere

3.2 Four fundamental postulates of quantum mechanics

Four basic postulates support the huge mansion of quantum mechanics. They describe essential issues of quantum mechanics in the mathematical forms and, therefore, abstract a physical world to a mathematical world. These issues are: what objects we are dealing with, how they change, how to measure them and how to build large systems from small ones. The descriptions of the four postulates are basically from [34].

State spaces

Postulate 1: Any isolated physical system is associated to a complex vector space with inner product (that is, a Hilbert space), which is known as the state space of the system. The system is completely described by its system state, a unit vector in the state space.

Postulate 1 is the foundation for the rest three postulates, for it gives us a mathematical definition of quantum systems. Though what it says is simple, its significance is great. A qubit is such a system: its state space is C^2, and its system state is a unit vector $|\varphi\rangle = a|0\rangle + b|1\rangle$, where $|a|^2 + |b|^2 = 1$. "a" is sometimes referred to as the *amplitude* for the state $|0\rangle$ and "b" is referred to as the *amplitude* for the state $|1\rangle$.

Evolution of closed quantum systems

Postulate 2: The evolution of a closed quantum system is described by unitary operations. The system state $|\varphi_1\rangle$ at time t_1 and the system state $|\varphi_2\rangle$ at time t_2 are connected by a unitary operator *U*, which only depends on the times t_1 and t_2:

$$U|\varphi_1\rangle = |\varphi_2\rangle \tag{3.6}$$

The Equation (3.6) stems from Schrodinger's famous equation [35] for physical systems:

$$i\hbar\frac{d|\varphi\rangle}{dt} = H|\varphi\rangle \tag{3.7}$$

where H is a fixed Hermitian operator known as the *Hamiltonian* of a closed system and $\hbar$ is a physical constant called *Planck's constant.*

The unitary operator in (3.6) indicates two things. First, the postulate 2 does not conflict with postulate 1, because if the vector $|\varphi_1\rangle$ is a unit vector, then the state $|\varphi_2\rangle$ is also a unit vector. Second, the evolution of a closed quantum system is *invertible*, because we can derive the state $|\varphi_1\rangle$ completely from the state $|\varphi_2\rangle$: $U^\dagger|\varphi_2\rangle = |\varphi_1\rangle$. In another word, during the process of the evolution of a closed quantum system, the information of this system has been preserved without any loss.

The unitary operator U is a function of time t. But, in the thesis, we only concern about the input state $|\varphi_1\rangle$ and the output state $|\varphi_2\rangle$. Therefore, we can regard U as an operator having nothing to do with time.

The Equation (3.6) does not indicate what kind of unitary operators can be used to describe the evolution of a closed quantum system. It turns out that, in the case of closed one-qubit systems, any unitary operator defines an evolution that can be realized in realistic world. Commonly used operators on qubits are Pauli matrices and the Hadamard operator H, which is defined as:

$$H \equiv \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3.8}$$

Hadamard operator has the properties: $H|0\rangle = |+\rangle, H|1\rangle = |-\rangle$.

Quantum measurements

Postulate 3: quantum measurements are described by a collection of quantum operators $\{M_m\}$, where m is the index of the measurement outcomes. These operators are applied on the state space associated with the system being measured. Suppose $|\psi\rangle$ is the system state just before the measurement. The probability that the measurement outcome m occurs is:

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle \tag{3.9}$$

The system state $|\psi'\rangle$ immediately after the measurement is:

$$|\psi'\rangle = \frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}} = \frac{M_m|\psi\rangle}{\sqrt{p(m)}} \tag{3.10}$$

The collection $\{M_m\}$ satisfies the *completeness equation*:

$$\sum M_m^\dagger M_m = I \tag{3.11}$$

The postulate 3 depicts a clear picture of quantum measurements. It gives us the probability of the occurrence of a possible outcome, the system state after the measurement and the property of the measurement operators. The meaning of equations (3.9) and (3.10) is very clear. Then, what is the meaning of the completeness Equation (3.11)? Suppose $|\psi\rangle$ is an arbitrary state of the system being measured by a collection $\{M_m\}$, and we calculate $\sum p(m)$ explicitly:

$$\begin{aligned}\sum p(m) &= \sum\langle\psi|M_m^\dagger M_m|\psi\rangle\\ &= \langle\psi|\sum M_m^\dagger M_m|\psi\rangle\\ &= \langle\psi|I|\psi\rangle\\ &= \langle\psi|\psi\rangle = 1\end{aligned} \tag{3.12}$$

Thus, the completeness equation is merely one of the expressions of the fact that probabilities sum to one.

The most used and important measurement in quantum information theory is the *projective measurement*, defined as:

Projective measurement: the measurement operators of a projective measurement are a collection of orthogonal projectors $\{P_m\}$, which satisfies the completeness equation: $\sum P_m = I$ (Note that $P_m^2 = P_m$). “orthogonal” means $P_iP_j = \delta_{i,j}P_i$. λ_m are the measurement outcomes. The operator $M = \sum \lambda_m P_m$ is called the *observable* of the projective measurement.

In fact, each λ_m is the eigenvalue of M, and P_m is the eigenspace associated to the eigenvalue λ_m. Later, if we say "measure in a basis $\{|m\rangle\}$", where $\{|m\rangle\}$ is an orthonormal basis, we mean a projective measurement whose projectors are $P_m = |m\rangle\langle m|$. For example, if we measure a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ in the basis $\{|0\rangle, |1\rangle\}$, then the projective measurement is $\{P_0 = |0\rangle\langle 0|,\ P_1 = |1\rangle\langle 1|\}$ and we will get the state

$$\frac{P_0|\psi\rangle}{\sqrt{\langle\psi|P_0|\psi\rangle}} = \frac{a|0\rangle}{|a|} \tag{3.13}$$

with the probability $|a|^2$ and the state $\frac{b|1\rangle}{|b|}$ with the probability $|b|^2$. If the measurement outcomes are $\lambda_0 = 1, \lambda_1 = -1$, then the observable

$$M = \lambda_0 P_0 + \lambda_1 P_1 = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.14}$$

is the Pauli matrix Z.

One of the important consequences of quantum measurement is that only orthogonal states can be reliably *distinguished* [5]. Two states are *distinguishable* if and only if there exists a measurement which can tell them apart certainly. Suppose the states $|\psi\rangle, |\varphi\rangle$ are orthogonal. We construct a projective measurement, $P_0 = |\psi\rangle\langle\psi|$, $P_1 = I - P_0$. If the state is $|\psi\rangle$, we get the measurement result: $p(0) = 1, p(1) = 0$; if the state is $|\varphi\rangle$, we get the measurement result: $p(0) = 0, p(1) = 1$. Thus, we can distinguish these two states, and what is more, do not destroy them. If two states are not orthogonal, then there is no such a measurement that can distinguish them without any error.

Now that we have familiarized ourselves with quantum measurements, let's have a look at the relationship between the postulate 2 and the postulate 3. Both of two postulates describe the changes of quantum systems. The difference between these two postulates is that: the postulate 2 describes the changes of a *closed* quantum system, where the force making the system to change is coming from the system itself, while the postulate 3 describes the changes of an *open* quantum system, where the force making the system to change is coming from the outside, from scientists' interference. Although it is possible to consider the system being measured, the quantum measurement device and some other quantum systems together as a closed quantum system, whether or not we can derive the postulate 3 as a consequence of the postulate 2 is still an open problem. Fortunately, the answer to this problem will not affect us very much. What we care about is when to apply postulate 2 and when to apply postulate 3, and in this thesis, the circumstances have always made this point very clear.

Composite systems

Postulate 4: The state space of a composite system is the tensor product of the state spaces of its constituent physical systems. If we number the component of the systems from 1 to n, and the ith component be prepared in the state $|\psi_i\rangle$, then the state of the composite system is: $|\psi_1\rangle\otimes|\psi_2\rangle\otimes\cdots\otimes|\psi_n\rangle$.

A straight conclusion of the postulate 4 is that the state of an n-qubit system is a unit vector over C^{2^n}, for:

$$\left(\langle\psi_1|\otimes\langle\psi_2|\otimes\cdots\otimes\langle\psi_n|\right)\left(|\psi_1\rangle\otimes|\psi_2\rangle\otimes\cdots\otimes|\psi_n\rangle\right)=\prod_{i=1}^{n}\langle\psi_i|\psi_i\rangle=1 \tag{3.15}$$

The basis of a tensor product space may be constructed from its component systems, as we mentioned in chapter 1. For example, for a two-qubit system, if the basis we select for the one-qubit system is $\{|0\rangle, |1\rangle\}$, then the basis of the two-qubit system is $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Normally, we often use two bases for an n-qubit system. One takes $\{|0\rangle, |1\rangle\}$ as the basis for the one-qubit system, and the other one takes $\{|+\rangle, |-\rangle\}$ as the basis for the one-qubit system.

The tensor product space contains a very special and important kind of vectors, called *entangled* states. These vectors are not tensor products themselves. For example, a famous and important entangled state is the *Bell* or *EPR* [36] state:

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{3.16}$$

It gets the name after the initials of Einstein, Podolsky and Rosen, who discovered it. It is easy to check that EPR state cannot be written in the form of a tensor product $|\psi_1\rangle \otimes |\psi_2\rangle$.

The entangled states are very important, for they are the key in most quantum mechanical applications. In an entangled state, the component systems are correlated. To see this correlation, we take the EPR state as an example. Suppose we measure the first qubit of the EPR state in the basis $\{|0\rangle, |1\rangle\}$. We can get two possible resulted states $|00\rangle$ and $|11\rangle$ with the same probability. In either case, the second qubit is in the same state as the first qubit. Then, we measure the second qubit in the basis $\{|0\rangle, |1\rangle\}$ too. Of course, the measurement result is always as same as the measurement result of the first qubit. This measurement correlation exists no matter how far away these two qubits are separated or which qubit we measure first. It has been proven that: the measurement correlation in the EPR state is stronger than could ever exist between classical systems [37]. The EPR state plays an important role in the entanglement enhanced teleportation

[11].

A frequently used notation for composite systems is the subscript notation. The subscript notation, A_i, means that the operator A acts on the *ith* system. For example, Z_3 is an operator acting on the third qubit.

3.3 Density operator

A quantum system can be described by an ensemble $\{p_i, |\psi_i\rangle\}$, which means that the quantum system may be in a states $|\psi_i\rangle$ with its corresponding probability p_i. As all of the possible states of this quantum system are included in the ensemble, the probabilities sum to one: $\sum p_i = 1$. If there is only one state in the ensemble, we say the quantum system is in a *pure* state. Otherwise, we say the quantum system is in a *mixed* state. The ensemble $\{p_i, |\psi_i\rangle\}$ is called an *ensemble of pure states*. So far, our explanation of quantum mechanics is limited in pure state. But more often than not, we need to treat the quantum systems in mixed states. *Density operators* or *density matrices*, which were first developed in [39], arise naturally to fulfill this purpose. In this section, we introduce their basic ideas.

3.3.1 The definition of density operators

For an n-dimensional quantum state space, a *density operator* ρ is an $n \times n$ Hermitian operator which satisfies the following two conditions:

1) (Trace condition) $Tr(\rho) = 1$;

2) (Positivity condition) ρ is a positive operator.

Suppose we have a quantum system described by an ensemble $\{p_i, |\psi_i\rangle\}$, its density operator is:

$$\rho = \sum p_i |\psi_i\rangle\langle\psi_i| \tag{3.17}$$

It is easy to check that the Equation (3.17) meets the definition of density operators. The trace:

$$Tr(\rho) = \sum p_i Tr(|\psi_i\rangle\langle\psi_i|) = \sum p_i = 1 \tag{3.18}$$

It satisfies the trace condition. Let $|\varphi\rangle$ be an arbitrary quantum system state, the inner product:

$$\langle\varphi|\rho|\varphi\rangle = \sum p_i \langle\varphi|\psi_i\rangle\langle\psi_i|\varphi\rangle = \sum p_i |\langle\varphi|\psi_i\rangle|^2 \geq 0 \tag{3.19}$$

Thus, it is a positive operator.

There is a simple criterion to judge whether a density operator is in a pure state or in a mixed state. It is easy to prove that if $Tr(\rho^2) = 1$, ρ is in a pure state; if $Tr(\rho^2) < 1$, ρ is in a mixed state.

For a given density operator ρ, there is always at least one quantum system corresponding to it, for ρ must have a diagonal representation [24, 25] $\rho = \sum \lambda_i |i\rangle\langle i|$, where $\{\lambda_i\}$ are the real and non-negative eigenvalues and $\{|i\rangle\}$ are orthonormal eigenvectors, and due to the property of trace one, the system described by the ensemble $\{\lambda_i, |i\rangle\}$ has the density operator ρ. As a matter of fact, we can often find several quantum systems which give the same density operator. An important fact about these systems is that we can not distinguish them by experiments, provided that no more extra side information is given. Thus, the density operator gives as much information as possible about experiments performed on the system corresponding to it.

3.3.2 Reformulate the four postulates by density operators

As density operators completely characterize quantum systems, we can reformulate the four postulates in the language of density operators. By combining the four postulates in section 3.2 and the Equation (3.17) together, we derive the following results. For the results which are simple and straight, we omit their reformulating process.

Postulate 1: Any isolated physical system is associated with a complex vector space equipped with inner product(in another word, a Hilbert space) which is known as the *system space* of the system. The system space is uniquely described by its density operator. If a quantum system is in the state ρ_i with probability p_i, the density operator of the system space is $\sum p_i \rho_i$.

Postulate 2: The evolution of a closed quantum system is described by a unitary operator *U*, which only depends on time t_1 and time t_2. Suppose at time t_1, the system state is ρ, and at time t_2, the system state is ρ'. Then

$$\rho' = U \rho U^\dagger \tag{3.20}$$

Postulate 3: The measurement of a quantum system is described by a set of operators $\{M_m\}$, where *m* is the index of the measurement results. Suppose the system state right before the measurement was ρ. The probability of getting result *m* is:

$$p(m) = Tr\left(M_m^\dagger M_m \rho\right) \tag{3.21}$$

The system state after the measurement is:

$$\rho_m' = \frac{M_m \rho M_m^\dagger}{Tr\left(M_m^\dagger M_m \rho\right)} \tag{3.22}$$

The measurement operators satisfy the completeness equation:

$$\sum M_m M_m^{\dagger} = I \tag{3.23}$$

***Reformulating process*:**

Suppose we perform a measurement described by the measurement operators M_m and the system was in the initial state $|\psi_i\rangle$ with probability p_i. The probability of getting the result *m* is:

$$p(m|i) = \langle\psi_i| M_m^{\dagger} M_m |\psi_i\rangle = Tr\left(M_m^{\dagger} M_m |\psi_i\rangle\langle\psi_i|\right) \tag{3.24}$$

According to the probability theory, the total probability *p(m)* is:

$$p(m) = \sum p(m|i) p_i = \sum Tr\left(M_m^{\dagger} M_m |\psi_i\rangle\langle\psi_i|\right) p_i = Tr\left(M_m^{\dagger} M_m \rho\right) \tag{3.25}$$

The state after getting the result *m* is:

$$|\psi_i^m\rangle = \frac{M_m |\psi_i\rangle}{Tr\left(M_m^{\dagger} M_m |\psi_i\rangle\langle\psi_i|\right)} \tag{3.26}$$

The density operator after getting the result *m* is:

$$\rho_m = \sum p(i|m) |\psi_i^m\rangle\langle\psi_i^m| \tag{3.27}$$

Using the *Bayes* rule, the probability $p(i|m)$ is:

$$p(i|m) = p(i,m) / p(m) = p(m|i) p_i / p(m) \tag{3.28}$$

We substitute (3.24), (3.25), (3.26) and (3.28) in (3.27) we get:

$$\rho_m = \sum p(i|m) p_i \Big/ p(m) |\psi_i^m\rangle\langle\psi_i^m| \tag{3.29}$$

$$= \sum \frac{p_i M_m |\psi_i\rangle\langle\psi_i| M_m^{\dagger}}{p(m)} \tag{3.30}$$

$$= \frac{M_m \rho M_m^{\dagger}}{Tr\left(M_m^{\dagger} M_m \rho\right)} \tag{3.31}$$

The Equation (3.31) is only the density operator after getting result *m*. What about if we want to know the post-measurement density operator ρ' before we perform the measurement? We know that the post-measurement system is described by the ensemble $\{p(i,m), |\psi_i^m\rangle\}$, then the density operator ρ' is:

$$\rho' = \sum_{i,m} p(i,m) |\psi_i^m\rangle\langle\psi_i^m| \tag{3.32}$$

$$= \sum_{i,m} p(m|i) p_i |\psi_i^m\rangle\langle\psi_i^m| \tag{3.33}$$

We substitute (3.24) and (3.26) in (3.33) to get:

$$\rho' = \sum_{i,m} p(m|i) p_i \frac{M_m |\psi_i\rangle\langle\psi_i| M_m^\dagger}{p(m|i)} = \sum_m M_m \rho M_m^\dagger \tag{3.34}$$

The Equation (3.34) is a very nice and compact formula for the measurement in the language of density operators.

Note that there is another way to express the density operator ρ':

$$\rho' = \sum_m p(m) \rho_m \tag{3.35}$$

The Equations (3.35) and (3.34) are mathematically same, but have different physical explanations. Thus, by choosing an appropriate form when we tackle a problem, we would be able to reveal and understand the physical meaning of the problem much easier.

***Postulate 4*:** The state space of a composite physical system is a tensor product of the state space of its element physical system. Suppose we number the element systems from *1* to *n*, and the *ith* element system is prepared in the state ρ_i, then the state of the composite system is: $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$.

3.3.3 The reduced density operator

Recall that two systems in an entangled state have a definite state, or a pure state, when considered together, but each of them can not be said to be in a definite state on its own.

Then, what do we get if we only observe one of them? Its behavior is described by its own *reduced density operator*. Suppose two systems A and B are in the state ρ^{AB}. The reduced density operator of A is defined as:

$$\rho^{A} \equiv Tr_B\left(\rho^{AB}\right) \tag{3.36}$$

On the right hand side of the Equation (3.36), the notation Tr_B is called the *partial trace over system B*. Let $\rho^{AB} = \sum_i p_i \rho_{A_i} \otimes \rho_{B_i}$, and the partial trace is defined as:

$$Tr_B\left(\rho^{AB}\right) \equiv \sum_i p_i \rho_{A_i} Tr\left(\rho_{B_i}\right) \tag{3.37}$$

A very useful and simple case is:

$$Tr_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) = \langle b_2|b_1\rangle |a_1\rangle\langle a_2| \tag{3.38}$$

Just like a density operator, a subsystem's reduced density operator gives as much information as possible about the experiments applied on this subsystem [40]. In fact, the *correlation* between two systems is defined by the reduced density operators. Let ρ^{AB} be the joint state of systems A and B. The two systems are said to be *correlated* if and only if $\rho^{AB} \neq \rho^{A} \otimes \rho^{B}$. We take the Bell state $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ as an example. The density operator ρ of the Bell state is:

$$\rho = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \tag{3.39}$$

The reduced density operator of the first qubit is:

$$\rho^{1} = Tr_2\left(\rho\right) \tag{3.40}$$

$$= Tr_2\left(\frac{|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|}{2}\right) \tag{3.41}$$

$$=\frac{|0\rangle\langle 0|\langle 0|0\rangle+|0\rangle\langle 1|\langle 0|1\rangle+|1\rangle\langle 0|\langle 1|0\rangle+|1\rangle\langle 1|\langle 1|1\rangle}{2} \qquad (3.42)$$

$$=\frac{I}{2} \qquad (3.43)$$

As $Tr\left(\left(\rho^1\right)^2\right)=\frac{1}{2}<1$, ρ^1 is a mixed state. In another word, the first qubit is not in a definite state. It is also clear that $\rho \neq \rho^1 \otimes \rho^2$, which means Bell state is an entangled state. Both of these conclusions are as we expect.

3.3.4 Unitary freedom in the ensemble for density operators

In the section 3.3.1, we have pointed out that the quantum systems with the same density operator can not be distinguished by experiments, provided that we are not given extra side information about them. In this section, we introduce an important theorem about these systems. This theorem is called *unitary freedom in the ensemble for density operators*, which was first discovered by Schrodinger [41], and later rediscovered and developed by Jaynes [42] and by Hughston, Jozsa and Wootters [40].

Suppose there are two ensembles of pure states $\{p_i, |\psi_i\rangle\}$ and $\{q_i, |\varphi_i\rangle\}$. For the convenience, we define the un-normalized vectors $|\tilde{\psi}_i\rangle=\sqrt{p_i}\,|\psi_i\rangle$ and $|\tilde{\varphi}_i\rangle=\sqrt{q_i}\,|\varphi_i\rangle$. If the number of the vectors in $\{|\tilde{\psi}_i\rangle\}$ is different from that in $\{|\tilde{\varphi}_i\rangle\}$, we can add zero vectors in the smaller set to make them have the same number of vectors.

Theorem 3.1: (Unitary freedom in the ensemble for density operators) two sets $\{|\tilde{\psi}_i\rangle\}$ and $\{|\tilde{\varphi}_i\rangle\}$ have the same density operator if and only if:

$$|\tilde{\psi}_i\rangle=\sum_j u_{ij}\,|\tilde{\varphi}_j\rangle \qquad (3.44)$$

where the u_{ij} is a unitary matrix.

Proof: First, suppose we two sets vectors satisfy Equation (3.44), then:

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{ijk} u_{ij}u_{ik}^* |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{3.45}$$

$$= \sum_{jk}\sum_i \left(u_{ij}u_{ik}^*\right)|\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{3.46}$$

$$= \sum_{jk} \delta_{jk} |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{3.47}$$

$$= \sum_j |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j| \tag{3.48}$$

This means that the two sets have the same density operator.

Second, suppose two sets have the same density operator $\rho = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_i |\tilde{\varphi}_i\rangle\langle\tilde{\varphi}_i|$. The density operator has the spectral decomposition: $\rho = \sum_k \lambda_k |k\rangle\langle k|$. We define $|\tilde{k}\rangle = \sqrt{\lambda_k}\,|k\rangle$. $\{|\tilde{k}\rangle\}$ is an orthogonal basis for the space of ρ, and satisfies:

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_k |\tilde{k}\rangle\langle\tilde{k}| \tag{3.49}$$

As the set $\{|\tilde{\psi}_i\rangle\}$ is totally in the space of ρ, $|\tilde{\psi}_i\rangle$ is the linear combination of $|\tilde{k}\rangle$:

$$|\tilde{\psi}_i\rangle = \sum_k c_{ik} |\tilde{k}\rangle \tag{3.50}$$

We substitute (3.50) into the left hand side of (3.49):

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{ikl} c_{ik}c_{il}^* |\tilde{k}\rangle\langle\tilde{l}| = \sum_{kl}\left(\sum_i \left(c_{ik}c_{il}^*\right)\right)|\tilde{k}\rangle\langle\tilde{l}| \tag{3.51}$$

Comparing the Equation (3.51) with the right hand side of the Equation (3.49), we can get:

$$\sum_i \left(c_{ik}c_{il}^*\right) = \delta_{kl} \tag{3.52}$$

This condition ensures that if c_{ij} is not a square matrix, we can append extra columns to obtain a unitary matrix v_{ij} and add zero vectors to $\{|\tilde{k}\rangle\}$ at the same time, such that: $|\tilde{\psi}_i\rangle = \sum_k v_{ik} |\tilde{k}\rangle$. In the same way, we can find a unitary matrix w_{ij} such that: $|\tilde{\varphi}_i\rangle = \sum_k w_{ik} |\tilde{k}\rangle$. From the unitary matrices v_{ij} and w_{ij}, there is a unitary matrix

$u = vw^\dagger$, which makes $|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$. □

3.4 The Schmidt decomposition and purifications

Entangled states are always a favorite in the research of quantum information theory and quantum computation, because they are the central resource in many applications. We have already introduced the surprising fact that while a composite system is in an entangled state, its subsystems are in mixed states. In this section, we go to more details about the entangled states.

The nature of the entangled states gives rise to two questions. 1) Do the subsystems of an entangled state have something in common? In some cases, a pure state is much easier to handle than a mixed state. Then, 2) can we find a way to purify a mixed-state into a pure-state? The search for answers to these questions has led to two tools of great value for entangled states - *Schmidt decomposition* and *purification*, which were developed by Schmidt [43].

Theorem 3.2*: *(The Schmidt decomposition) Suppose $|\psi\rangle$ is a pure state of a composite system *AB*. There must exist an orthonormal basis $\{|i_A\rangle\}$ of A and an orthonormal basis $\{|i_B\rangle\}$ of B, such that:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \tag{3.53}$$

λ_i are known as the *Schmidt coefficients* and satisfy:

1) $\lambda_i \geq 0$ for any i;

2) $\sum \lambda_i^2 = 1$ (3.54)

Proof: Let $\{|j\rangle\}$ be any fixed orthonormal basis for *A*, and $\{|k\rangle\}$ be any fixed orthonormal basis for B. Then the state $|\psi\rangle$ can be written as:

$$|\psi\rangle = \sum_{jk} a_{jk} |j\rangle |k\rangle \tag{3.55}$$

If $\{|j\rangle\}$ and $\{|k\rangle\}$ do not have the same dimension, we can add zero vectors to the smaller one to make it have the same size as the other one so that a_{jk} is a square matrix. By the singular value decomposition, $a = udv$, where d is a diagonal matrix with non-negative entries, and u and v are unitary matrices. Then:

$$|\psi\rangle = \sum_{ijk} u_{ji} d_{ii} v_{ik} |j\rangle |k\rangle \tag{3.56}$$

If we define $|i_A\rangle \equiv \sum_j u_{ji} |j\rangle$, $|i_B\rangle \equiv \sum_k v_{ik} |k\rangle$ and $d_{ii} \equiv \lambda_i$, we get the result:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \tag{3.57}$$

□

From the theorem 3.2, we can get the reduced density operators of A and B respectively:

$$\rho^A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A| \tag{3.58}$$

$$\rho^B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B| \tag{3.59}$$

The two reduced density operators have the same eigenvalues! As many properties of a system are uniquely determined by the eigenvalues of its density operator, this result is very significant. It seems that, to a certain degree, the pure state of the composite system is a reflection of the common points between its component systems. For example, the non-symmetry pure state $\frac{|00\rangle + |01\rangle + |11\rangle}{\sqrt{3}}$ does not give us any apparent properties. Yet, if we calculate $Tr\left(\left(\rho^A\right)^2\right)$ and $Tr\left(\left(\rho^B\right)^2\right)$, we find that they give the same result, i.e., $\frac{7}{9}$.

The bases $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ are called the *Schmidt bases* for A and B respectively. The number of λ_i's is called the *Schmidt number* of the state $|\psi\rangle$. The Schmidt number is an important parameter to measure the degree of entanglement of an entangled state. Imagine that if the Schmidt number is only one: $|\psi^{AB}\rangle = |i_A\rangle|i_B\rangle$, obviously, there is no correlation. If the Schmidt number is more than one, the correlation exists. The bigger the Schmidt number of a composite system is, the more entanglement this system involves.

We now introduce the second tool, the *Schmidt purification*, which, to some degrees, is the inverse of the Schmidt decomposition. The Schmidt decomposition is from an entangled state to a mixed state, while the Schmidt purification, on the other hand, is from a mixed-state to an entangled state. Suppose that ρ^A is the density operator of A, and its spectral decomposition is $\rho^A = \sum_i \lambda_i |i_A\rangle\langle i_A|$. We mathematically introduce an ancillary system R that has the same state space as A, and construct an entangled state $|\psi^{AR}\rangle = \sum_i \sqrt{\lambda_i}|i_A\rangle|i_R\rangle$. $|i_R\rangle$ is an orthonormal basis of R. The state $|\psi^{AR}\rangle$ is a purification of ρ^A, for $Tr_R\left(|\psi^{AR}\rangle\langle\psi^{AR}|\right) = \rho^A$.

The purification of ρ^A is not unique, for the selection of the orthonormal basis $\{|i_R\rangle\}$ is not unique. Let U_R be an arbitrary unitary operator in R. The state

$$|\psi^{AR'}\rangle = (I_A \otimes U_R)|\psi^{AR}\rangle \tag{3.60}$$

must be a purification of ρ^A, because U_R just turns the orthonormal basis $\{|i_R\rangle\}$ to another orthonormal basis of R. The Equation (3.60) is called the *freedom in purification*.

Summary

In this chapter, we go to details to introduce the primary ideas of quantum mechanics: qubits, postulates, density operators and the Schmidt decomposition and the Schmidt purification. All of these issues are very important for quantum information theory and quantum computation. Qubits are the mathematical models for quantum systems in reality, just as bits in classical world. A qubit is the smallest unit we will deal with in quantum information theory. The four postulates are the fundamental rules which support the whole mansion of quantum mechanics. Density operators give us as much of information as possible about the experiments applied on quantum systems. They are very useful when we study mixed states. The Schmidt decomposition reveals the common properties of the subsystems in an entangled state, and the Schmidt purification tells us how to introduce entanglement to purify a mixed state.

There are many excellent books on quantum mechanics and we suggest some of them: Peres' superb book [44], Sakurai's book [45], Volume III of the excellent series by Feynman, Leighton, and Sands [46], and the book of Cohen-Tannoudji, Diu and Laloe [47, 48]. But all of these books involve too much physics. Therefore, we suggest readers to take them just as references. The contents in this chapter also appear in many books on quantum information theory and quantum computation, and we suggest readers to read these books, for the theory of quantum mechanics in them is easy to understand. Here, we list some of these books: the books of Lo [49], Gruska [50], Nielsen and Chuang [34], Bouwmeester et. al. [51], Alber et. al. [52], the lecture notes of Preskill [53] and the collection of references by Cabello [54] which contains many references to other reviews.

Chapter 4

Open-system dynamics

In Chapter 3, we have introduced that the evolution of a closed quantum system is described by a unitary operator. But, in real life, most quantum systems we are interested in are *open-systems*, which interact with other systems. In this Chapter, we will review the description of open-system dynamics and quantum noise in terms of quantum operations.

4.1 Quantum operations

Throughout this and following sections, we use a set of inventions introduced by Schumacher [58]. The *primary system* is denoted by Q, which is assumed to be a finite-dimensional, with a Hilbert space H_Q. The primary system interacts with an *environment* E, and there may be an *auxiliary system* R. Suppose the initial state of Q is ρ. After interacting with E, the state of Q becomes ρ'. There is a map ε to describe this process: $\rho' = \varepsilon(\rho)$. The map ε is called a *quantum operation* if and only if it satisfies the following three axiomatic properties [55, 56, 57]:

1) $Tr(\varepsilon(\rho))$ is the probability that the process represented by ε occurs. Therefore, $0 \le Tr(\varepsilon(\rho)) \le 1$ for any state ρ.

2) The quantum operation ε is a *convex-linear map* on the set of density operators, that is, for probabilities $\{p_i\}$, we have:

$$\varepsilon\left(\sum_i p_i\rho_i\right)=\sum_i p_i\varepsilon(\rho_i). \tag{4.1}$$

3) ε is a *completely positive* map. That is, for any positive operator A of Q, $\varepsilon(A)$ is always a positive operator. What is more, for any joint positive operator A^{QR} of Q and an arbitrary reference system R, $(I_R\otimes\varepsilon)(A^{QR})$ is always a positive operator as well.

These three axiomatic properties reflect our requirements for a map being a quantum operation. At first, we need the quantum operation to be able to represent part or entire whole physical process, so we have the first property, allowing $0\leq Tr(\varepsilon(\rho))\leq 1$. Since a real density operator is required to meet the condition $Tr(\rho)=1$, we must normalize $\varepsilon(\rho)$ to get the true state of the primary system. Thus, the density operator after the map is $\frac{\varepsilon(\rho)}{Tr(\varepsilon(\rho))}$. But more often than not, we say that $\varepsilon(\rho)$ is the density operator, for the difference between the true density operator and $\varepsilon(\rho)$ is a mere normalization factor.

For example, suppose ρ is the initial state of the primary system and M_i is one of the operators of a quantum measurement described by the collection $\{M_m\}$. We define the quantum operation $\varepsilon_i(\rho)\equiv M_i\rho M_i^{\dagger}$, then the probability that the process represented by ε_i occurs is $Tr(\varepsilon_i(\rho))=Tr(M_i^{\dagger}M_i\rho)$ and the state after ε_i is $\rho_i=\frac{M_i\rho M_i^{\dagger}}{Tr(M_i^{\dagger}M_i\rho)}$. The results are exactly same as what we got in Chapter 3. If the primary system is a closed quantum system associated with a unitary operator U, we can define the quantum operation $\varepsilon(\rho) \equiv U\rho U^{\dagger}$, then $Tr(\varepsilon(\rho)) = 1$ and the final state is $\varepsilon(\rho) = U\rho U^{\dagger}$. From these two examples, we can see the flexibility of the quantum operation. To some degree, we can regard ε as a random variable, and $Tr(\varepsilon) = p(\varepsilon)$.

The second property comes from the physical requirement on quantum operations. Suppose the initial state of the primary system is $\rho = \sum_i p_i \rho_i$. Naturally, the state after ε is:

$$\frac{\varepsilon(\rho)}{Tr(\varepsilon(\rho))} = \sum_i p(i|\varepsilon) \frac{\varepsilon(\rho_i)}{Tr(\varepsilon(\rho_i))} \tag{4.2}$$

According to the basic probability theory, $p(i|\varepsilon) = \frac{p(\varepsilon|i) p_i}{p(\varepsilon)}$, and note that

$Tr(\varepsilon) = p(\varepsilon)$ and $Tr(\varepsilon(\rho_i)) = p(\varepsilon|i)$,

$$\varepsilon(\rho) = p(\varepsilon) \sum_i \frac{p(\varepsilon|i) \cdot p_i}{p(\varepsilon)} \frac{\varepsilon(\rho_i)}{p(\varepsilon|i)} = \sum_i p_i \varepsilon(\rho_i) \tag{4.3}$$

The third property stems from an important requirement on the quantum operation. The quantum operation is a map from a set of density operators to another set of density operators. As long as the input is a valid density operator, the output should also be a valid density operator. This fact gives rise to the third property. This should be true whether the primary system is an independent system or is a subsystem of a larger quantum system. The following example will enable the readers to appreciate the importance of the third property. Suppose the operation T performs the transpose operation:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{T} \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{4.4}$$

It is clear that T takes any one-qubit density operator into another density operator.

But when we apply $I \otimes T$ on a two-qubit state $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$, we get:

$$\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.5}$$

The eigenvalues of this matrix is $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$ and $-\frac{1}{2}$. Apparently, it is not a valid density operator. The result means that T can not do its job for a two-qubit system. Therefore, T is not a quantum operation.

4.2 Representations of quantum operations

The three axiomatic properties are abstracted so highly in the mathematical form that they are not convenient to use when we do calculations. We need a more practical form of the quantum operation. It is quite surprising that these axiomatic properties are enough to deduce the practical representation of quantum operations, the *operator-sum* representation [55, 56, 57]

Theorem 4.1: The map ε is a quantum operation if and only if

$$\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger \tag{4.6}$$

where the operators $\{E_i\}$ map the input Hilbert space to the output Hilbert space, and satisfy $\sum_i E_i^\dagger E_i \le I$.

Proof:

Suppose $\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger$. Obviously, ε satisfies the first and the second axiomatic properties. We only need to check whether it is completely positive. Suppose A is a positive operator of the joint system RQ, and $|\psi\rangle$ is an arbitrary vector of RQ,

$$\langle\psi|(I \otimes \varepsilon)(A)|\psi\rangle = \sum_i \langle\psi|(I \otimes E_i) A (I \otimes E_i^\dagger)|\psi\rangle \tag{4.7}$$

Defining $\langle\psi_i| \equiv \left(I \otimes E_i^\dagger\right)|\psi\rangle$,

$$(4.7) = \sum_i \langle\psi_i|A|\psi_i\rangle \geq 0 \tag{4.8}$$

So $(I\otimes\varepsilon)(A)$ is also a positive operator. Thus, we complete the *if* part of the proof.

Now, we begin to prove the *only if* part. Suppose $\varepsilon(\rho)$ satisfies the three axiomatic properties. What we should do is to find out how to express this map in the form of the operator-sum. We introduce an auxiliary system R, which has the same dimension as Q. Let $\{|i_R\rangle\}$ and $\{|i_Q\rangle\}$ be the orthonormal bases for R and Q, respectively. Because two systems have the same dimension, we can use the same index i to represent the two bases. Define a joint state $|\alpha\rangle$ of RQ by

$$|\alpha\rangle \equiv \sum_i |i_R\rangle|i_Q\rangle \tag{4.9}$$

Here, we neglect the normalization factor of $|\alpha\rangle$ for notational simplicity. The positive operator of $|\alpha\rangle$ is:

$$A = |\alpha\rangle\langle\alpha| = \sum_{ij} |i_R\rangle\langle j_R| \otimes |i_Q\rangle\langle j_Q| \tag{4.10}$$

Then we define a positive operator σ of RQ by

$$\begin{aligned}\sigma &\equiv (I\otimes\varepsilon)(A)\\ &= \sum_{ij} |i_R\rangle\langle j_R| \otimes \varepsilon\left(|i_Q\rangle\langle j_Q|\right)\end{aligned} \tag{4.11}$$

Up to now, we have settled ε into a larger system and we have introduced a new positive operator. Why do we spend time in constructing this new operator? We will soon see that a very interesting thing happens when we try to recover ε from this operator, and all of the new items introduced become part of the operator-sum representation of ε. Let $|\psi\rangle = \sum_i \psi_i |i_Q\rangle$ be any state of system Q. Define a corresponding state $|\tilde{\psi}\rangle$ of system R by

$$|\tilde{\psi}\rangle = \sum_i \psi_i^* \, |i_R\rangle \tag{4.12}$$

Then, we form the following inner product

$$\langle\tilde{\psi}|\sigma|\tilde{\psi}\rangle = \langle\tilde{\psi}|\sum_{ij}|i_R\rangle\langle j_R|\otimes\varepsilon\left(|i_Q\rangle\langle j_Q|\right)|\tilde{\psi}\rangle \tag{4.13}$$

$$= \sum_{ij}\psi_i\psi_j^*\varepsilon\left(|i_Q\rangle\langle j_Q|\right) \tag{4.14}$$

$$= \varepsilon\left(|\psi\rangle\langle\psi|\right) \tag{4.15}$$

Let $\sigma = \sum_i |s_i\rangle\langle s_i|$ be some decomposition of σ. Note that the vectors $|s_i\rangle$ are not necessarily eigenvectors and, obviously, they are not normalized. Then define a map

$$E_i|\psi\rangle \equiv \langle\tilde{\psi}|s_i\rangle \tag{4.16}$$

It is easy to see that E_i is a linear operator on Q. Combining (4.16) and (4.13), we get

$$\langle\tilde{\psi}|\sigma|\tilde{\psi}\rangle = \sum_i \langle\tilde{\psi}|s_i\rangle\langle s_i|\tilde{\psi}\rangle \tag{4.17}$$

$$= \sum_i E_i|\psi\rangle\langle\psi|E_i^\dagger \tag{4.18}$$

$$= \varepsilon\left(|\psi\rangle\langle\psi|\right) \tag{4.19}$$

So, for any pure state of Q, we have

$$\varepsilon\left(|\psi\rangle\langle\psi|\right) = \sum_i E_i|\psi\rangle\langle\psi|E_i^\dagger \tag{4.20}$$

By convex-linearity, we can generalize the formula (4.20) for an arbitrary state ρ,

$$\varepsilon(\rho) = \sum_i E_i\rho E_i^\dagger \tag{4.21}$$

From the first axiomatic property of ε, $0 \le Tr\left(\varepsilon(\rho)\right) \le 1$, we have the result

$$\sum_i E_i^\dagger E_i \le I \tag{4.22}$$

□

There are more interesting things in the proof of Theorem 4.1. Let's look back at the

Equation (4.16)

$$E_i|\psi\rangle \equiv \langle\tilde{\psi}|s_i\rangle$$

The definition of the operators $\{E_i\}$ are based on the vectors $|s_i\rangle$, which are from a decomposition of σ. However, the decomposition of σ is not unique. Naturally, for each decomposition we can define a new set of operators, and all of these sets of operators are valid operator-sum representations for ε. Now that all of these sets of operators represent the same quantum operation, is it possible that there is a definite relation between them? Suppose that $\sigma = \sum_i |s_i\rangle\langle s_i|$ and $\sigma = \sum_i |t_i\rangle\langle t_i|$ are two decompositions. Theorem 3.1 in Chapter 3 tells us that if two sets of vectors have the same density operator, there is a unitary matrix mapping one set of vectors to the other. Thus, we have

$$|s_i\rangle = \sum_j u_{ij}|t_j\rangle \tag{4.23}$$

Where u_{ij} is a unitary matrix. Define two sets of operators $E_{\ i} = \langle\tilde{\psi}|s_i\rangle$ and $F_i = \langle\tilde{\psi}|t_i\rangle$. Obviously, $\varepsilon(\rho) = \sum_i E_i\rho E_i^\dagger = \sum_i F_i\rho F_i^\dagger$. By (4.23), we can find out the relation between $\{E_i\}$ and $\{F_i\}$

$$E_i = \langle\tilde{\psi}|s_i\rangle = \sum_j u_{ij}\langle\tilde{\psi}|t_j\rangle = \sum_j u_{ij}F_j \tag{4.24}$$

We generalize the Equation (4.24) as the following important theorem.

Theorem 4.2*: *(Unitary freedom in the operator-sum representation) Suppose operators $\{E_i\}$ and $\{F_i\}$ correspond to the quantum operations ε and f, respectively. By adding zero operators to the shorter list of two operators, we may ensure that $\{E_i\}$ and $\{F_i\}$ have the same number of elements. Then $\varepsilon = f$ if and only if there exist a unitary matrix u_{ij} such that $E_i = \sum_j u_{ij}F_j$.

Theorem 4.2 is very useful for quantum error-correction. Quantum noise is represented in terms of operator-sum. Different kinds of noise have different operator-sum representations. Theorem 4.2 shows that though some kinds of noise are seemingly different, they give rise to the same dynamics. If we can find a scheme to get rid of one of them, we can use the same scheme to get rid of all of them. Thus, it provides us with a simple and useful way to deal with noise.

From Theorem, 4.2 and 4.1, we can deduce easily another useful theorem.

Theorem 4.3: All quantum operations ε on a system of Hilbert space with dimension d can be expressed by an operator-sum representation having at most d^2 elements,

$$\varepsilon(\rho)=\sum_{i=1}^{M}E_i\rho E_i^\dagger \tag{4.25}$$

where $1\le M\le d^2$.

The proof of this theorem is simple. The Equation (4.16) shows that the number of operators in $\{E_i\}$ depends on the number of vectors $|s_i\rangle$. The largest number of vectors $|s_i\rangle$ is the dimension of σ, which is d^2. Therefore, $\{E_i\}$ has at most d^2 operators.

Operator-sum representations are not the only choice to represent quantum operations. The condition $\sum_i E_i^\dagger E_i=I$ is called *trace-preserving* condition, for it implies that $Tr(\varepsilon(\rho))=1$. The condition $\sum_i E_i^\dagger E_i\le I$ is called *non-trace-preserving* condition. The references [56, 57, 59] provide us an important conclusion that every trace-preserving quantum operation ε has a unitary representation, and now we discuss this situation. Suppose ρ is the initial state of the primary system Q, and $\rho^E=\sum\lambda_l|v_l\rangle\langle v_l|$ is the initial state of the environment E, where states $|v_l\rangle$ are the eigenstates of ρ^E. Two systems form a closed system. They interact for a time, and the interaction is described by a unitary operator U. Then the state of primary system after interaction is:

$$\varepsilon(\rho) = Tr_E\left(U\left(\rho \otimes \rho^E\right)U^\dagger\right) \tag{4.26}$$

The equation (4.26) is the unitary representation of the quantum operation ε. Note that:

$$Tr\left(\varepsilon(\rho)\right) = Tr\left(U\left(\rho \otimes \rho^E\right)U^\dagger\right) = 1 \tag{4.27}$$

The explicit calculation of the Equation (4.26) gives us an operator-sum representation. Let $\{|g_k\rangle\}$ be an orthonormal basis of E, and we get:

$$\varepsilon(\rho) = \sum_k \langle g_k| \left(U \left(\rho \otimes \sum_l \lambda_l |v_l\rangle\langle v_l| \right) U^\dagger \right) |g_k\rangle \tag{4.28}$$

$$= \sum_{kl} \sqrt{\lambda_l} \langle g_k|U|v_l\rangle \rho \langle v_l|U^\dagger|g_k\rangle \sqrt{\lambda_l} \tag{4.29}$$

$$= \sum_{kl} E_{kl} \rho E_{kl}^\dagger \tag{4.30}$$

where $E_{kl} \equiv \sqrt{\lambda_l} \langle g_k|U|v_l\rangle$.

Now, we give an example of unitary representation. Let's find the operator-sum of the quantum operation ε for the system showed in Figure 4.1.

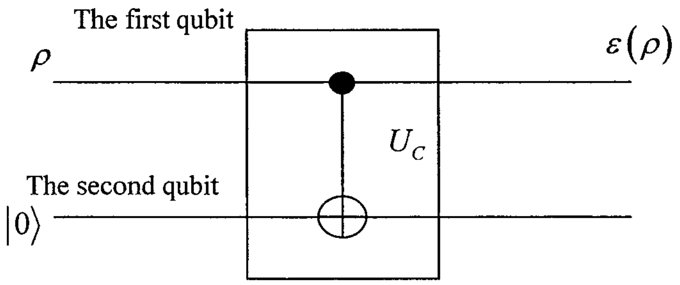

Figure 4.1. Controlled-NOT gate as an elementary example of a single qubit gate

The Controlled-Not gate, U_C, is a two-qubit unitary operator, which is defined as:

$$U_C = |00\rangle\langle 00| + |01\rangle\langle 01| + |11\rangle\langle 10| + |10\rangle\langle 11| \tag{4.31}$$

In this process, the first qubit is the primary system and we can regard the second qubit as the environment. Two qubits establish a closed system whose evolution is described by the Controlled-Not gate. The quantum operation is:

$$\varepsilon(\rho) = Tr_E\left(U_C\left(\rho \otimes |0\rangle\langle 0|\right)U_C^\dagger\right) \tag{4.32}$$

If we choose $\{|0\rangle, |1\rangle\}$ as an orthonormal basis for the second-qubit system, we have:

$$\varepsilon(\rho) = \langle 0|\left(U_C\left(\rho \otimes |0\rangle\langle 0|\right)U_C^\dagger\right)|0\rangle + \langle 1|\left(U_C\left(\rho \otimes |0\rangle\langle 0|\right)U_C^\dagger\right)|1\rangle \tag{4.33}$$

$$= \langle 0|U_C|0\rangle(\rho)\langle 0|U_C^\dagger|0\rangle + \langle 1|U_C|0\rangle(\rho)\langle 0|U_C^\dagger|1\rangle \tag{4.34}$$

Define $E_1 \equiv \langle 0|U_C|0\rangle$ and $E_2 \equiv \langle 1|U_C|0\rangle$, and we get the operator-sum representation,

$$\varepsilon(\rho) = E_1\rho E_1^\dagger + E_2\rho E_2^\dagger \tag{4.35}$$

To find out the exact forms of E_1 and E_2, we substitute the Equation (4.31) in the Equation (4.35), and get the final result,

$$\varepsilon(\rho) = |0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1| \tag{4.36}$$

$$= P_1\rho P_1 + P_2\rho P_2 \tag{4.37}$$

where $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1|$ are projectors.

The way used in unitary representations provides us with some hints to find operator-sum representations for some non-trace-preserving quantum operations. Consider following example. Suppose ρ is the initial state of Q and $\rho^E = \sum \lambda_l |v_l\rangle\langle v_l|$ is the initial state of the environment E, where states $|v_l\rangle$ are the eigenstates of ρ^E. The joint system QE is a closed system and is described by a unitary operator U. After U

acts, we perform a measurement $\{M_m\}$. The quantum operation of the primary system is:

$$\varepsilon_m(\rho) \equiv Tr_E\left(M_m U\left(\rho \otimes \rho^E\right)U^\dagger M_m^\dagger\right) \tag{4.38}$$

and

$$Tr\left(\varepsilon_m(\rho)\right) = Tr\left(M_m U\left(\rho \otimes \rho^E\right)U^\dagger M_m^\dagger\right) \tag{4.39}$$

Let $\{|g_k\rangle\}$ be an orthonormal basis of E, and we get the operator-sum representation: $\varepsilon_m(\rho) = \sum_{kl} E_{kl}\rho E_{kl}{}^\dagger$, where $E_{kl} \equiv \sqrt{\lambda_l}\langle g_k|M_m U|v_l\rangle$.

4.3 Examples of quantum noise

In this section, we present some typical kinds of quantum noise on a single qubit. By doing so, we not only show the power of the operator-sum representation, but also introduce several important quantum noise models which are used frequently in our study of quantum error-correction.

The bit-flip channel

A bit-flip channel flips a qubit from the state $|0\rangle$ to $|1\rangle$, or from $|1\rangle$ to $|0\rangle$ with probability *p*. The bit-flip action is described by the Pauli matrix *X*. Suppose the initial state of the qubit is ρ, the quantum operation is

$$\varepsilon(\rho) = (1-p)\rho + pX\rho X \tag{4.40}$$

Therefore, we have the operation elements:

$$E_0 = \sqrt{1-p}I = \sqrt{1-p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad E_1 = \sqrt{p}X = \sqrt{p}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.41}$$

Note that $\sum_i E_i^\dagger E_i = I$. The physical meaning of the Equation (4.40) is that with probability $(1-p)$, the initial state ρ remains untouched, and with probability p the initial state is replaced by $X\rho X$.

The phase-flip channel

The phase-flip channel has the operation elements:

$$E_0 = \sqrt{1-p}I = \sqrt{1-p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad E_1 = \sqrt{p}Z = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{4.42}$$

Similar to the bit-flip channel, we can see that the physical meaning of this channel is that with the probability p the phase-flip channel turns the state of a qubit from $|1\rangle$ to $-|1\rangle$, or from $-|1\rangle$ to $|1\rangle$, and with the probability $(1-p)$ the channel does nothing to the state. As a special case of the phase-flip channel, we consider the quantum operation when $p = \frac{1}{2}$. From the freedom in the operator-sum representation, this operation can be written in the form

$$\varepsilon(\rho) = P_0\rho P_0 + P_1\rho P_1 \tag{4.43}$$

where $P_0 = |0\rangle\langle 0|$, $P_1 = |1\rangle\langle 1|$. Equation (4.43) describes a quantum measurement of the qubit in the $|0\rangle$, $|1\rangle$ basis. We get the final state $\frac{P_0\rho P_0}{Tr(P_0\rho P_0)}$ with probability $Tr(P_0\rho P_0)$, or the final state $\frac{P_1\rho P_1}{Tr(P_1\rho P_1)}$ with probability $Tr(P_1\rho P_1)$.

The bit-phase flip channel

By combining the bit-flip channel and the phase-flip channel, we get the bit-phase flip channel. The operation elements are

$$E_0 = \sqrt{1-p}I = \sqrt{1-p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad E_1 = \sqrt{p}Y = \sqrt{p}iXZ = \sqrt{p}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{4.44}$$

The depolarizing channel

The depolarizing channel is an important type of quantum noise. The quantum operation is defined as:

$$\varepsilon(\rho) = \frac{pI}{2} + (1-p)\rho \tag{4.45}$$

The meaning of Equation (4.45) is that with the probability p, the density operator has been replaced by the completely mixed state, $\frac{I}{2}$, and with the probability $(1-p)$ the density operator has been left intact. The process of density operator of the qubit being replaced by $\frac{I}{2}$ is called *depolarization*.

Equation (4.45) is not in the operator-sum form. Before we try to find its operator-sum representation, we give two useful results. First, the state of a single qubit can always be written in the *Bloch representation*,

$$\rho = \frac{I + \vec{r}\cdot\vec{\sigma}}{2} \tag{4.46}$$

where $\vec{r}$ is a three component real unit vector, and $\vec{r}\cdot\vec{\sigma} = r_x\sigma_x + r_y\sigma_y + r_z\sigma_z$, where σ_x, σ_y, and σ_z are Pauli matrices. By calculation, we get the explicit representation,

$$\rho = \frac{1}{2}\begin{bmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1-r_z \end{bmatrix} \tag{4.47}$$

Second, we define a map

$$\varepsilon(A) \equiv \frac{A + XAX + YAY + ZAZ}{4} \tag{4.48}$$

It is easy to show that

$$\varepsilon(I) = I; \;\; \varepsilon(X) = \varepsilon(Y) = \varepsilon(Z) = 0 \tag{4.49}$$

Using Equation, (4.47) to (4.49), we can prove that for an arbitrary ρ

$$\frac{I}{2} = \frac{\rho + X\rho X + Y\rho Y + Z\rho Z}{4} \tag{4.50}$$

Then we substitute for $\frac{I}{2}$ in (4.45). At last we come to the equation

$$\varepsilon(\rho)=\left(1-\frac{3p}{4}\right)\rho+\frac{p}{4}(X\rho X+Y\rho Y+Z\rho Z) \qquad (4.51)$$

Equation (4.51) shows that the elements of the quantum operation are $\left\{\sqrt{1-\frac{3p}{4}}I, \sqrt{p}\frac{X}{2}, \sqrt{p}\frac{Y}{2}, \sqrt{p}\frac{Z}{2}\right\}$. There is another convenient form for the depolarizing channel. In Equation (4.51), if we define $p'=\frac{3}{4}p$, we have

$$\varepsilon(\rho)=(1-p')\rho+\frac{p'}{3}(X\rho X+Y\rho Y+Z\rho Z) \qquad (4.52)$$

The physical meaning of Equation (4.52) is that with probability $(1-p')$ the density operator remains intact, and with probability $\frac{p'}{3}$ any one of three operators *X*, *Y*, *Z* is applied.

Summary

In this Chapter, we reviewed open-system dynamics and presented two representations of quantum operations: operator-sum representations and unitary representations. Operator-sum representations are extremely useful in quantum information theory, for they are a universal form to describe the dynamics of quantum systems, no matter open-systems or closed-systems, and they can be put to varied uses. There are many good texts for quantum operations, and we suggest two of them here, Carlton M. Caves's paper [60] and Benjamin Schumacher's paper [61]. Readers can find many useful references in these two papers.

Quantum noise is an application of quantum operations and the kinds of noise reviewed in this chapter are typical ones and are the main concerns in quantum error-correction theory. In real world, quantum noise is very complex. The reason that we can establish the simple noise models is that we do not consider the time, just as what we did to the unitary operators which are associated with the closed quantum systems. There are a great many papers on quantum noise, and we suggest two of them: Davies paper [62] provided a rather mathematical perspective on quantum noise, Gardiner's paper [63] studied quantum noise from the perspective of quantum optics.

Nielsen, M. and Chuang's book [34] also provides a vast amount of useful information on quantum operation and quantum noise.

Chapter 5

Quantum error-correcting codes

For a long time, the scientists had been puzzled by the problem of how to protect quantum information from quantum noise, for there were two seemingly insurmountable obstacles. First, an unknown quantum state cannot be cloned [5]. Second, it is greatly possible that the measurement of a quantum state may collapse the original state to one of a series of outcomes, and therefore the useful information be lost. The first fact seems to make it impossible to add redundant information, and the second fact seems to prevent us from detecting the quantum states. By comparison, classic information, 0 and 1, has never been faced with these serious problems. Copying a bit and measuring the value of a bit are easy to do. However, the discovery of the first quantum error-correcting code by Shor [6] and Steane [64], which could correct an arbitrary one-qubit error, marked the prelude to a series of great progresses in quantum error-correcting codes. In this chapter, we will present some of the most remarkable achievements, including the general error-correction theory, the construction of quantum error-correcting codes and some typical examples of quantum error-correcting codes.

5.1 Independent error model

When we transmit a block of qubits over a noisy quantum channel, the qubits will be affected by noise. The "noise" here means that each qubit being transmitted may, with a small probability p, become entangled with the channel. We consider the independent error model. When we say "independent error model", we mean that the interactions between the qubits and the channel are independent from qubit to qubit. In another word, each operator of operator-sum representation for quantum noise is a tensor product of one-qubit operators. For example, an operator E_i for a two-qubit system may be $E_i = \sigma_x \otimes \sigma_y$. This error model is analogous to the one used very often in the classical theory of error-correction, and is physically reasonable in many situations. In this Chapter, we will always use this model.

5.2 Quantum error-correction theory

At the beginning of this Chapter, we mentioned two seemingly insurmountable obstacles. However, the reason why these seem to be impossible to overcome is that we think quantum mechanics in a classical way. Quantum mechanics has its own unique properties. What is easy in classical information theory becomes almost impossible in quantum information theory, and vice versa. Therefore, we have to consider the problems of quantum error-correction in a quantum-mechanical way.

It is true that it is impossible to clone an unknown quantum state, but cloning is not the only way to add redundancy. Quantum mechanics has already offered us a unique and peculiar tool to solve this problem, the *entanglement*. We can introduce some auxiliary qubits and make them entangled with the qubits we want to transmit. The redundancy is stored in the new entangled state, which will be sent through the noisy channel. When we receive the corrupted states at the other end of the channel, we make use of the auxiliary qubits to recover the original state. Suppose that $|\psi_k\rangle$ is a *k*-qubit state that we would like transmit, and $|0_{n-k}\rangle$ is the state of the auxiliary $n-k$ qubits, which initially are all in the state $|0\rangle$, then the following mapping illustrates the encoding process:

$$|\psi_k\rangle \otimes |0_{n-k}\rangle \xrightarrow{U_{encoding}} |\psi_n\rangle \tag{5.1}$$

where $|\psi_n\rangle$ is the encoded state of n-qubit. We define an $[[n, k]]$ quantum code C to be a unitary mapping of H_2^k into H_2^n [7], where H_2^k is the Hilbert space for k qubits and H_2^n is the Hilbert space for n qubits. Strictly speaking, we can see from (5.1) that the unitary mapping is actually from $H_2^k \otimes H_2^{n-k}$ into a 2^k-dimensional subspace of H_2^n. Later, when we talk about quantum code C, we will use the notation C itself to represent the 2^k-dimensional subspace associated with it.

Let's consider the second obstacle: how to measure a quantum state without destroying it. Although, for most cases, a quantum state is likely to be destroyed by a measurement, there is an exception. For example, suppose that the state $|\psi\rangle = |0\rangle$ is measured by the projective measurement $Z = P_1 - P_{-1}$, where $P_1 = |0\rangle\langle 0|$ and $P_{-1} = |1\rangle\langle 1|$. The measurement result is 1 with probability 1, and the state is still itself, $|\psi\rangle = |0\rangle$. If the state is $|\psi\rangle = |1\rangle$, then the measurement result will be -1 with probability 1, but the state will be changed, from $|1\rangle$ to $-|1\rangle$. From this exception, we can see that if a projective measurement M has a projector with eigenvalue 1, and the code space C happens to fall into this projector space thoroughly, we are able to detect the codewords of C with probability one without changing them. In another word, if all of the codewords are the eigenvectors corresponding to the eigenvalue 1 of M, we can detect this code safely.

A deeper analysis of the above example can lead to the physical approach of the general quantum error-correction theory. Suppose that $M = P_0 + \sum_{i \neq 0} \lambda_i P_i$ is a projective measurement of code C and the code space is in P_0, where $\lambda_i \neq 1$. Let $\{E_e\}$ represent the noise in the channel. If no error acts on C, it is, of course, in the projector P_0. If the error $\{E_e\}$ occurs, we hope that the codewords will be scattered in the projectors $\{P_i\}$ in such a way that we would be able to detect them and turn them back into the code space. Thus, we need to design the code C and the measurement M according to the noise $\{E_e\}$. Without the knowledge of the noise $\{E_e\}$, it is impossible to build an error-correcting code.

Suppose that ρ is the state of the code C, $\varepsilon = \{E_e\}$ is the noise in the quantum channel. C is an error-correcting code for ε if and only if there is a recovery operation $R = \{R_r\}$ which makes

$$R \circ \varepsilon(\rho) = \alpha\rho \tag{5.2}$$

where α is a complex number and $\sum R_r^\dagger R_r = I$. The notation $R \circ \varepsilon(\rho)$ means $R(\varepsilon(\rho))$. The reason that we need α in Equation (5.2) is because the operation ε may be a trace-preserving operation or may not. Therefore, we need α for the purpose of normalization. Of course, the operation R must be a trace-preserving operation, for the correction step must succeed. Obviously, the recovery operation R is the combination of detecting and correcting. Equation (5.2) is equivalent to the following condition: for any $E_e \in \varepsilon$ and $R_r \in R$, we have [66, 67]:

$$R_r E_e = \lambda_{re} I \tag{5.3}$$

where λ_{re} is a complex number.

Two necessary and sufficient conditions for the existence of the recovery operation R were proven independently by Emanuel Knill, Raymond Laflamme [66] and C. H. Bennett, D. P. Divincenzo, J. A. Smolin, and W. K. Wotters [67], who based their work on the paper of Artur Ekert and Chiara Macchiavello [65]. Suppose that $\{|i_L\rangle\}$ is an orthonormal basis of the code space C, and $\{E_e\}$ is the quantum noise. Then the necessary and sufficient conditions are stated as [66]:

Theorem 5.1: The recovery operation R on code C exists if and only if for all basis elements $|i_L\rangle$, $|j_L\rangle$, $(i \neq j)$, and operators E_e, E_f in $\{E_e\}$

$$\langle i_L | E_e^\dagger E_f | i_L \rangle = \langle j_L | E_e^\dagger E_f | j_L \rangle \tag{5.4}$$

and

$$\langle i_L | E_e^\dagger E_f | j_L \rangle = 0 \tag{5.5}$$

The proof of this theorem is long and not easy, but very worth reading. If not interested, readers can skip the proof and go on to the next section. The proof is mainly from [66].

Proof:

Part1: Assume that the error-correction scheme R exists. We can calculate Equation (5.5) explicitly:

$$\langle i_L | E_e^\dagger E_f | j_L \rangle = \langle i_L | E_e^\dagger I E_f | j_L \rangle \tag{5.6}$$

$$= \langle i_L | E_e^\dagger \sum_r R_r^\dagger R_r E_f | j_L \rangle \tag{5.7}$$

$$= \sum_r \langle i_L | E_e^\dagger R_r^\dagger R_r E_f | j_L \rangle \tag{5.8}$$

$$= \sum_r \langle i_L | \bar{\lambda}_{re} \lambda_{rf} | j_L \rangle \tag{5.9}$$

$$= \alpha_{ef} \delta_{ij} \tag{5.10}$$

It is easy to see that Equation (5.10) is the combination of Equations (5.4) and (5.5).

Part 2: Now we start from Equations (5.4) and (5.5) to deduce the recovery operation R. For a basis element $|i_L\rangle$, let V^i be the subspace spanned by the vectors $\{E_e |i_L\rangle\}$ (for all e). Let $\{|v_{(i,r)}\rangle\}$ be an orthonormal basis of V^i. From Equation (5.5), it is easy to see that any two subspaces, V^i and V^j, are orthogonal to each other. Therefore, all of the elements in $\left\{\left|v_{(i,r)}\right\rangle\right\}$ are mutually orthogonal for all i and r. We define the error space:

$$P_E \equiv \sum_{i,r} \left|v_{(i,r)}\right\rangle \left\langle v_{(i,r)}\right| \tag{5.11}$$

$$= \sum_{r} P_r \tag{5.12}$$

where $P_r \equiv \sum_i \left|v_{(i,r)}\right\rangle \left\langle v_{(i,r)}\right|$. Both P_r and V^i are subspaces of error space P_E. To see the relation between them, we arrange all of the elements in $\left\{\left|v_{(i,r)}\right\rangle\right\}$ in a square:

$$\begin{bmatrix} \left|v_{(0,0)}\right\rangle & \cdots & \left|v_{(0,r)}\right\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \left|v_{(i,0)}\right\rangle & \cdots & \left|v_{(i,r)}\right\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{5.13}$$

P_E is spanned by all of the elements in $\left\{\left|v_{(i,r)}\right\rangle\right\}$, V^i is spanned by all of the elements in the ith row, P_r is formed by the rth column. As the number of rows in the square (5.13) is as same as the number of elements in the basis $\left\{\left|i_L\right\rangle\right\}$, we can turn the rth column of square (5.13) into the basis $\left\{\left|i_L\right\rangle\right\}$ by defining the unitary matrix $V_r \equiv \sum_i \left|i_L\right\rangle \left\langle v_{(i,r)}\right|$:

$$V_r \left|v_{(i,r)}\right\rangle = \left|i_L\right\rangle \tag{5.14}$$

The recovery operation R is defined as:

$$R \equiv \left\{I - P_E, R_1, \cdots, R_r, \cdots\right\} \tag{5.15}$$

where

$$R_r \equiv V_r P_r \tag{5.16}$$

Before we perform R to recover a corrupted codeword, we need to introduce another unitary matrix U_i which 1) turns the $0th$ row of $\left|v_{(i,r)}\right\rangle$ to the ith row:

$$U_i\left|v_{(0,r)}\right\rangle=\left|v_{(i,r)}\right\rangle \tag{5.17}$$

and 2) for all E_e,

$$U_iE_e\left|0_L\right\rangle=E_e\left|i_L\right\rangle \tag{5.18}$$

The existence of U_i meeting Equation (5.18) comes from Equation (5.4). Since $\left\langle 0_L\right|E_e^\dagger E_f\left|0_L\right\rangle=\left\langle i_L\right|E_e^\dagger E_f\left|i_L\right\rangle$ for all e and f, the inner product relationships between $\left\{E_e\left|0_L\right\rangle\right\}$ and $\left\{E_e\left|i_L\right\rangle\right\}$ are identical. So, there must be a unitary matrix U_i such that $U_iE_e\left|0_L\right\rangle=E_e\left|i_L\right\rangle$ [68]. To meet Equation (5.17), we can define $\left|v_{(i,r)}\right\rangle$ as $\left|v_{(i,r)}\right\rangle\equiv U_i\left|v_{(0,r)}\right\rangle$.

Now we show how R recovers a corrupted codeword. Suppose $\left|\psi\right\rangle=\sum_i\alpha_i\left|i_L\right\rangle\in C$, after being transmitted through the channel, it becomes:

$$E_e\left|\psi\right\rangle=E_e\sum_i\alpha_i\left|i_L\right\rangle \tag{5.19}$$

$$=\sum_i\alpha_iE_e\left|i_L\right\rangle \tag{5.20}$$

$$=\sum_i\alpha_iU_iE_e\left|0_L\right\rangle \tag{5.21}$$

$$=\sum_{i,r}\alpha_iU_i\beta_{0,r}^e\left|v_{(0,r)}\right\rangle \tag{5.22}$$

$$=\sum_{i,r}\alpha_i\beta_{0,r}^e\left|v_{(i,r)}\right\rangle \tag{5.23}$$

where, $E_e\left|0_L\right\rangle=\sum_r\beta_{0,r}^e\left|v_{(0,r)}\right\rangle$. Equation (5.23) means that the error E_e scatters $\left|\psi\right\rangle$ throughout the error space P_E. Applying R on (5.23), we have:

$$R_rE_e\left|\psi\right\rangle=V_rP_r\sum_{i,k}\alpha_i\beta_{0,k}^e\left|v_{(i,k)}\right\rangle \tag{5.24}$$

$$=V_r\sum_i\alpha_i\beta_{0,k}^e\left|v_{(i,r)}\right\rangle \tag{5.25}$$

$$=\sum_i \alpha_i \beta_{0,r}^e |i_L\rangle \tag{5.26}$$

$$=\beta_{0,r}^e |\psi\rangle \tag{5.27}$$

The physical meaning of Equation (5.24) is that we select the information falling into the projector P_r from the whole error space P_E, and then apply V_r to turn P_r into the code space. $\{P_r\}$ serves as a projective measurement. Because Equation (5.27) is true for all of codewords of C, it indicates that $R_r E_e = \beta_{0,r}^e I$. The fact that R is the recovery operation for code C follows. □

A very special and useful case is when the condition (5.4) becomes

$$\langle i_L | E_e^\dagger E_f | i_L \rangle = \delta_{ef} \tag{5.28}$$

Under this condition, the set of vectors $\{E_e |i_L\rangle\}$ (for all *e*) itself forms a basis for the space V^i. By defining $|v_{(i,e)}\rangle \equiv E_e |i_L\rangle$, we can re-write some definitions:

$$P_E \equiv \sum_{i,e} |v_{(i,e)}\rangle\langle v_{(i,e)}| \tag{5.29}$$

$$=\sum_e P_e \tag{5.30}$$

where $P_e \equiv \sum_i |v_{(i,e)}\rangle\langle v_{(i,e)}|$.

$$\begin{bmatrix} |v_{(0,0)}\rangle & \cdots & |v_{(0,e)}\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ |v_{(i,0)}\rangle & \cdots & |v_{(i,e)}\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{5.31}$$

$$R_e \equiv V_e P_e = V_e \sum_i |v_{(i,e)}\rangle\langle v_{(i,e)}| \tag{5.32}$$

where $V_e |v_{(i,e)}\rangle = |i_L\rangle$.

The projector P_e corresponds to the error operator E_e. The physical meaning of R_e

is that if a state is in P_e, then we apply V_e to turn the state back into the code space C. Equations (5.28) and (5.5) are sufficient conditions for the existence of the recovery operation R. They were first discovered by Artur Ekert and Chiara Macchiavello [65].

5.3 Stabilizer codes

Theorem 5.1 sets a framework for the study of error-correcting codes, but it does not tell us how to construct error-correcting codes. In this section, we introduce a very important way of constructing a class of error-correcting codes, which are called stabilizer codes. Stabilizer codes were discovered independently by A. R. Calderbank, E. M. Rains, P. W. Shor, N. N. A. Sloane [9] and D. Gottesman [10]. The content of this section is mainly from their work.

Theorem 5.1 shows that if a code C is an error-correcting code for the errors $\{E_e\}$, there is a definite relationship between the code space C and error operators $\{E_e\}$. On the other hand, the code space C is part of the eigenvector space corresponding to eigenvalue +1 of a projective measurement M. Therefore, there is a definite relationship between M and $\{E_e\}$ as well (Note that the projector P_r or P_e should be one of projectors of M). If we are able to construct the projective measurement M, it is not a difficult task to find the error-correcting code.

Our discussion is based on the *Pauli group,* G_n, on n qubits, which is defined to consist of all *n*-fold tensor products of Pauli matrices with multiplicative factors $\pm 1, \pm i$. The group is closed under the operation of matrix multiplication. For example,

$$G_1 \equiv \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\} \qquad (5.33)$$

If g is an element of G_n, then g is in the form $aw_1 \otimes \cdots \otimes w_i \cdots \otimes w_n$, where a is one of $\pm 1, \pm i$, and w_i is one of $I, \sigma_x, \sigma_y, \sigma_z$. Obviously, there are 4^{n+1} elements in G_n.

Andrew Steane [8], and Artur Ekert, Chiara Macchiavello [65] have proven independently that if a quantum error-correcting code can correct bit-flip error and phase-flip error on a qubit, this code can correct arbitrary errors on this qubit. Therefore, we only need to take the Pauli matrices $I, \sigma_x, \sigma_y, \sigma_z$ into account when we consider errors. σ_y can be looked as the combination of σ_x and σ_z, $\sigma_y = i\sigma_x\sigma_z$, and identity matrix I can be regarded as a special error. In the light of this result, G_n is actually the ensemble of all possible errors we need to consider.

Instead of one projective measurement, a stabilizer code uses an ensemble of projective measurements S to fix itself, and $S \subset G_n$. Suppose $S = \{g_i\}$, then the code C is the intersection of the eigenvector spaces corresponding to eigenvalue +1 of all of the elements of S. In another word, $C = \{|\psi\rangle : g|\psi\rangle = |\psi\rangle, \forall g \in S\}$. This ensemble S of the projective measurements is called the *stabilizer* of the code C.

Not any set of elements in G_n can be a stabilizer. S has to have some special properties. As any two of Pauli matrices, $I, \sigma_x, \sigma_y, \sigma_z$, either commute or anti-commute, any two elements of G_n must either commute or anti-commute. All elements of S must commute with each other, for if $g_1, g_2 \in S$ anti-commute, $\forall |\psi\rangle \in C$, we have

$$|\psi\rangle = g_1|\psi\rangle = g_1 g_2|\psi\rangle = -g_2 g_1|\psi\rangle = -|\psi\rangle \tag{5.34}$$

Equation (5.34) is impossible to hold for a non-trivial code C. Therefore, All elements of S must commute with each other. If $g_1, g_2 \in S$, then $g_1 g_2$ also satisfies $g_1 g_2|\psi\rangle = |\psi\rangle$. We define the stabilizer S to contain all of these possible measurements. Thus, S is an Abelian subgroup of G_n (note that an Abelian group means that for any two elements in the group, $g * h = h * g$). To guarantee that the code C is non-trivial, we still need one more condition: $-I \notin S$, for $-I$ commutes with any elements of G_n but does not have eigenvalue +1. Now we can generalize the properties of a stabilizer: *a stabilizer is an Abelian subgroup of* G_n *that does not contain* $-I$.

It is much more convenient to describe a group G by its *generators*, which is defined to *be a set of independent elements* $g_1, \cdots, g_l$ *of* G *such that every element of* G *can be written as a product of elements from* $g_1, \cdots, g_l$. The word "*independent*" here means that any g_i in $g_1, \cdots, g_l$ cannot be written as a product of the elements from $g_1, \cdots, g_{i-1}, g_{i+1}, \cdots g_l$. The generators of S is denoted as $\langle g_1, \cdots, g_l \rangle$ and if S has l generators, it is l-dimensional. The generators $\langle g_1, \cdots, g_l \rangle$ are enough to measure the code C,, for other measurements in S do not give us any more information.

Now, let's discuss how the generators partition the total 2^n-dimensional Hilbert space. Any element g in S has the eigenvalues ± 1 and has the property: $g = P_{+1} - P_{-1}$, where P_{+1} is the projector of +1 and equal to $\frac{I+g}{2}$, P_{-1} is the projector of -1 and equal to $\frac{I-g}{2}$, for the simple reasons that:

1) for any eigenvector $|\psi\rangle$ corresponding to eigenvalue +1, $P_{+1}|\psi\rangle = \frac{I+g}{2}|\psi\rangle = |\psi\rangle$ and $P_{-1}|\psi\rangle = \frac{I-g}{2}|\psi\rangle = 0$;
2) for any eigenvector $|\psi\rangle$ corresponding to eigenvalue -1, $P_{+1}|\psi\rangle = \frac{I+g}{2}|\psi\rangle = 0$ and $P_{-1}|\psi\rangle = \frac{I-g}{2}|\psi\rangle = |\psi\rangle$;
3) $P_{+1}P_{-1} = 0$ and $P_{+1} + P_{-1} = I$;

Thus, one element of S partitions the 2^n-dimensional Hilbert space into two orthogonal subspaces. What about two elements g_1 and g_2? g_1 and g_2 have four projectors: $\{P_{+1}^1, P_{-1}^1, P_{+1}^2, P_{-1}^2\}$ and it is easy to prove that all of them commute with each other. They partition the total space into four equal size orthogonal subspaces: $P_{+1}^1P_{+1}^2, P_{-1}^1P_{+1}^2, P_{+1}^1P_{-1}^2, P_{-1}^1P_{-1}^2$. In the same way, the generators $g_1,\cdots,g_l$ of S partition the 2^n-dimensional Hilbert space into 2^l subspaces, each of which is 2^{n-l}-dimensional [9,69]. To represent these subspaces, let $\vec{x} = (x_1,\cdots x_i,\cdots,x_l)$, $x_i \in \{0,1\}$, and define

$$P^{\vec{x}} = \prod_{i=1}^{l} \frac{I+(-1)^{x_i} g_i}{2} \tag{5.35}$$

Each vector $\vec{x}$ represents a subspace. Obviously, the code space C is equal to the subspace corresponding to the zero vector $\vec{x} = (0,\cdots,0)$:

$$P^{(0,\cdots,0)} = \prod_{i=1}^{l} \frac{I+g_i}{2} \tag{5.36}$$

If we use a stabilizer code to encode k qubits into n qubits, the code space C is a 2^k-dimensional subspace. Therefore, the relation between the dimension of the code space and the dimension l of the stabilizer is $k = n - l$. The maximum stabilizer has n generators. In this case, the code C has merely one vector.

To see what kind of errors a stabilizer code can correct, we still need the following knowledge. There are some special elements in G_n which do not belong to S, but commute with every element in S. The *centralizer* $C(S)$ of S is defined to *consist of all elements in* G_n *which commute with every element in* S. Obviously, S is contained in $C(S)$, $S \subseteq C(S)$. Due to the properties of S and G_n, the centralizer $C(S)$ is equal to the *normalizer* $N(S)$ of S in G_n, which is defined to consist of any element A in G_n such that for $\forall s \in S, A^\dagger s A \in S$. To see this [10], suppose that A is any element in G_n, and $s \in S$,

$$A^\dagger s A = \pm A^\dagger A s = \pm s \tag{5.37}$$

As $-I$ is not in S, A is in $N(S)$ if and only if A is in $C(S)$. Therefore, $N(S) = C(S)$. For an $[[n,k]]$ stabilizer code, the stabilizer S has 2^{n-k} elements, and $N(S)$ has 2^{n+k} elements [69].

With the conception of normalizer and the properties of stabilizers, we can introduce an extremely important theorem of stabilizer codes:

Theorem 5.2*: *(Error-correction conditions for stabilizer codes) Let S be the stabilizer for a stabilizer code $C(S)$ and $\{E_e\}$ be a set of operations in G_n such that $E_e^\dagger E_f \notin N(S) - S$ for all e and f. Then $\{E_e\}$ is a correctable set of errors for the stabilizer code $C(S)$.

Proof: Let $E = E_e^\dagger E_f$ and $\{|i\rangle\}$ be a set of orthonormal basis for the code space $C(S)$. As $E \notin N(S) - S$, E is either in S or in $G_n - N(S)$. If E is in S, then

$$\langle i|E|j\rangle = \langle i|j\rangle = \delta_{i,j} \tag{5.38}$$

If E is in $G_n - N(S)$, then there is at least one element g in S anti-commuting with it. Thus:

$$\langle i|E|j\rangle = \langle i|Eg|j\rangle = -\langle i|gE|j\rangle = -\langle i|E|j\rangle = 0 \tag{5.39}$$

Both Equations (5.38) and (5.39) satisfy Equations (5.4) and (5.5). Thus, we get the conclusion. □

Why we cannot correct the errors $E \in N(S) - S$? Suppose $|\psi\rangle$ is a state in the code space $C(S)$, and the operator E is in $N(S) - S$. For any $g \in S$,

$$E|\psi\rangle = Eg|\psi\rangle = g(E|\psi\rangle) \tag{5.40}$$

Equation (5.40) means that the state $E|\psi\rangle$ is also a valid codeword. The operator E maps one codeword to another codeword. Because our error-correction theory is devised for the errors happening outside the code space, we cannot find and correct errors happening inside the code space. Similar phenomena exist in the error-correction theory in classic information theory as well [20].

The generators $\langle g_1, \cdots, g_i, \cdots g_{n-k}\rangle$ of S partition the 2^n-dimensional Hilbert space into 2^{n-k} subspaces and the detection of the stabilizer code $C(S)$ is tightly combined with these subspaces. Each correctable error turns the code space $C(S) = P^{(0,\cdots,0)}$ to one of the 2^{n-k} subspaces $\{P^{\vec{x}}\}$. By measuring the corrupted state $|\psi\rangle$ by every generator g_i, we can fix which subspace the state $|\psi\rangle$ is in. Noting the relation between the vector $\vec{x} = (x_1, \cdots, x_i, \cdots x_{n-k})$ and the measurement results:

$$x_i = \begin{cases} 0, & if\ g_i|\psi\rangle = |\psi\rangle \\ 1, & if\ g_i|\psi\rangle = -|\psi\rangle \end{cases} \tag{5.41}$$

it should not be surprising to the readers if we define the vector $\vec{x}$ as the *error syndrome*.

Now, let's look at the definition of *distance* for a quantum code. It is an analogous to the definition of distance in classical error-correction theory. The *weight* of an error $E \in G_n$ is defined to be the number of terms in the one-qubit operator tensor product which are not equal to the identity. For example, the weight of the operator $I \otimes X \otimes Y \otimes Z \otimes I \otimes I \otimes X \otimes Y$ is 5. The distance of a stabilizer code $C(S)$ is defined to be the minimum weight of an element in $N(S) - S$. We denote an $[[n,k]]$ code with distance d as $[[n,k,d]]$. If the distance d is at least $2t+1$, this code can correct arbitrary errors on any t qubits. It is not difficult to prove this conclusion: if the weights of the elements of an error set $\{E_i\}$ are all smaller or equal to t, then the weight of $E_i^\dagger E_j$, for all i and j, is surely smaller than d, which means that $E_i^\dagger E_j$ is not in $N(S) - S$. The result follows.

There is a very useful and important way [9] to represent the stabilizer of a stabilizer code $C(S)$. Suppose the stabilizer is $S = \langle g_1, \cdots, g_i, \cdots g_{n-k} \rangle$, we define a $1 \times 2n$ binary vector $(a|b)$ for each generator, where both $a = (a_1, \cdots, a_j, \cdots a_n)$ and $b = (b_1, \cdots, b_j, \cdots b_n)$ are $1 \times n$ binary vectors. Let $g_i = w_1 \otimes \cdots \otimes w_j \otimes \cdots \otimes w_n$ be an arbitrary generator of $S = \langle g_1, \cdots, g_i, \cdots g_{n-k} \rangle$, then the corresponding vector $(a|b)$ is defined as:

$$a_j = \begin{cases} 0, & \mathit{if}\ w_j = I\ \mathit{or}\ Z \\ 1, & \mathit{if}\ w_j = X\ \mathit{or}\ Y \end{cases} \qquad b_j = \begin{cases} 0, & \mathit{if}\ w_j = I\ \mathit{or}\ X \\ 1, & \mathit{if}\ w_j = Z\ \mathit{or}\ Y \end{cases} \tag{5.42}$$

Thus, $(n-k)$ independent generators create $(n-k)$ binary vectors, which form an $(n-k)\times 2n$ matrix. We call this matrix the *check matrix* of the stabilizer code $C(S)$. For example, suppose the stabilizer of a stabilizer code $C(S)$ is $\langle Z_1Z_2, Z_2Z_3\rangle$. Then its check matrix is:

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \tag{5.43}$$

$(n-k)$ generators are independent if and only if the rows of the check matrix are linearly independent.

It is not difficult to prove that any two binary vectors $(a|b)$, $(a'|b')$ of a check matrix must satisfy the following condition:

$$a\cdot b' + a'\cdot b = 0 \tag{5.44}$$

or

$$\sum_j a_j b'_j + \sum_j a'_j b_j = 0 \tag{5.45}$$

where "+" denotes the modulo two addition. This condition is just the translation of the commuting relation between any two generators. If two binary vectors do not satisfy (5.44) or (5.45), their corresponding operators must anti-commute. Therefore, the check matrix can be used to tell whether the set of elements it represents forms a stabilizer or not, and that is where its name comes from.

The check matrix has the great advantage of offering the possibility to make use of classical error-correction theory, for we know more about binary vectors than Pauli matrices. Furthermore, the check matrix turns the multiplication in S into the modulo two addition:

$$\forall g, g' \in S, gg' \to (a|b) + (a'|b') = (a+a'|b+b') \qquad (5.46)$$

where $(a|b)$ is the binary vector representation of g and $\left(a'|b'\right)$ is the binary vector representation of g'.

5.4 Examples of stabilizer codes

In this section, we give some examples of stabilizer codes. These include three qubit bit flip code, Shor code and CSS code. Although these codes were discovered before the advent of stabilizer codes, it is good to explain them in terms of the stabilizer language.

Three qubit bit flip code

The code given by Equation (5.43) is called three qubit bit flip code for the reason that it can correct any one-qubit bit-flip errors. It is easy to check that each of the operators $\{X_1, X_2, X_3\}$ anti-commutes with at least one of the elements in the stabilizer $\langle g_1 = Z_1Z_2, g_2 = Z_2Z_3\rangle$. Two generators divide the 2^3-dimensional Hilbert space into 2^2 subspaces, with each subspace being a 2-dimensional space. Therefore, it is a $[[3,1]]$ quantum code. The stabilizer fixes the code space spanned by the basis $\{|000\rangle, |111\rangle\}$, and we take the map $|0\rangle \to |000\rangle, |1\rangle \to |111\rangle$. Suppose the initial state $\alpha|0\rangle + \beta|1\rangle$ has been encoded into the state $\alpha|000\rangle + \beta|111\rangle$. Table 5.1 lists the relationship among the errors, measurement results, and the subspaces.

Table 5.1. Error Syndromes for the three qubit bit flip code

Error syndrome $\vec{x}$	Z_1Z_2	Z_2Z_3	Error type	Projector of the subspace
(0,0)	+1	+1	No error	$P_0 = \lvert 000\rangle\langle 000\rvert + \lvert 111\rangle\langle 111\rvert$
(1,0)	-1	+1	Bit 1 flipped	$P_1 = \lvert 100\rangle\langle 100\rvert + \lvert 011\rangle\langle 011\rvert$
(1,1)	-1	-1	Bit 2 flipped	$P_2 = \lvert 010\rangle\langle 010\rvert + \lvert 101\rangle\langle 101\rvert$
(0,1)	+1	-1	Bit 3 flipped	$P_3 = \lvert 001\rangle\langle 001\rvert + \lvert 110\rangle\langle 110\rvert$

After detecting the error, we can flip the erroneous bit back to its original state.

A very similar code is the three qubit phase flip code. Noting that the operator Z changes the state $|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ to the state $|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$, and vice versa, it is not surprising that we take the map $|+\rangle \to |+++\rangle, |-\rangle \to |---\rangle$ to encode the qubit. The phase flip error is actually the "bit flip error" in the basis $\{|+\rangle, |-\rangle\}$. As the relationship between the bases $\{|+\rangle, |-\rangle\}$ and $\{|0\rangle, |1\rangle\}$ is: $|+\rangle = H|0\rangle$, $|-\rangle = H|1\rangle$, where H is the Hadamard matrix $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, we can easily get the stabilizer of the three qubit phase flip code:

$$g_1 = (H_1Z_1H_1)(H_2Z_2H_2) = X_1X_2,\quad g_2 = (H_2Z_2H_2)(H_3Z_3H_3) = X_2X_3 \qquad (5.47)$$

The Shor code

The three qubit bit flip code and the three qubit phase flip code do not have any practical values, but by combining them together, we get to the Shor code [6, 64], the first quantum error-correcting code which is able to correct arbitrary one qubit errors. This code proved the possibility of the quantum error correction, and the quantum error-correction theory has been developing fast since then.

The generators of the stabilizer of the Shor code are listed in Table 5.2.

Table 5.2. The generators for the Shor code.

g_1	$ZZIIIIIII$
g_2	$IZZIIIIII$
g_3	$IIIZZIIII$
g_4	$IIIIZZIII$
g_5	$IIIIIIZZI$
g_6	$IIIIIIIZZ$
g_7	$XXXXXXIII$
g_8	$IIIXXXXXX$

As there are eight independent generators and any one qubit error anti-commutes with, at least, one of these generators, the Shor code is an $[[9,1,3]]$ code. The mapping of one qubit to nine qubits is carried out in two steps. First, we encode the qubit by phase flip code: $|+\rangle \to |+++\rangle, |-\rangle \to |---\rangle$. Second, we encode the phase flip code by bit flip code: $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}} \to \frac{|000\rangle+|111\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} \to \frac{|000\rangle-|111\rangle}{\sqrt{2}}$. Thus, the mapping of one qubit to nine qubits is:

$$|0\rangle \to \frac{(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)}{2\sqrt{2}} \tag{5.48}$$

$$|1\rangle \rightarrow \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}} \qquad (5.49)$$

Now, let's have a look at how it works. Suppose a bit flip error occurs on any qubit in the first block of three qubits, we can use the generators g_1, g_2 to measure this block. Similar to what we do in the three qubit bit flip code, we can detect the error and correct it. The generators g_3, g_4 are for the second block, and g_5, g_6 are for the third block. So, the Shor code can correct one qubit bit flip errors. Suppose a phase flip error occurs, then no matter in which block it happens, the effect is the same: the sign of that block has been changed, and become different from the other two blocks.

$$|000\rangle + |111\rangle \rightarrow |000\rangle - |111\rangle, |000\rangle - |111\rangle \rightarrow |000\rangle + |111\rangle \qquad (5.50)$$

We can use generators g_7, g_8 to measure these three blocks. Table 5.3 lists all possibilities.

Table 5.3. Phase flip errors for the Shor code.

$X_1X_2X_3X_4X_5X_6$	$X_4X_5X_6X_7X_8X_9$	Error type
+1	+1	No phase flip errors
-1	+1	Block 1 has been affected
-1	-1	Block 2 has been affected
+1	-1	Block 3 has been affected

Because the Shor code can correct one qubit bit flip error and one qubit phase flip error on any of nine qubits, it can correct arbitrary one qubit errors. The Shor code is not an efficient code, but it is a very good example to explain quantum error-correcting codes, for it is easy to understand.

CSS (Calderbank-Shor-Steane) codes

The CSS codes [7, 8] are an important subclass of stabilizer codes. It is important not only because it was the first method which enables us to construct a series of quantum codes, but also because it proved the connection between the classical error-correcting codes and the quantum error-correcting codes. The construction of the CSS codes completely depends on finding some particular classical error-correcting codes. To be prepared to discuss the CSS codes, we need some knowledge of linear codes [20] (in this section, we are only concerned about linear codes over $GF(2)$).

An $[n, k, d]$ linear code encodes k-bit messages into n-bit codewords. In another word, it maps the k-dimensional space into a k-dimensional subspace of the n-dimensional space. We denote this k-dimensional subspace as the code C. There are 2^k codewords. The *(Hamming) weight* of a codeword x is the number of ones in this codeword, and is denoted as $wt(x)$. For example, suppose $x=(0,1,1,0,1)$, then $wt(x)=3$. The *(Hamming) distance* between two codewords is the number of positions where the two codewords differ. For example, suppose $y=(0,0,1,1,1)$, then distance between x and y is $d(x,y)=2$. It is easy to see that:

$$d(x,y)=wt(x+y) \tag{5.51}$$

The *(Hamming) distance* of a linear code C is the minimum distance between two codewords in C, and is denoted as d. It is also equal to the minimum weight of the non-zero codewords in C. A linear code with distance $d \geq 2t+1$, for some integer t, can correct errors on up to t bits.

The relationship between the k-bit messages and the n-bit codewords is determined by the *generator matrix* G, which is a $k \times n$ matrix. Suppose m is an arbitrary k-bit message, then its corresponding codeword u is generated by $u = mG$. The k rows of G are a basis for the k-dimensional subspace. By selecting a different basis to form G, we map the same message to a different codeword. Therefore, the messages and codewords themselves are fixed, but the mapping relation between them depends on the generator matrix we choose.

The *scalar product* of two vectors $u = (u_1, \cdots, u_i, \cdots, u_n)$ and $v = (v_1, \cdots, v_i, \cdots, v_n)$ is defined as $u \cdot v = \sum_i u_i \cdot v_i$, where the addition is modulo two addition. If $u \cdot v = 0$, u and v are called *orthogonal*. For a linear code C, there are many vectors in the n-dimensional space which are orthogonal to all of codewords in C. The *dual* of C, denoted $C^{\perp}$, is defined as $C^{\perp} = \{v : v \cdot u = 0, \forall u \in C\}$. $C^{\perp}$ is an $(n-k)$-dimensional subspace. The *parity check matrix* H, an $(n-k) \times n$ matrix, of the linear code C is defined to be any generator matrix of $C^{\perp}$. Apparently, for any codeword $u \in C$, we have:

$$GH^T = 0 \rightarrow uH^{\mathrm{T}} = 0 \tag{5.52}$$

If a codeword $u \in C$ has been corrupted by an error $e \notin C$, we can detect this error by H:

$$(u+e)H^T = uH^T + eH^T = eH^T \neq 0 \tag{5.53}$$

H gets its name from the properties showed in Equations (5.52) and (5.53). A very important and useful result from Equation (5.52) is that if the distance of the linear code C is d, then any $d-1$ columns of H are linearly independent, but there exists a set of d columns that are linearly dependent.

With this knowledge, we can go on to discuss the CSS codes. For any CSS code, the check matrix A of its stabilizer S takes the form:

$$A = \left[\begin{array}{c|c} H\left(C_2^{\perp}\right) & 0 \\ 0 & H\left(C_1\right) \end{array} \right] \tag{5.54}$$

where C_1 and C_2 are $[n,k_1]$, $[n,k_2]$ classical linear codes so that $C_2 \subset C_1$ and both C_1 and $C_2^{\perp}$ correct t errors. $H\left(C_2^{\perp}\right)$ is the parity check matrix of $C_2^{\perp}$, and $H\left(C_1\right)$ is the parity check matrix of C_1. A is an $\left(n-k_1+k_2\right)\times 2n$ matrix.

To see why A is a check matrix, we only need to check whether any two rows of A satisfy the condition (5.44) or (5.45). Noting that $H\left(C_2^{\perp}\right)=G\left(C_2\right)$, where $G\left(C_2\right)$ is a generator matrix of C_2, and $C_2 \subset C_1$, we get:

$$G(C_2)\left(H(C_1)\right)^{\mathrm{T}}=0 \leftrightarrow H\left(C_2^{\perp}\right)\left(H(C_1)\right)^{\mathrm{T}}=0 \tag{5.55}$$

Thus, A really represents a stabilizer, and fix a stabilizer code. Suppose $\{E_i\}$ is a set of error operations in G_n, and the weight of every element of $\{E_i\}$ is less than or equal to t. Then for any $E_i, E_j \in \{E_i\}$, the weight of $E_i^{\dagger}E_j$ is certainly less than $d=2t+1$. Let $\left(a|b\right)$ be the binary vector representation for $E_i^{\dagger}E_j$, and the hamming weights of a and b be, of course, less than d. This indicates that both a and b are not in $C_2^{\perp}$ or C_1. So,

$$a\left(H(C_1)\right)^{T} \neq 0, \quad b\left(H\left(C_2^{\perp}\right)\right)^{T} \neq 0 \tag{5.56}$$

Equation (5.56) means that $E_i^{\dagger}E_j$ anti-commutes with at least one of generators in S. Therefore, $E_i^{\dagger}E_j \notin N\left(S\right)$. According to Theorem 5.2, the set of errors $\{E_i\}$ can be corrected by the CSS code S. So, the CSS code S is an $\left[\left[n, k_1-k_2, d\right]\right]$ code, which can correct errors on up to t qubits.

To encode k qubits to n qubits by the CSS code S, we still need to construct an orthonormal basis $\{|i_L\rangle\}$ for the code space, the 2^k-dimensional subspace fixed by S. The construction of the basis relies on the linear codes C_1 and C_2. The number of cosets [20] of C_2 in C_1 is:

$$\frac{|C_1|}{|C_2|} = 2^{k_1-k_2} = 2^k \tag{5.57}$$

We denote these 2^k cosets as $\{Coset_i\}$, $1 \le i \le 2^k$. The number of the elements in $\{|i_L\rangle\}$ is the same as the number of the cosets $\{Coset_i\}$. We construct the basis $\{|i_L\rangle\}$ in the following way:

$$|i_L\rangle \equiv \frac{1}{\sqrt{|C_2|}} \sum_{j=1}^{|C_2|} |c_{ij}\rangle, \qquad c_{ij} \in Coest_i \tag{5.58}$$

Each element of $\{|i_L\rangle\}$ corresponds to a coset in $\{Coset_i\}$.

Now, we show that Equation (5.58) really constructs an orthonormal basis. First, we show that $|i_L\rangle$ is in the code space. Let g be an arbitrary generator in $\left[H\left(C_2^{\perp}\right)|0\right]$, the "upper part" of the check matrix A, and $(a|0)$ be its binary vector representation. Since g consists only of X operators, we have

$$g|i_L\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{j=1}^{|C_2|} |c_{ij} + a\rangle \tag{5.59}$$

Since $a \in C_2$,

$$\sum_{j=1}^{|C_2|} |c_{ij} + a\rangle = \sum_{j=1}^{|C_2|} |c_{ij}\rangle \rightarrow g|i_L\rangle = |i_L\rangle \tag{5.60}$$

Let g' be an arbitrary generator in $\left[0|H(C_1)\right]$, the "lower part" of the check matrix

A, and $(0|b)$ be its binary vector representation. Since g' consists only of Z operators, we have

$$g'|i_L\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{j=1}^{|C_2|} (-1)^{c_{ij}\cdot b} |c_{ij}\rangle \tag{5.61}$$

Since $b \in H(C_1), c_{ij} \in C_1$,

$$c_{ij} \cdot b = 0 \to g'|i_L\rangle = |i_L\rangle \tag{5.62}$$

Equations (5.60) and (5.62) have proved that $|i_L\rangle$ is in the code space. Second, because any two cosets of $\{Coset_i\}$ are totally different, having no common vectors, we draw the conclusion that for any two $|i_L\rangle, |j_L\rangle \in \{|i_L\rangle\}$, $\langle i_L | j_L\rangle = \delta_{i,j}$. Therefore, $\{|i_L\rangle\}$ is an orthonormal basis for the code space.

The Steane code

The Steane code is an important example of the CSS code. It first appeared in [4,6]. In the check matrix A, we select C_1 as the $[7,4,3]$ *Hamming code* [20], whose parity check matrix is:

$$H(C_1) = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{5.63}$$

This code has the property $C_1^{\perp} \subset C_1$. If we let $C_2 \equiv C_1^{\perp}$, then $C_2 \subset C_1$ and both $C_2^{\perp}$ and C_1 can correct one bit error. The parity check matrix of $C_2^{\perp}$ is equal to the parity check matrix of C_1, $H(C_2^{\perp}) = H(C_1)$. As C_2 is a $[7,3,3]$ linear code, the Steane code is a $[[7,1,3]]$ quantum error correcting code, taking the map:

$$|0\rangle \rightarrow |0_L\rangle = \frac{1}{\sqrt{8}}\Big[|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle \\ + |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle\Big] \tag{5.64}$$

$$|1\rangle \rightarrow |1_L\rangle = \frac{1}{\sqrt{8}}\Big[|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle \\ + |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle\Big] \tag{5.65}$$

Summary

In this chapter, we introduced the general quantum error-correction theory, the stabilizer codes, and some important examples of stabilizer codes. The puzzles and the solutions for the quantum error-correction lie in the intrinsic properties of quantum mechanics. We have seen how the basic postulates of quantum mechanics play their own roles. The stabilizer of a stabilizer code is actually a series of projective measurements, which can detect certain quantum states without destroying them. The status of the stabilizer codes in the quantum error-correcting codes is very high, for they brought group theory to the construction of codes. Group theory is the mathematical foundation for the classical error-correcting codes as well. Thus, both kinds of codes grow from the same root. The CSS codes were the first codes to reveal the connection between the classical error-correcting codes and the quantum error-correcting codes. The original proof of the CSS codes in [7, 8] is not simple. However, as we already have the knowledge of the stabilizer codes, we explain the principles of the CSS codes in a simple way. Though we gave some important results in this chapter, much of quantum error-correction theory and error-correcting codes has not been touched. If interested, readers can go to the references given in this Chapter.

Chapter 6

The additive cyclic quantum error-correcting codes

The stabilizer codes [9, 10] offer us a general description of quantum error-correcting codes which is based on the independent error model. However, finding the stabilizers, especially the stabilizers of good quantum codes, has proven to be difficult. Though the CSS codes [7, 8] can construct a series of quantum codes, they are just small part of the stabilizer codes. We still needed a more powerful method to find all of the stabilizer codes. This problem now has been solved. In 1996, A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane [70] devised a method which sounds surprisingly simple, but successfully turns the problem of finding stabilizers into the problem of finding *additive self-orthogonal codes* over $GF(4)$ [20] with respect to a certain trace inner product. In this Chapter, we will first introduce this method. Then based on their work, we will focus on a particular class of quantum codes, the *additive cyclic quantum error-correcting codes*. A new search algorithm will be discussed and a series of good quantum error-correcting codes obtained from this search algorithm are presented.

6.1 Stabilizer codes and self-orthogonal codes over $GF(4)$

The key idea of [70] is that a stabilizer code is equivalent to a special code over $GF(4)$. In this section, we introduce this idea step by step. The check matrices provide us with a binary representation for the stabilizers. First, we give some definitions and then discuss the stabilizer codes completely in the language of binary vectors. Let $\overline{E}$ denote a $2n$-dimensional binary vector space. Any vector in $\overline{E}$ can be written in the form $(a|b)$. We define the inner product:

$$\left((a|b),(a'|b')\right) \equiv a \cdot b' + a' \cdot b \tag{6.1}$$

where $a \cdot b'$ and $a' \cdot b$ are the scalar products defined in Chapter 5. We define the *weight* of $(a|b) = (a_1, \cdots, a_i, \cdots, a_n | b_1, \cdots, b_i, \cdots, b_n)$, written $wt((a|b))$, to be the number of coordinates i where either a_i or b_i is not equal to 0. Note that $((a|b),(a|b)) = 0$. We define the *distance* between two vectors $(a|b)$ and $(a'|b')$ as $d((a|b),(a'|b')) \equiv wt((a|b)+(a'|b'))$. Two vectors, $(a|b)$ and $(a'|b')$, are said to be *orthogonal* if their inner product, as defined in (6.1), is zero. The *dual* of a subspace $\overline{S}$ in $\overline{E}$ is defined to be:

$$\overline{S}^{\perp} = \left\{u \in \overline{E} : u \cdot v = 0 \textit{ for all } v \in \overline{S}\right\} \tag{6.2}$$

So far, the definitions of the weight and distance are equivalent to the corresponding definitions of an element in G_n, and the orthogonal and non-orthogonal correspond to the commute and anti-commute, respectively. The problem of finding a stabilizer S has therefore been changed into the problem of finding a subspace $\overline{S}$ in $\overline{E}$ which is contained in its dual, $\overline{S} \subseteq \overline{S}^{\perp}$. $\overline{S}$ is actually the binary representation of S, and $\overline{S}^{\perp}$ is the binary representation of $N(S)$. Recalling Theorem 5.2 in Chapter 5, readers will not find it difficult to understand the following theorem [70]:

***Theorem 6.1*:** Suppose that $\overline{S}$ is an $(n-k)$-dimensional subspace in $\overline{E}$ which is contained in its dual $\overline{S}^{\perp}$, and the weights of the vectors in $\overline{S}^{\perp} - \overline{S}$ are all $> d$, then there exists an $[[n, k, d]]$ quantum error-correcting code.

A quantum error-correcting code obtained from Theorem 6.1 is called an *additive code*, for the simple reason that $\overline{S}$ is closed under the addition. The stabilizer codes and the additive codes refer to the same thing.

Describing the stabilizers in the language of a binary space is our first step to introduce the method of [70]. Now, we take the second step by moving from the binary space $\overline{E}$ to the Galois field $GF(4)$. $GF(4)$ is denoted as $\{0,1,w,\overline{w}\}$. The *trace* of an element $\beta \in GF(4)$ is defined as:

$$Tr(\beta)=\beta+\overline{\beta} \tag{6.3}$$

$GF(4)$ is a vector space of dimension 2 over $GF(2)$. If we choose $\{w,\overline{w}\}$ as a basis, any element in $GF(4)$ can be expressed as a combination of these two elements over $GF(2)$: $0=0\cdot w+0\cdot\overline{w}$, $1=1\cdot w+1\cdot\overline{w}$, $w=1\cdot w+0\cdot\overline{w}$ and $\overline{w}=0\cdot w+1\cdot\overline{w}$. To make it simple, we can use a 2-dimensional binary vector space to represent $GF(4)$:

$$0:(0,0) \qquad 1:(1,1) \qquad w:(1,0) \qquad \overline{w}:(0,1)$$

In the same way, we can use the $2n$-dimensional binary vector space $\overline{E}$ to represent the vectors in $GF(4)^n$. The relation Φ between a vector $u\in GF(4)^n$ and its binary vector representation $(a|b)$ is defined as:

$$u\equiv\Phi((a|b))\equiv aw+b\overline{w} \qquad (a|b)\equiv\Phi^{-1}(u) \tag{6.4}$$

Φ has the property:

$$\Phi((a|b)+(a'|b'))=\Phi((a|b))+\Phi((a'|b')) \tag{6.5}$$

For example, the vector $(0,1,w,\overline{w},w,1,1)$ in $GF(4)^7$ has the binary representation $(0,1,1,0,1,1,1|0,1,0,1,0,1,1)$.

The *Hamming weight* of a vector $u\in GF(4)^n$, written $wt(u)$, is defined to be the number of its nonzero components. The *Hamming distance* between two vectors $u,u' \in GF(4)^n$ is defined as $d(u,u') \equiv wt(u+u')$. The minimal Hamming distance between the vectors of a subset C of $GF(4)^n$ will be denoted by $d(C)$. It is easy to see that $wt(u) = wt(\Phi^{-1}(u))$ and $d(u,u') = d(\Phi^{-1}(u), \Phi^{-1}(u'))$.

Suppose $u,u' \in GF(4)^n$ and $\Phi^{-1}(u) = (a|b)$, $\Phi^{-1}(u') = (a'|b')$, we define the *trace inner product* of u and u' as $u * u' \equiv Tr(u \cdot \overline{u'})$, where $u \cdot \overline{u'} \equiv \sum_i u_i \overline{u'_i}$ is the classical *Hermitian inner product*. After calculating the trace inner product explicitly:

$$Tr(u \cdot \overline{u'}) \equiv Tr((wa + \overline{w}b) \cdot (\overline{w}a' + wb')) \tag{6.6}$$

$$= (a \cdot a')Tr(1) + (a \cdot b')Tr(\overline{w}) + (b \cdot a')Tr(w) + (b \cdot b')Tr(1) \tag{6.7}$$

$$= a \cdot b' + b \cdot a' \tag{6.8}$$

we draw the conclusion that the trace inner product of $u,u' \in GF(4)^n$ is equivalent to the inner product of $\Phi^{-1}(u), \Phi^{-1}(u') \in \overline{E}$ with respect to (6.1). We say that $u,u' \in GF(4)^n$ are orthogonal if and only if $Tr(u \cdot \overline{u'}) = 0$, or in another word, $u,u' \in GF(4)^n$ are orthogonal if and only if $\Phi^{-1}(u), \Phi^{-1}(u') \in \overline{E}$ are orthogonal.

If C is a subgroup in $GF(4)^n$, the *dual* of C is defined to be:

$$C^{\perp} = \{u \in GF(4)^n : u * v = 0 \textit{ for all } v \in C\} \tag{6.9}$$

As C is closed under addition, we will refer to it as an *additive code* over $GF(4)$, and if it has 2^{n-k} vectors, we say that it is an $(n, 2^{n-k})$ code. Then, $C^{\perp}$ is an $(n, 2^{n+k})$ code. Note that we use $(\ )$ to refer to the codes over $GF(4)$, and $[[\ ]]$ to refer to the quantum codes. If $C \subseteq C^{\perp}$, we say C is a *self-orthogonal* code, and if $C = C^{\perp}$, we say C is a *self-dual* code. Now, we can reformulate Theorem 6.1.

Theorem 6.2: Suppose C is an $\left(n, 2^{n-k}\right)$ additive self-orthogonal code of $GF\left(4\right)^n$, and there are no vectors of weight $<d$ in $C^{\perp}-C$, then there is an $\left[\left[n, k, d\right]\right]$ additive quantum error-correcting code.

We say that C is *pure* if $d\left(C\right) \geq d$; otherwise, we say C is *impure*. The associated quantum error-correcting code is said to be *pure* if C is pure; otherwise, we call the quantum code *impure*. If C is not only closed under addition, but also closed under multiplication by w as well, we say it is *linear* and the associated quantum code is also *linear*.

Theorem 6.2 has finished the transformation: turning a stabilizer S into an additive self-orthogonal code C over $GF(4)$. As we have already accumulated a vast amount of knowledge about the codes over $GF(4)$, we have therefore found a resourceful source for the construction of the stabilizers. We should emphasize that the parameters of an additive self-orthogonal code over $GF\left(4\right)$ are not the parameters for an additive quantum error-correcting code, but are the parameters of the stabilizer of that quantum code. For example, if an additive self-orthogonal code C is an $\left(n, 2^l\right)$ code, the corresponding quantum error-correcting code is not an $\left[\left[n, l\right]\right]$ code. The parameter 2^l is the number of the elements in the corresponding stabilizer S, and l is the number of the independent generators which generate S. These independent generators divide the total 2^n-dimensional Hilbert space into 2^l subspaces, each of which is a 2^{n-l}-dimensional subspace, and the quantum code space is one of them. Thus, the corresponding quantum code is $\left[\left[n, k = n-l\right]\right]$. The distance of this code is found from the set $C^{\perp} - C$. In a word, the self-orthogonal codes over $GF(4)$ are just the representations of the stabilizers: $C \leftrightarrow \overline{S} \leftrightarrow S, C^{\perp} \leftrightarrow \overline{S}^{\perp} \leftrightarrow N(S)$.

6.2 Additive cyclic quantum codes

The paper [70] offers many methods of constructing the additive self-orthogonal codes and tabulates $[[n,k,d]]$ additive quantum error-correcting codes for $n \leq 30$ with the highest achievable minimal distance d. It also gives the possible upper bounds for the minimal distance d. Though we are not going to introduce all of these methods, there is one kind of codes appealing to us very much: the *additive cyclic quantum error-correcting codes*. An additive code C over $GF(4)^n$ is *cyclic* if and only if $(u_0, u_1, \cdots, u_{n-1}) \in C$ implies $(u_{n-1}, u_0, u_1, \cdots, u_{n-2}) \in C$. The additive quantum codes associated with the additive self-orthogonal cyclic codes over $GF(4)^n$ are called *additive cyclic quantum error-correcting codes*. The reason why we are interested in this kind of codes is simple and straight. On the one hand, the cyclic codes are the most studied of all codes in the classical error-correction theory, for they are easy to encode. The cyclic codes have provided us a great many good codes and, what is more, they are building blocks for many other codes [20]. They play an extremely important role in the classical error-correction theory. When we need some codes to meet our requirements, the cyclic codes will always be the choice with high priority. Inevitably, we wonder if the additive cyclic quantum codes also play an important role in the quantum error-correction theory. The table in [70] is an ensemble of good quantum codes which are found by different methods. It does not point out which method has higher priority than others. When we need a new quantum code, we would hardly know which method to select. We will have to try these methods one by one. In another word, we will be looking for a good code by chance. If we can show that the additive cyclic quantum codes are of great value, we will be able, at least, to offer researchers a worthy choice when they begin the search for good codes. On the other hand, the study of cyclic codes in the classical error-correction theory has accumulated a great deal of knowledge which is a prized treasure for our exploration of the additive self-orthogonal cyclic codes over $GF(4)$.

In this section, we will introduce a theorem from [70] and explain a new searching algorithm different from [70]. In Section 6.3, we will present the best additive cyclic quantum codes via an exhaustive search for the odd value of n ranging from 5 to 23.

At first, we explain some notations which we will use frequently. The *generator matrix* G of a self-orthogonal $(n,2^k)$ code C, is a $k \times n$ matrix, which consists of k independent vectors in C. The relation between G and the check matrix A of the stabilizer associated with C is $G = \Phi(A)$. For example, the generator matrix of an $(5,2^4)$ additive self-orthogonal code is:

$$\begin{bmatrix} 1 & w & w & 1 & 0 \\ 0 & 1 & w & w & 1 \\ 1 & 0 & 1 & w & w \\ w & 1 & 0 & 1 & w \end{bmatrix} \tag{6.10}$$

The corresponding additive quantum code is a $[[5,1,3]]$ code, the smallest quantum code which can correct one qubit error. It was independently discovered by [67] and [71]. The additive self-orthogonal cyclic codes have a more concise and convenient way to represent themselves than the generator matrices. We can use only one or two vectors to represent the entire generator matrix, for the rest of vectors of the generator matrix can be obtained by the right-shift operation of the given ones. We still take the $(5,2^4)$ code as an example. This code is a cyclic code, so we only need the vector

$(1, w, w, 1, 0)$ to represent the matrix (6.10). We will call such vectors the *generators of the cyclic code,* and write them in the form $\langle\ \rangle$, for example, $\langle 1ww10\rangle$. We can also represent a generator in the polynomial form. If $u = \langle u_0\ u_1 \cdots u_{n-1}\rangle$ is a generator of a cyclic code, the *polynomial representation* is defined as $p(x) = u_0 + u_1 x + \cdots + u_{n-1}x^{n-1}$. Then the generator $\langle 1ww10\rangle$ can be rewritten as $\langle 1 + wx + wx^2 + x^3\rangle$.

Now we introduce an important theorem [70] about the additive cyclic codes of $GF(4)^n$.

***Theorem 6.3*:**

a) Any $(n, 2^k)$ additive cyclic code C has two generators which can be represented as $\langle wp(x) + q(x), r(x)\rangle$, where $p(x), q(x), r(x)$ are binary polynomials, $p(x)$ and $r(x)$ divide $x^n - 1$ (mod 2), $r(x)$ divides $q(x)(x^n-1)\big/p(x)$ (mod 2), and $k = 2n - \deg p - \deg r$.

b) If $\langle wp'(x) + q'(x), r'(x)\rangle$ is another such representation, then $p'(x) = p(x), r'(x) = r(x)$ and $q'(x) \equiv q(x)$ (mod $r(x)$).

c) C is self-orthogonal if and only if

$$p(x)r(x^{n-1}) \equiv p(x^{n-1})r(x) \equiv 0 \quad (\text{mod } x^n - 1) \tag{6.11}$$

$$p(x)q(x^{n-1}) \equiv p(x^{n-1})q(x) \quad (\text{mod } x^n - 1) \tag{6.12}$$

The proof is given in [70]. The theorem makes it possible to search all of the additive cyclic self-orthogonal codes. It is necessary to define the search ranges of the polynomials $p(x), q(x), r(x)$ before we discuss the search algorithm.

1) The range for $p(x)$ is between 1 and x^n-1, not including x^n-1. $p(x)$ can not be 0, for if $p(x)$ is 0, the code C will be a binary code.

2) The range for $r(x)$ is between 1 and x^n-1, including x^n-1. When $r(x)$ is x^n-1, we can also find some valid and good $p(x), q(x)$, but $r(x)$ cannot be considered as a generator, for $r(x)$ is actually 0 (mod x^n-1). In this case, the generator of the code is simply $\langle wp(x)+q(x)\rangle$.

3) The range for $q(x)$ is between 1 and $r(x)$, including $r(x)$. Note that $q(x)=r(x)$ is equivalent to $q(x)=0$, for the generators $\langle wp(x)+r(x), r(x)\rangle$ and $\langle wp(x), r(x)\rangle$ generate the same code.

The search algorithm consists of two parts. At first, we have to find three polynomials $p(x), q(x), r(x)$ which meet the requirements of Theorem 6.3. Secondly, we need to find the parameters of the corresponding $[[n,k,d]]$ quantum code (Note that for the self-dual code C, the corresponding quantum code is an $[[n,0,d]]$ code, where d is equal to $d(C)$). Such undertakings are very time-consuming. Searching all of the polynomials with degrees less than n, and finding out d for each self-orthogonal cyclic code will occupy a great deal of computer resources and take a very long time to calculate. Great care, therefore, has to be taken to design an efficient algorithm.

An important problem needed to be solved is how to make the process of finding the appropriate polynomials as simple as possible. As both $r(x)$ and $p(x)$ divide x^n-1, they must be the factors of x^n-1. Therefore, we do not need to try all of the polynomials. If we can find all of the irreducible binary factors of $x^n - 1$, then $r(x)$ and $p(x)$ are just combinations of these factors. Thus, the process will be greatly simplified. Now, we begin to discuss the problem of factoring $x^n - 1$ over $GF(q)$. The following discussion is mainly from [20]. The *polynomials over* $GF(q)$ are the polynomials whose coefficients are all from $GF(q)$. We always assume that n and q are relatively prime, for example, if $q = 2$, then n should always be odd. There is a smallest integer m such that n divides $q^m - 1$. This m is called the *multiplicative order of* q *modulo* n. As the $q^m - 1$ roots of $x^{q^m-1} - 1$ form the non-zero elements in $GF(q^m)$ and the roots of $x^n - 1$ are also the roots of $x^{q^m-1} - 1$, all of the roots of $x^n - 1$, which are called n^{th} *roots of unity*, lie in the field $GF(q^m)$, and in no smaller field. $GF(q^m)$ is therefore called the *splitting field* of $x^n - 1$.

Suppose β is the primitive element of $GF(q^m)$ and $a = \frac{q^m - 1}{n}$, we define $\alpha \equiv \beta^a$ to be the *primitive* n^{th} *roots of unity*, for $\alpha^0, \alpha^1, \cdots, \alpha^{n-1}$ form the n distinct roots of $x^n - 1$:

$$x^n - 1 = \prod_{i=0}^{n-1} (x - \alpha^i) \tag{6.13}$$

Equation (6.13) factors $x^n - 1$ over $GF(q^m)$, for 1 and α^i are the elements in $GF(q^m)$. Let's move on to see how to factor $x^n - 1$ over $GF(q)$. There is no doubt that each irreducible polynomial over $GF(q)$ consists of one or several factors of $\{(x - \alpha^i)\}$. Then, by properly partitioning the factors $\{(x - \alpha^i)\}$ into different groups, we can make each group an irreducible polynomial over $GF(q)$. The method of partitioning is described below. For $s \in \{0,1,\cdots,n-1\}$, the *cyclotomic coset mod* n *over* $GF(q)$ which contains s is defined to be:

$$C_s = \{s, sq, sq^2, \cdots, sq^{m_s-1}\} \tag{6.14}$$

where $sq^{m_s} \equiv s \bmod n$. Thus the integers mod n are partitioned into cyclotomic cosets, i.e.,

$$\{0,1,\cdots,n-1\} = \bigcup_s C_s \tag{6.15}$$

where s runs through a set of *coset representatives mod n*. For example, for $n=9, q=2$,

$$C_0 = \{0\}, \quad C_1 = \{1,2,4,8,7,5\}, \quad C_3 = \{3,6\} \tag{6.16}$$

It has been proven [20] that a irreducible polynomial over $GF(q)$ containing the factor $(x-\alpha^s)$ is:

$$M^{(s)}(x) = \prod_{i \in C_s}(x-\alpha^i) \tag{6.17}$$

Thus,

$$x^n - 1 = \prod_s M^{(s)}(x) \tag{6.18}$$

Where s runs through a set of coset representatives mod n. For example, for $n=9, q=2$,

$$x^9 - 1 = M^{(0)}(x)M^{(1)}(x)M^{(3)}(x) \tag{6.19}$$

where

$$M^{(0)}(x) = x+1 \tag{6.20}$$

$$M^{(1)}(x) = x^6 + x^3 + 1 \tag{6.21}$$

$$M^{(3)}(x) = x^2 + x + 1 \tag{6.22}$$

Note that the degree of the irreducible polynomial $M^{(s)}(x)$ is equal to the number of elements of C_s and the number of the irreducible polynomials is equal to the number of cyclotomic cosets. Thus, by calculating the cyclotomic cosets mod n over $GF(2)$, we will know how many irreducible binary factors x^n+1 has and the degree of each factors. With these pieces of information, it is not difficult to find all of the irreducible binary factors of x^n+1.

Finding d is another problem we need to solve. It is not difficult to find the code C, and list all of its vectors, but finding its dual code has proven difficult. The key to the solution of this problem lies in the relationship between the *weight distributions* of a code C and its dual code $C^{\perp}$. The weight distribution of a code C is a sequence $A_0, A_1, \cdots A_n$, where A_j is the number of the vectors in C whose weights are j. The polynomial

$$W_C(x,y)=\sum_{j=0}^{n} A_j x^{n-j} y^j \tag{6.23}$$

is called the *weight enumerator*[20] of C. It is a surprising fact that the weight enumerator of $C^{\perp}$ is uniquely determined by the weight enumerator of C.

Theorem 6.4: If C is an $(n.2^k)$ additive code with weight enumerator $W_C(x,y)$, then the weight enumerator of $C^{\perp}$ is given by:

$$W_{C^{\perp}}(x,y)=2^{-k}W(x+3y,x-y) \tag{6.24}$$

This theorem follows from the general theory of additive codes developed by Delsarte [38]. Because the codes we are interested in are self-orthogonal codes, we can find the minimum distance of $C^{\perp}-C$ by comparing the coefficients of $W_C(x,y)$ with those of

$W_{C^{\perp}}(x,y)$.

With these two problems being solved, we can describe the search algorithm below.

1) Factor $x^n - 1$ over $GF(2)$, find all of the irreducible binary factors.

2) Consider all of the pairs of $p(x)$ and $r(x)$ which satisfy the equation

$$p(x)r(x^{n-1}) \equiv p(x^{n-1})r(x) \equiv 0 \qquad (\text{mod } x^n - 1)$$

3) For each pair of $p(x)$ and $r(x)$ coming from step 2), consider all of the possible $q(x)$ which satisfy

   a) $q(x)(x^n - 1) \equiv 0 \quad (\text{mod } p(x)r(x))$

   b) $p(x)q(x^{n-1}) \equiv p(x^{n-1})q(x) \qquad (\text{mod } x^n - 1)$

4) For each set of qualified polynomials $p(x), q(x), r(x)$, calculate the weight enumerators of the code and its dual code in order to find d.

6.3 Search results

In this section, we present the results from the search algorithm described above. We have made an exhaustive search, during which n began at 5 and ended at 23 (n is an odd number). In Table 6.1, we list all of the additive cyclic quantum codes with the highest minimum distance d. The search has shown that there are many additive cyclic quantum codes with the same parameters $[[n,k,d]]$. It is not necessary to list them all, so we only give one example for each $[[n,k,d]]$. Theorem 6.3 tells us that for additive cyclic quantum codes, there is a certain relationship between the parameters n and k: $k = 2n - \deg p - \deg r$, which indicates that there is a limitation on the value of k. To find all of the valid ("valid" means that an $[[n,k]]$ additive quantum code exists) k for n, we made another exhaustive search, during which n is from 5 to 31. In Table 6.2, we list all of the valid (n,k) pairs.

Table 6.1

Additive cyclic code with highest minimum distance

Parameters	Generators
$[[5,0,3]]$	$\langle w\ w\ 0\ 1\ 0\rangle\ \langle 1\ 1\ 1\ 1\ 1\rangle$
$[[5,1,3]]$	$\langle \bar{w}\ \bar{w}\ 1\ 0\ 1\rangle$
$[[5,4,1]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\rangle$
$[[7,0,3]]$	$\langle w\ w\ w\ 0\ w\ 0\ 0\rangle\ \langle 1\ 0\ 1\ 1\ 0\ 0\ 0\rangle$
$[[7,1,3]]$	$\langle \bar{w}\ \bar{w}\ 1\ 0\ 0\ 0\ 1\rangle$
$[[7,3,2]]$	$\langle \bar{w}\ w\ 0\ \bar{w}\ 0\ 1\ 1\rangle$
$[[7,4,2]]$	$\langle \bar{w}\ 1\ w\ w\ \bar{w}\ 0\ 1\rangle$
$[[7,6,1]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\rangle$
$[[9,0,4]]$	$\langle \bar{w}\ \bar{w}\ w\ 1\ 0\ 1\ 0\ 0\ 0\rangle\ \langle 1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 0\rangle$
$[[9,1,3]]$	$\langle \bar{w}\ \bar{w}\ 1\ 0\ 0\ 0\ 0\ 0\ 1\rangle$
$[[9,2,3]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ 1\ 0\ 1\ 1\ 0\ 1\rangle$
$[[9,3,3]]$	$\langle \bar{w}\ 0\ 0\ \bar{w}\ 1\ 1\ 0\ 1\ 1\rangle$
$[[9,6,2]]$	$\langle \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\rangle$
$[[9,7,1]]$	$\langle \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\rangle$
$[[9,8,1]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\rangle$

[[11,0,4]]	$\langle \bar{w} \bar{w} 0 0 0 1 0 1 0 0 0 \rangle$ $\langle 1 1 1 1 1 1 1 1 1 1 1 \rangle$
[[11,1,3]]	$\langle \bar{w} \bar{w} 1 0 0 0 0 0 0 0 1 \rangle$
[[11,10,1]]	$\langle \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \rangle$
[[13,0,5]]	$\langle w w 0 0 1 0 1 1 1 0 1 0 0 \rangle$ $\langle 1 1 1 1 1 1 1 1 1 1 1 1 1 \rangle$
[[13,1,5]]	$\langle \bar{w} \bar{w} 1 0 0 1 1 0 1 1 0 0 1 \rangle$
[[13,12,1]]	$\langle \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \bar{w} \rangle$
[[15,0,6]]	$\langle w 1 w \bar{w} w 1 w 0 0 0 0 0 0 0 0 \rangle$ $\langle 1 0 1 1 0 0 1 1 0 1 0 0 0 0 0 \rangle$
[[15,1,5]]	$\langle \bar{w} \bar{w} 1 0 0 1 1 0 0 0 1 1 0 0 1 \rangle$
[[15,2,5]]	$\langle \bar{w} \bar{w} \bar{w} 1 1 0 1 1 0 0 1 1 0 1 1 \rangle$
[[15,3,5]]	$\langle \bar{w} 0 0 \bar{w} 0 1 1 0 1 0 1 0 1 1 0 \rangle$
[[15,4,4]]	$\langle \bar{w} w 1 0 w 1 0 0 1 1 1 0 1 1 1 \rangle$
[[15,5,4]]	$\langle \bar{w} 0 w 1 w \bar{w} 0 0 1 1 1 1 0 0 1 \rangle$
[[15,6,4]]	$\langle \bar{w} 0 0 \bar{w} \bar{w} w \bar{w} 1 0 1 0 1 0 0 1 \rangle$
[[15,7,3]]	$\langle \bar{w} \bar{w} 0 w 0 0 1 w 0 1 0 1 1 1 1 \rangle$
[[15,8,3]]	$\langle \bar{w} 1 1 1 w 0 \bar{w} \bar{w} w 0 0 1 1 0 1 \rangle$
[[15,9,3]]	$\langle \bar{w} w 0 0 \bar{w} w w 1 1 w 1 0 1 1 1 \rangle$
[[15,10,2]]	$\langle \bar{w} 1 1 1 1 \bar{w} 1 1 1 1 \bar{w} 1 1 1 1 \rangle$
[[15,11,2]]	$\langle \bar{w} \bar{w} 1 0 1 \bar{w} \bar{w} 1 0 1 \bar{w} \bar{w} 1 0 1 \rangle$
[[15,12,2]]	$\langle \bar{w} 1 1 \bar{w} 1 1 \bar{w} 1 1 \bar{w} 1 1 \bar{w} 1 1 \rangle$
[[15,13,1]]	$\langle \bar{w} \bar{w} 0 \bar{w} \bar{w} 0 \bar{w} \bar{w} 0 \bar{w} \bar{w} 0 \bar{w} \bar{w} 0 \rangle$

$[[15,14,1]]$	$\langle \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w} \rangle$
$[[17,0,7]]$	$\langle w\ w\ 0\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 0\ 0 \rangle$ $\langle 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \rangle$
$[[17,1,7]]$	$\langle \overline{w}\ \overline{w}\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1 \rangle$
$[[17,8,4]]$	$\langle \overline{w}\ 1\ 0\ \overline{w}\ w\ \overline{w}\ 0\ 1\ \overline{w}\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1 \rangle$
$[[17,9,4]]$	$\langle \overline{w}\ \overline{w}\ 1\ w\ 1\ 1\ w\ 1\ \overline{w}\ \overline{w}\ 0\ 0\ 1\ 0\ 1\ 0\ 0 \rangle$
$[[17,16,1]]$	$\langle \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w} \rangle$
$[[19,0,7]]$	$\langle w\ w\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0 \rangle$ $\langle 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \rangle$
$[[19,1,7]]$	$\langle \overline{w}\ \overline{w}\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1 \rangle$
$[[19,18,1]]$	$\langle \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w}\ \overline{w} \rangle$
$[[21,0,8]]$	$\langle w\ w\ 0\ 1\ 0\ \overline{w}\ 0\ 0\ 1\ 1\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \rangle$, $\langle 1\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0 \rangle$
$[[21,1,7]]$	$\langle \overline{w}\ \overline{w}\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1 \rangle$
$[[21,2,6]]$	$\langle \overline{w}\ w\ \overline{w}\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 1 \rangle$
$[[21,3,6]]$	$\langle \overline{w}\ w\ 0\ w\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 1 \rangle$
$[[21,4,6]]$	$\langle \overline{w}\ 0\ w\ \overline{w}\ w\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 1 \rangle$
$[[21,5,6]]$	$\langle \overline{w}\ 1\ 1\ 0\ w\ \overline{w}\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1 \rangle$
$[[21,6,5]]$	$\langle \overline{w}\ w\ w\ 0\ \overline{w}\ 1\ \overline{w}\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1 \rangle$
$[[21,7,5]]$	$\langle \overline{w}\ 0\ 0\ w\ w\ \overline{w}\ \overline{w}\ \overline{w}\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1 \rangle$
$[[21,8,4]]$	$\langle \overline{w}\ 0\ w\ 1\ 0\ \overline{w}\ 0\ w\ w\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1 \rangle$
$[[21,9,4]]$	$\langle \overline{w}\ 0\ 0\ 0\ 1\ 1\ w\ 1\ 1\ w\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1 \rangle$

$[[21,10,4]]$	$\langle \bar{w}\ 0\ w\ 1\ \bar{w}\ 0\ w\ w\ 0\ 1\ w\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 1\ 1 \rangle$
$[[21,11,4]]$	$\langle \bar{w}\ 0\ 1\ w\ w\ 0\ 1\ 1\ 1\ w\ 0\ w\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1 \rangle$
$[[21,12,3]]$	$\langle \bar{w}\ 1\ w\ \bar{w}\ 1\ 1\ w\ 0\ \bar{w}\ 0\ 0\ w\ w\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1 \rangle$
$[[21,13,3]]$	$\langle \bar{w}\ \bar{w}\ w\ 0\ w\ 1\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ 0\ w\ 0\ w\ 0\ 0\ 1\ 0\ 0\ 1\ 1 \rangle$
$[[21,14,3]]$	$\langle \bar{w}\ w\ 0\ 0\ 0\ \bar{w}\ w\ w\ 1\ w\ \bar{w}\ w\ 1\ 0\ w\ 0\ 1\ 0\ 1\ 1\ 1 \rangle$
$[[21,15,3]]$	$\langle \bar{w}\ 1\ \bar{w}\ 1\ 1\ w\ 0\ 1\ \bar{w}\ \bar{w}\ 0\ 0\ w\ 1\ w\ w\ 0\ 1\ 1\ 0\ 1 \rangle$
$[[21,16,2]]$	$\langle \bar{w}\ \bar{w}\ w\ \bar{w}\ w\ 1\ w\ 0\ \bar{w}\ 1\ 0\ w\ w\ 1\ 0\ 0\ w\ 0\ 1\ 1\ 1 \rangle$
$[[21,17,2]]$	$\langle \bar{w}\ w\ 0\ \bar{w}\ 0\ 1\ 1\ \bar{w}\ w\ 0\ \bar{w}\ 0\ 1\ 1\ \bar{w}\ w\ 0\ \bar{w}\ 0\ 1\ 1 \rangle$
$[[21,18,2]]$	$\langle \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1\ \bar{w}\ 1\ 1 \rangle$
$[[21,19,1]]$	$\langle \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0\ \bar{w}\ \bar{w}\ 0 \rangle$
$[[21,20,1]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w} \rangle$
$[[23,0,8]]$	$\langle w\ w\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0 \rangle$, $\langle 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \rangle$
$[[23,1,7]]$	$\langle w\ w\ w\ w\ w\ 0\ 0\ w\ 0\ 0\ w\ 0\ w\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \rangle$, $\langle 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \rangle$
$[[23,11,4]]$	$\langle w\ w\ \bar{w}\ w\ w\ 0\ 0\ \bar{w}\ 0\ 1\ \bar{w}\ 0\ \bar{w}\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \rangle$, $\langle 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \rangle$
$[[23,12,4]]$	$\langle w\ \bar{w}\ w\ w\ \bar{w}\ 1\ 0\ \bar{w}\ 0\ 1\ w\ 0\ \bar{w}\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0 \rangle$
$[[23,22,1]]$	$\langle \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w}\ \bar{w} \rangle$

Table 6.2

All of the valid (n,k) for additive cyclic codes. (“E” means exist)

$n\backslash k$	0	1	2	3	4	5	6	7
5	E	E			E			
7	E	E		E	E		E	
9	E	E	E	E			E	E
11	E	E						
13	E	E						
15	E	E	E	E	E	E	E	E
17	E	E						
19	E	E						
21	E	E	E	E	E	E	E	E
23	E	E						
25	E	E			E	E		
27	E	E	E	E			E	E
29	E	E						
31	E	E				E	E	

Table 6.2 (continued)

All of the valid (n,k) for additive cyclic codes("E" means exist)

$n\backslash k$	8	9	10	11	12	13	14	15
9	E							
11			E					
13					E			
15	E	E	E	E	E	E	E	
17	E	E						
19								
21	E	E	E	E	E	E	E	E
23				E	E			
25								
27	E	E						
29								
31			E	E				E

Table 6.2 (continued)

All of the valid (n,k) for additive cyclic codes("E" means exist)

$n\backslash k$	16	17	18	19	20	21	22	23
17	E							
19			E					
21	E	E	E	E	E			
23							E	
25					E	E		
27			E	E	E	E		
29								
31	E				E	E		

Table 6.2 (continued)

All of the valid (n,k) for additive cyclic codes("E" means exist)

$n\backslash k$	24	25	26	27	28	29	30
25	E						
27	E	E	E				
29					E		
31		E	E				E

6.4 Analysis of the search results

In the last section, we showed our search results. There is something hiding behind these data, and in this section, we will discuss these results and the algorithm we used.

At first, from Tables 6.1 and 6.2, we can see that for some values of n we have many additive cyclic quantum codes, while for some we only have few. For example, for $n=21$, there are so many additive cyclic quantum codes, but for $n=11,13$, there are only three codes. The reason for this difference is the number of the irreducible factors of x^n-1. The more irreducible factors a number n has, the more additive cyclic quantum codes it can provide. The number 11 has two cyclotomic cosets mod 2,

$$C_0=\{0\}$$

$$C_1=\{1,2,3,4,5,6,7,8,9,10\}$$

while the number 21 has six cyclotomic cosets mod 2,

$$C_0=\{0\}$$

$$C_1=\{1,2,4,8,11,16\}$$

$$C_3=\{3,6,12\}$$

$C_5 = \{5,10,13,17,19,20\}$

$C_7 = \{7,14\}$

$C_9 = \{9,15,18\}$

Obviously, $x^{21}-1$ has more irreducible factors than $x^{11}-1$ has. Then the polynomials $r(x), p(x)$ of $x^{21}-1$ have more choices than the polynomials of $x^{11}-1$ have. Thus, it is not surprising that $x^{21}-1$ provides us with so many additive cyclic quantum codes, while $x^{11}-1$ only gives us three.

Second, for any odd number n, there are always $[[n,0]],[[n,1]],[[n,n-1]]$ additive cyclic quantum codes for the simple reason that x^n-1 always has the factors $x-1$ and $\sum_{i=0}^{n-1} x^i$. If we let $r(x)=q(x)=x^n-1, p(x)=1$, we get the $[[n,0]]$ additive cyclic quantum code, if we let $r(x)=q(x)=x^n-1, p(x)=x-1$, we get the $[[n,1]]$ additive cyclic quantum code, if we let $r(x)=q(x)=x^n-1, p(x)=\sum_{i=0}^{n-1} x^i$, we get the $[[n,n-1]]$ additive cyclic quantum code. Of course, the codes generated by these polynomials perhaps are not the best additive cyclic codes, but they prove the existence of the $[[n,0]],[[n,1]],[[n,n-1]]$ codes.

Third, although it is a good thing that the algorithm can find all of the additive cyclic quantum codes, it takes a really long time to search them as n becomes larger and larger. Suppose that we are searching an $[[n,k]]$ code, in order to find the minimum distance d, we have to list all of the 2^{n-k} codewords to find the weight enumerator. Perhaps there are hundreds or thousands of $[[n,k]]$ codes, and we have to list 2^{n-k} codewords for hundreds or thousands times. If $n=31,k=1$, the calculation will be thousands of times of 2^{30}. It is not surprising that the calculation time will increase exponentially. However, if k is large, the calculation time is still acceptable, for 2^{n-k} is not a too big deal for a fast computer. A partial search for $n=31$ has found the following additive cyclic quantum codes:

Table 6.3

Some additive cyclic code with highest minimum distance for $n=31$

Parameters	Generators
$[[31,15,5]]$	$\langle \bar{w}\, w\, \bar{w}\, w\, 0\, 0\, \bar{w}\, 1\, w\, \bar{w}\, 1\, w\, 0\, \bar{w}\, 1\, 1\, w\, 0\, 0\, \bar{w}\, 1\, w\, 0\, 0\, 0\, 0\, 0\, 0\, 0\, 0\, 0 \rangle$, $\langle 1\, 1\, 0\, 0\, 1\, 0\, 1\, 1\, 0\, 1\, 1\, 1\, 1\, 0\, 1\, 0\, 1\, 0\, 0\, 0\, 1\, 0\, 0\, 1\, 1\, 1\, 0\, 0\, 0\, 0\, 0 \rangle$
$[[31,16,5]]$	$\langle w\, \bar{w}\, \bar{w}\, \bar{w}\, 1\, 0\, w\, 1\, w\, \bar{w}\, 1\, w\, 1\, w\, 1\, 0\, w\, 1\, 0\, w\, 1\, w\, 1\, 0\, 0\, 0\, 0\, 0\, 0\, 0\, 0 \rangle$, $\langle 1\, 0\, 1\, 0\, 1\, 1\, 1\, 0\, 1\, 1\, 0\, 0\, 0\, 1\, 1\, 1\, 1\, 1\, 0\, 0\, 1\, 1\, 0\, 1\, 0\, 0\, 1\, 0\, 0\, 0\, 0 \rangle$
$[[31,20,4]]$	$\langle \bar{w}\, 1\, 0\, 0\, 1\, w\, \bar{w}\, 0\, 1\, \bar{w}\, 0\, \bar{w}\, w\, 0\, 1\, \bar{w}\, w\, \bar{w}\, 0\, \bar{w}\, w\, 0\, 0\, 0\, 1\, 1\, 0\, 1\, 0\, 0\, 1 \rangle$
$[[31,21,4]]$	$\langle \bar{w}\, w\, 0\, 0\, 1\, \bar{w}\, 0\, \bar{w}\, 1\, w\, w\, w\, 1\, \bar{w}\, 1\, w\, 0\, 1\, w\, \bar{w}\, 1\, \bar{w}\, 1\, 0\, 1\, 1\, 1\, 1\, 1\, 1\, 1 \rangle$

These codes are not mentioned in [70].

6.5 Conclusion

To see the significance of our results, we compare Tables 6.1, 6.2, 6.3 with the table of [70] closely. We can see the following two facts:

1) All of the codes listed in Table 6.1, except the codes $[[11,0,4]],[[11,1,3]]$, meet the lower bounds in the table of [70].

2) Additive cyclic quantum codes offer us a great many quantum codes. For each odd value of n, we can always find some additive cyclic codes. By comparison, other

methods in [70] are not so resourceful and flexible. Many codes in the table of [70] which were found by different methods are now unified under one method.

These facts are very exciting, for they have proved the success of the search algorithm and the great value of additive cyclic quantum codes. The lower bounds in the table of [70] are the best codes found until now, while the upper bounds came from the theoretical calculation, so such codes may exist or may not exist. The fact that all of the best additive cyclic quantum codes we found, except $[[11,0,4]],[[11,1,3]]$, meet the lower bounds means that, in most cases, the best additive cyclic quantum codes are also the best quantum codes we can obtain. It is a very good thing, because, to a great degree, searching for the best quantum codes can be replaced by searching for the best additive cyclic quantum codes. The search complexity will therefore be greatly reduced.

Thus, we draw the conclusion that just as the cyclic codes play an extremely important role in the classical error-correcting codes, the additive cyclic quantum codes are also an important class of the quantum error-correcting codes and are worth great attention.

Chapter 7

Contributions and suggestions

Quantum information theory is a new field which has been developed during the last decade. Although this theory is far from mature for practical applications, it is developing very fast and appeals to many scientists, for it showed many striking properties which are greatly different from the classical information theory. It may be possible that the successful applications of quantum information theory will, one day, trigger another industrious revolution. However, due to its physical nature, this field seems closed to most electrical engineers. Acting on the great interest in quantum information theory, we entered into this field, and did some research in the area of quantum error-correcting codes. We hope that this thesis serves not only as a good introduction of quantum information theory to electrical engineers, but a valuable exploration of the quantum error-correcting codes as well. As there are too many areas in quantum information theory and the space of the thesis is limited, we only included some primary concepts and theories relevant to our research in the thesis. In a word, the thesis has the following contributions:

A) An introduction of quantum information theory:

1) Mathematical foundation of quantum mechanics: linear algebra;
2) The prime principles of quantum mechanics;
3) Quantum operations and quantum noise;
4) Quantum error-correction theory and the stabilizer codes;
5) The connection between the additive quantum error-correcting codes and the classical error-correcting codes.

B) The exploration of additive cyclic quantum error-correcting codes:

1) A new algorithm has been designed to search additive cyclic quantum error-correcting codes;

2) All of the best additive cyclic quantum error-correcting codes, which range from $n = 5$ to $n = 23$, have been found; we also gave a list of pairs of (n, k), $n \leq 31$ for which an $[[n, k]]$ additive cyclic quantum code exists. In addition, we found some best additive cyclic quantum error-correcting codes for $n = 31$: $[[31,15,5]]$, $[[31,16,5]]$, $[[31,20,4]]$, $[[31,21,4]]$. Most of the additive cyclic codes showed in this thesis are not presented in Calderbank's paper [70];

3) The great value of the additive cyclic quantum error-correcting codes was proven from two aspects:

   a) Most of the best additive cyclic quantum error-correcting codes are also the best additive quantum error-correcting codes obtained up to now. In the future, we can search for the best additive quantum codes by only searching for the best additive cyclic quantum codes. Thus, the search complexity will be greatly reduced.

   b) The additive cyclic quantum error-correcting codes offer us a large number of good additive quantum codes, which means that this class of additive quantum codes is very resourceful.

Although we made some good progress in our research in the additive cyclic quantum error-correcting codes, there are many things needed to be done in the future:

1) When n becomes large, say, around 31, the search will become difficult for small k. We wonder whether there is any thing we can do to improve the search algorithm so that we can finish such undertakings faster;

2) The classical cyclic codes are easily encoded, so we are also interested in the encoding circuit for the additive cyclic quantum error-correcting codes: is it also simple?
3) The classical cyclic codes are the building blocks for many other codes, so is it possible that we can build some additive quantum error-correcting codes from additive cyclic quantum codes?

Finally, we hope that this thesis can serve its functions well and attract more people to this new and fantastic field.

Bibliography

[1]. J. Bardeen, L. N. Cooper, and J. R. Schreiffer, "*Theory of Superconductivity*", *Physical Review*, Vol 108, Issue 5, pp. 1175-1204 (1957).

[2]. Nobel lectures of Chu, Cohen-Tannoudji, and Phillips, Rev. Mod. Phys. 70, 685-742 (1998).

[3]. C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[4]. B. Schumacher, "*Quantum coding*," *Phys. Rev. A,* vol. 51, pp. 2738–2747, 1995.

[5]. W. K. Wooters and W. H. Zurek, "*A single quantum can not be cloned*," Nature 299, 802(1982); D. Dieks, "Communication by EPR devices," Phys, Lett. A92, 271 (1982).

[6]. P. W. Shor, "*Scheme for reducing decoherence in quantum memory*," Phys, Rev. A52, 2493 (1995).

[7]. A. R. Calderbank and Peter W. Shor. "*Good quantum error-correcting codes exist*". arXive quant-ph/9512032.

[8]. Andrew Steane, "*Multiple Particle Interference and Quantum Error Correction*", arXiv: quant-ph/9601029, 1996.

[9]. A. R. Calderbank, E. M. Rains, P. W. Shor, and N. N. A. Sloane. "*Quantum Error Correction and Orthogonal Geometry*". arXive quant-ph/9605005.

[10]. D. Gottesman. "*Stabilizer Codes and Quantum Error Correction*". Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997.

[11]. C.H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W.K. Wootters, "*Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*", *Phys. Rev. Lett.* **70** (1993) 1895-1899.

[12]. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger, Experimental Quantum Teleportation, *Nature* **390** (1997) 575.

[13]. Charles H. Bennett, Gilles Brassard, and Artur K. Ekert in *Scientific American*, Vol.

267, No. 4, pages 50---57; October 1992.

[14]. C. H. Bennett and S. J. Wiesner. "*Communication via one- and two-particles operators on Einstein-Podolsky-Rosen states*". *Phys. Rev. Lett.*, 69(20):2881-2884, 1992.

[15]. A. Pais. *Subtle is the Lord: The Science and the Life of Albert Einstein.* Oxford University Press, Oxford, 1982.

[16]. A. Pais. *Inward Bound: Of Matter and Forces in the Physical World.* Oxford University Press, Oxford, 1986.

[17]. A. Pais. *Niels Bohr's Times: In Physics, Philosophy, and Polity.* Oxford University Press, Oxford, 1991.

[18]. G. J. Milburn. *Schrödinger's Machines: the Quantum Technology Reshaping Everyday Life.* W. H. Freeman, New York, 1997.

[19]. G. J. Milburn. *The Feynman Processor: Quantum Entanglement and the Computing Revolution.* Perseus Books, Reading, Mass., 1998.

[20]. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (NorthHolland, Amsterdam, 1977).

[21]. T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley and Sons, New York, 1991.

[22]. N. J. A. Sloane and A. D. Wyner, editors. *Claude Elwood Shannon: Collected Papers.* IEEE Press, New York, 1993.

[23]. D. Slepian, editor. *Keys Papers in the Development of Information Theory.* IEEE Press, New York, 1974.

[24]. R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge. 1985.

[25]. R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge. 1991.

[26]. P. R. Halmos. *Finite-dimensional Vector Spaces.* Van Nostrand, Princeton, N. J.,

1958.

[27]. G Strang. *Linear Algebra and Its Applications.* Academic Press, New York, 1976.

[28]. R. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[29]. R. Jozsa and B. Schumacher, "*A new proof of the quantum noiseless coding theorem*," *J. Modern Optics,* vol. 41, pp. 2343-2349, 1994.

[30]. J. I. Cirac and P. Zoller, "*Quantum computations with cold trapped ions*," Phys. Rev. Lett. 74, 4091 (1995).

[31]. C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, "*Demonstration of a fundamental quantum logic gate*," Phys. Rev. Lett. 75, 4714 (1995).

[32]. Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble, "*Measurement of conditional phase shifts for quantum logic*," Phys. Rev. Lett. 75, 4710 (1995).

[33]. N. Gershenfeld and I. Chuang, "*Bulk spin resonance quantum computation*," Science 275, 350 (1997).

[34]. Nielsen, M. and Chuang, I. "*Quantum Computation and Quantum Information*". Cambridge, England: Cambridge University Press, 2000.

[35]. Schrodinger, E , "*The Relation between the Quantum Mechanics of Heisenberg, Born and Jordan and that of Schrodinger*" . Ann. Phys. 1926, *79*, 734-756. {in German}

[36]. A. Einstein, B. Podolsky, and N. Rosen. "*Can quantum-mechanical description of physical reality be considered complete*?" *Phys. Rev.*, 47:777-780, 1935.

[37]. J. S. Bell. On the Einstein-Podolsy-Rosen paradox. *Physics*, 1:195-200, 1964. Reprinted in J. S. Bell, "*Speakable and Unspeakable in Quantum Mechanics*", Cambridge University Press, Cambridge, 1987.

[38]. P. Delsarte, *"Bounds for unstricted codes, by linear programming," Philips Res. Rep.*, vol. 27, pp. 272-289, 1972.

[39]. J. von Neumann. *Gottinger Nachrichten*, page 245, 1927.

[40]. L. P. Hughston, R. Jozsa, and W. K. Wooters, Physics Letters A 183, 14 (1993).

[41]. E. Schrodinger. "*Probability relations between separated systems*". *Proc. Cambridge Philos. Soc.*, 32:446-452, 1936.

[42]. E. T. Jaynes. *Information theory and statistical mechanics*. ii. *Phys. Rev.*, 108(2):171-190, 1957.

[43]. E. Schmidt. Zur theorie der linearen und nichtlinearen integralgleighungen. *Math. Annalen.*, 63:433-476, 1906.

[44]. A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic, Dordrecht, 1993.

[45]. J. J. Sakurai. *Modern Quantum mechanics*. Addison-Welsey, Reading, Mass., 1995.

[46]. R. P. Feynman, R. B. Leighton, and M. Sands. *Volume III of The Feynman Lectures on Physics*. Addison-Wesley, Reading, Mass., 1965.

[47]. C. Cohen-Tannoudji, B. Diu and F. Laloe. *Quantum Mechanics,* Vol. I. John Wiley and Sons, New York, 1977.

[48]. C. Cohen-Tannoudji, B. Diu and F. Laloe. *Quantum Mechanics,* Vol. II. John Wiley and Sons, New York, 1977.

[49]. H.-K. Lo and and T. Spiller S. Popescu (editors). *Introduction to quantum computation and information*. World Scienti_c, Singapore (1998).

[50]. J. Gruska. *Quantum computing.* McGraw-Hill, New York (1999).

[51]. D. Bouwmeester, A. K. Ekert and A. Zeilinger (editors). *The physics of quantum information: Quantum cryptography, quantum teleportation, quantum computation.* Springer, Berlin (2000).

[52]. G. Alber, T. Beth, M. Horodecki, R. Horodecki, M. Rotteler, H. Weinfurter, R. Werner and A. Zeilinger (editors). *Quantum information.* Springer, Berlin (2001).

[53]. J. Preskill. Lecture notes for the course *`information for physics 219/computer science 219, quantum computation*'. Caltech, Pasadena, California (1999). www.theory.caltech.edu/people/preskill/ph229.

[54]. A. Cabello. *Bibliographic guide to the foundations of quantum mechanics and quantum information*. quant-ph/0012089 (2000).

[55]. K.-E. Hellwig and K. Kraus, Comm. Math. Phys. 11, 214 (1969).

[56]. K. Hellwig and K. Krauss, "*Operations and measurements II*," *Communications in Mathematical Physics,* vol. 16, pp. 142-147, 1970.

[57]. K. Kraus, States, Effects, and Operations: Fundamental Notions of Quantum Theory (Springer, Berlin, 1983).

[58]. B. Schumacher, Phys. Rev. A 54, 2615 (1996).

[59]. M.-D. Choi, *Linear Algebra and Its Applications* 10, 285 (1975).

[60]. Carlton M. Caves, "*Quantum Error Correction and Reversible Operations",* arXiv:quant-ph/9811082 v1 30 Nov 1998.

[61]. Benjamin Schumacher, "*Sending entanglement through noisy quantum channels*", arXiv:quant-ph/9604023 v1 22 Apr 1996.

[62]. E. B. Davies. *Quantum Theory of Open Systems.* Academic Press, London, 1976.

[63]. C. W. Gardiner. *Quantum Noise.* Springer-Verlag, Berlin, 1991.

[64]. A. M. Steane. *Error correcting codes in quantum theory. Phys. Rev. Lett.,* 77:793, 1996

[65]. Artur Ekert, Chiara Macchiavello, *Quantum error correction for communication.* arXiv: quant-ph/9602022.

[66]. Emanuel Knill, Raymond Laflamme, *A theory of Quantum Error-Correcting Codes. Phys. Rev. A, 55:900, 1997. arXive e-print quant-ph/9604034*

[67]. C. H. Bennett, D. P. Divincenzo, J. A. Smolin, and W. K. Wotters. *Mixed state entanglement and quantum error correction. Phys. Rev. A, 54:3824, 1996. arXive*

e-print quant-ph/9604024

[68]. See for example the appendix of E. Knill, *Approximation by Quantum Circuits*, Los Alamos National Laboratory preprint LA-UR-95-2225.

[69]. A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel, "Z4 Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets," *Proc. London Math. Soc.*

[70]. A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "*Quantum error-correction via codes over* $GF(4)$," quant-ph/9608026 (1996).

[71]. R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, "*Perfect quantum error correction code,"* *Phys. Rev. Lett.*, vol. 77, pp. 198-201, 1996; also LANL e-print quant-ph/9602019.