Lecture Notes on
Quantum Information and Computation

Yong Zhang

School of Physics and Technology, Wuhan University
(Fall 2015)

Abstract
These lectures notes are written for both advanced undergraduate students and first-year graduate students in the School of Physics and Technology, University Wuhan. They are mainly based on both online lecture notes of John Preskill from Caltech and the standard textbook of Michael Nielsen and Issac Chuang, so I do not claim any originality. These notes certainly have all kinds of typos or errors, so they will be updated from time to time. I do take the full responsibility for all kinds of typos or errors (besides errors in English writing), and please let me know of them.

* The second version of these notes are typeset by
  Yi Peng
  Undergraduate student (Id: 2010301020027),
  Participant in the Fall semester, 2013.

* The third version of these notes are typeset by
  Kun Zhang
  Graduate student (Id: 2013202020002),
  Participant in the Fall semester, 2012-2013-2014.

1 yong_zhang@whu.edu.cn
2 pengyi_py@whu.edu.cn
3 kun_zhang@whu.edu.cn
Acknowledgements

I thank all participants in class including advanced undergraduate students, first-year graduate students and French students for their patience and persistence and for their various enlightening questions. I especially thank students who are willing to devote their precious time to the typewriting of these lecture notes in Latex.

Main References to Lecture Notes


Main References to Homeworks


Research Projects

* See Yong Zhang’s English and Chinese homepages.
To our parents and our teachers!

To be the best researcher is to be the best person first of all:
   Respect and listen to our parents and our teachers always!
This course focuses on fundamental principles of quantum mechanics.

The aim of this course is to study the simulation of both quantum field theory and quantum gravity on quantum computer.
# Contents

## I Introduction to Quantum Information and Computation

1 Overview

1.1 Reasons to learn Quantum Information and Computation ......................... 1
1.2 What’s Quantum Information and Computation? ................................. 1
1.3 Research topics ........................................ 2

2 Thermodynamics and Statistical Mechanics ..................................... 4

3 Quantum Mechanics (I): Axioms ........................................... 5

3.1 Axioms of quantum mechanics for closed system .............................. 5
3.1.1 State ........................................ 5
3.1.2 Observable ..................................... 6
3.1.3 Projective measurement ................................ 7
3.1.4 Schrödinger equation .................................. 8
3.1.5 Composite system ................................... 9

4 Qubit, Quantum Gates and Bell States ...................................... 10

4.1 Overview ........................................................................ 10
4.2 Pure state formalism of a qubit ........................................... 11
4.2.1 Single-qubit gate $SU(2)$ in the spin-1/2 case ............................ 12
4.2.1.1 Spin-1/2 operator and Pauli matrices .......................... 12
4.2.1.2 Spinor representation of $SU(2)$ group ....................... 12
4.2.2 Properties ........................................ 13
4.2.3 Physical realization of qubit ...................................... 16
4.3 Bell states ..................................................................... 16
4.3.1 Notation ........................................ 17
4.3.2 Parity-bit $(i)$ and Phase-bit $(j)$ .................................. 21
4.3.3 Orthonormal basis of the two-qubit Hilbert space ................... 22
4.3.4 How to distinguish (create) $|ψ(i,j)\rangle$. ........................... 23
4.4 Quantum circuit model of Bell states .................................... 25
4.4.1 Quantum circuit model ........................................ 25
4.4.2 Hadamard gate: $H$ ........................................ 25
4.4.3 CNOT gate ........................................ 28
4.4.4 Quantum circuit model of Bell states ............................... 32

5 No-Cloning, Dense Coding, Teleportation and Cryptography ............ 35

5.1 No-cloning theorem .................................................. 35
5.2 Dense coding ...................................................... 38
5.3 Quantum teleportation ................................................ 40
10.3.2 Classical algorithm ........................................ 96
10.3.3 Deutsch’s problem ......................................... 96
10.3.4 Phase kick-back ........................................... 96
10.3.5 Deutsch’s algorithm ........................................ 98
10.4 Deutsch-Jozsa’s algorithm ................................. 99
10.4.1 Constant and balanced function in n-qubit .......... 99
10.4.2 Constant or balanced function? ......................... 99
10.4.3 Notation and lemma ..................................... 100
10.4.4 Deutsch-Jozsa’s algorithm ............................... 101
10.5 Bernstein-Vazirani’s algorithm ......................... 103
10.6 Simon’s algorithm .......................................... 104
10.7 Grover’s algorithm ......................................... 104
10.7.1 Overview of the problem ................................. 105
10.7.2 Grover’s algorithm ....................................... 105
10.7.3 Example: \( N = 4 \) ..................................... 108
10.7.4 Quantum circuit model of Grover’s algorithm .... 110

11 Quantum Circuit Complexity ............................... 113
11.1 Circuit complexity ........................................... 113
11.1.1 Definitions .................................................. 113
11.1.2 Complexity class ......................................... 114
11.1.3 Quantum complexity ..................................... 114
11.1.4 Accuracy .................................................... 115

12 Quantum Simulation ............................................ 116

III Density Matrix and Quantum Entanglement ............. 117
13 Quantum Mechanics (II): Density Matrix ................. 118
13.1 Density matrix as state of quantum open system .... 118
13.1.1 State ensemble formalism of density matrix ........ 119
13.1.2 Operator formalism of density matrix ............... 121
13.1.3 Reduced density matrix (State for subsystem) .... 122
13.2 Mixed state formalism of a qubit ....................... 124
13.2.1 Why polarization vector? ............................... 125
13.2.2 Pure state and mixed state in two-dimensional Hilbert space \( \mathcal{H}_2 \) .... 126
13.3 Convexity of density matrix ............................... 128
13.4 Two-qubit system and its subsystem .................... 129
13.4.1 Example: EPR pair (Bell state) ......................... 129
13.4.2 Maximally entangled two-qubit pure states ......... 130
13.4.3 Monogamy of maximal entanglement ................. 131

14 Schmidt Decomposition, Purification and GHJW Theorem 133
14.1 Introduction .................................................. 133
14.2 Schmidt decomposition and quantum entanglement .... 133
14.2.1 Schmidt decomposition ................................. 133
14.2.2 Quantum entanglement ................................ 134
14.2.3 Proof for the theorem of the Schmidt decomposition 135
14.3 Example for the Schmidt decomposition .................. 136
Part I

Introduction to Quantum Information and Computation
Chapter 1

Overview

References:

- [Preskill] Chapter 1: Introduction and overview;

1.1 Reasons to learn Quantum Information and Computation

For an advanced undergraduate major in physics, he or she has to learn Quantum Information and Computation, because

- Quantum Information and Computation can be seen as a new type of advanced Quantum Mechanics between Quantum Mechanics and Quantum Field Theory;
- Quantum Information and Computation represents a further development of Quantum Mechanics;
- what Quantum Information and Computation focuses is the logic of the Quantum Mechanics.

The reason for a graduate student major in physics to learn Quantum Information and Computation is that

- if a graduate want to do great in modern theoretical physics (especially in Quantum Field Theory or High Energy Physics or Condensed Matter Physics), he or she has to understand Quantum Information and Computation very well, because he or she must understand Modern Quantum Mechanics which is represented by Quantum Information and Computation.

1.2 What’s Quantum Information and Computation?

There are many different opinions about this question:

- In Michael A. Nielsen and Isaac L. Chuang’s opinion,
  … Quantum computation and quantum information is the study of the information process tasks that can be accomplished using quantum mechanical systems (or using fundamentals of quantum mechanics) …
According to Rolf Landauer (1961),

... Information is physical ...

which says that information is something that is encoded in the states of physical systems.

According to David Deutsch (1985), computation is a physical process and is a task that can be performed on an actual physically realized device. What computers can or can not compute is determined by the law of physics alone, and not by mathematics.

We indeed can see some similarities between computers and physical systems as shown in Table 1.1.

<table>
<thead>
<tr>
<th>Computer</th>
<th>Physical System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation</td>
<td>Motion</td>
</tr>
<tr>
<td>Input</td>
<td>Initial State</td>
</tr>
<tr>
<td>Rules</td>
<td>Laws of Motion</td>
</tr>
<tr>
<td>Output</td>
<td>Final State</td>
</tr>
</tbody>
</table>

1.3 Research topics

See Issac Chuang’s homepage at MIT physics department.

There are two main research topics in Quantum Information and Computation:

- How can physical system represent and process information?
- Can nature be better understood in terms of information or computation?

In the book of Nielson & Chuang [NC] P.P. 203: “A detailed examination and attempted justification of the physics underlying quantum computer (the quantum circuit model) is outside the scope of the present discussions and indeed outside the scope present knowledge.”

Nowadays, there are even more radical ideas

- Quantum Mechanics + Special Relativity = Quantum Field Theory.
  Quantum Information and Computation + Special Relativity = ?
- Physics is information.
- Physics is computation.
- The universe is a computer.

Gift (open problem) to fresh students in Quantum Information and Computation: $P$ versus $NP$ problem ($P = NP$ or $P \neq NP$)

- One of seven Millenium problem by Clay Mathematical Institute.
• Experts intend to believe $P \not= NP$ or $P \subset NP$, but no proof up to now.

• $P \not= NP$ means there may exist problems which can not be solved efficiently, and it may put a new constraint on Nature like the light speed or the uncertainty principle.

• Google & Wiki for details
Chapter 2

Thermodynamics and Statistical Mechanics
Chapter 3

Quantum Mechanics (I): Axioms

For those who are not shocked when they first come across quantum theory can not possibly have understood it.

—Niels Bohr

I think I can safely say that nobody understands quantum mechanics.

—Richard Feynman

Quantum mechanics: Real black magic calculus

—Albert Einstein

References:

● [Preskill] Chapter 2: Foundations I: states and ensembles;

3.1 Axioms of quantum mechanics for closed system

Principles of quantum mechanics can be classified into two parts: the static part includes “States” and “Observables, and the dynamic part includes “Evolution” and “Measurement”.

When we talk about axioms of quantum mechanics, we introduce axioms for quantum closed systems and axioms for quantum open systems respectively, see the following table. The axioms for quantum open system will be discussed in detail in Chapter 17.

3.1.1 State

Axiom 3.1.1. A state $|\psi\rangle$ or a ray $\{e^{i\alpha}|\psi\rangle\}$, from the Hilbert space $\mathcal{H}$, can make a complete description of a physical system (with no hidden variable).

Ray is an equivalent class of vectors, where global phase has no physical meaning. But, notice that the relative phase however is of physical significance. For example, the state vector $|0\rangle + |1\rangle$ is physically different from $|0\rangle + e^{i\alpha}|1\rangle$ with $e^{i\alpha} \neq 1$. Note that the superposition principle is defined only for state vectors, not for state rays.
### Axioms of Quantum Mechanics

<table>
<thead>
<tr>
<th>Closed Systems</th>
<th>Open Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space</strong></td>
<td><strong>Hilbert space ( \mathcal{H} )</strong></td>
</tr>
<tr>
<td><strong>State</strong></td>
<td>pure state vector (</td>
</tr>
<tr>
<td></td>
<td>pure state ray ( {e^{i\alpha}</td>
</tr>
<tr>
<td></td>
<td>density matrix (operator) ( \rho )</td>
</tr>
<tr>
<td></td>
<td>called mixed state</td>
</tr>
<tr>
<td><strong>Observable</strong></td>
<td>self-adjoint operator ( A = \sum_n a_n P_n )</td>
</tr>
<tr>
<td></td>
<td>with ( {a_n} \subset \mathbb{R} ) and ( {P_n} ) being orthogonal projections</td>
</tr>
<tr>
<td><strong>Measurement</strong></td>
<td>projective measurement (orthogonal)</td>
</tr>
<tr>
<td></td>
<td>( \langle A \rangle = \langle \psi</td>
</tr>
<tr>
<td></td>
<td>general measurement (non-orthogonal)</td>
</tr>
<tr>
<td></td>
<td>( \langle A \rangle = \text{tr}(A \rho) )</td>
</tr>
<tr>
<td><strong>Dynamics</strong></td>
<td>unitary evolution</td>
</tr>
<tr>
<td></td>
<td>( i\hbar \frac{d}{dt}</td>
</tr>
<tr>
<td></td>
<td>non-unitary evolution via superoperator</td>
</tr>
<tr>
<td></td>
<td>( i\hbar \frac{d}{dt} \rho(t) = [H, \rho(t)] )</td>
</tr>
</tbody>
</table>

#### 3.1.2 Observable

An observable in physics is a property of a physical system that can be measured.

**Axiom 3.1.2.** *With every observable, there exists an associated linear, self-adjoint operator \( A \), which acts in the Hilbert space \( \mathcal{H} \),

\[ A^\dagger = A \iff \langle \psi | A \phi \rangle = \langle A \psi | \phi \rangle, \quad \text{with} \quad |\psi\rangle, |\phi\rangle \in \mathcal{H}. \]  

Let \( a_n \) be one of the eigenvalues of \( A \) and \( |a_n\rangle \) is the associated eigenvector,

\[ A |a_n\rangle = a_n |a_n\rangle. \]  

All the eigenvalues \( a_n \) of \( A \) are real, while the eigenvectors \( \{|a_n\rangle\} \) form a complete orthogonal basis of the Hilbert space \( \mathcal{H} \).

It would be easy to prove that the eigenvalues of the observable \( A \) is real, and the its eigenvectors are orthogonal.

**Proof.**

(i) The eigenvalues of the observable \( A \) are real.

\[ (\langle a_n | A |a_n\rangle)^* = \langle a_n | A^\dagger |a_n\rangle = \langle a_n | A |a_n\rangle \Rightarrow a_n^* = a_n, \]  

i.e., \( a_n \) is real.

(ii) The eigenvectors of the observable \( A \) are orthogonal.

- In the case of two eigenvectors with different eigenvalues, namely

\[ \left\{ \begin{array}{l}
A |a_k\rangle = a_k |a_k\rangle \\
A |a_\ell\rangle = a_\ell |a_\ell\rangle
\end{array} \right. \]  

with \( k \neq \ell \) and \( a_k \neq a_\ell \). Therefore,

\[ \langle a_k | A |a_\ell\rangle = a_\ell \langle a_k | a_\ell\rangle \Rightarrow (a_k - a_\ell) \langle a_k | a_\ell\rangle = 0 \Rightarrow \langle a_k | a_\ell\rangle = 0. \]  

6
In the circumstance of degeneration, namely, there are at least two mutually independent eigenvectors of \( A \), but with the same eigenvalue. The set of all these eigenvectors of \( A \) associated with the specific eigenvalue would span a subspace of the Hilbert space \( \mathcal{H} \). Thus, we can employ the Schmidt orthogonalization process, to make an orthogonal basis for such a subspace, with the basis vectors still being the eigenvectors of the observable \( A \) and the corresponding eigenvalue unchanged.

The spectral theorem. We can define the orthogonal projection operators \( P_n \) onto the subspace associated with the eigenvalue \( a_n \) of the Hilbert space \( \mathcal{H} \), namely

\[
\mathcal{H}_{a_n} = \{ A | \phi \rangle = a_n | \phi \rangle \mid \forall | \phi \rangle \epsilon \mathcal{H} \},
\]

as

\[
P_n = \sum_{k=1}^{\text{Degeneracy}} \left| a_n^{(k)} \right\rangle \left( a_n^{(k)} \right| ,
\]

with \( \left\{ \left| a_n^{(k)} \right\rangle \mid k = 1, \ldots, \text{Degeneracy} \right\} \) be an orthonormal basis for the subspace \( \mathcal{H}_{a_n} \). It’s easy to see that

\[
\begin{align*}
P_n^\dagger &= P_n, \\
P_n^2 &= P_n, \\
P_n P_m &= 0, \quad \text{if } n \neq m.
\end{align*}
\]

Hence, we can expand the operator \( A \) as

\[
A = \sum_n a_n P_n,
\]

which is the spectral theorem.

3.1.3 Projective measurement

Axiom 3.1.3. For an observable \( A \) with eigenvalues \( a_n \) and eigenvectors \( |a_n\rangle \), given the system is in the state \( |\psi\rangle \), the probability of obtaining \( a_n \) (in the non-degeneration case) as the outcome of the measurement of \( A \) is

\[
\text{Prob}(a_n) = |\langle a_n | \psi \rangle|^2.
\]

And the expectation value (mean value) of the observable \( A \) would be

\[
\langle A \rangle = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}.
\]

After the measurement, the system is left in the state within the subspace corresponding to the eigenvalue \( a_n \) (the so called wavepacket collapse).

The name “projection measurement” is evident as we shall see. For example, we can see that the operator \( P_n = |a_n\rangle \langle a_n| \), itself is a self-adjoint operator, which has the mean value,

\[
\langle \psi | P_n | \psi \rangle = \langle \psi | \left( |a_n\rangle \langle a_n| \right) | \psi \rangle = \langle \psi | a_n \rangle \langle a_n | \psi \rangle = | \langle a_n | \psi \rangle |^2 = \text{Prob}(a_n),
\]

7
where $|\psi\rangle$ is the normalized state vector of the physical system.

$$
|\psi\rangle \xrightarrow{P_n=|a_n\rangle\langle a_n|} P_n|\psi\rangle = \frac{P_n|\psi\rangle}{\sqrt{\langle P_n|\psi\rangle}} = |a_n\rangle \langle a_n| \psi\rangle \langle a_n| \psi\rangle = a_n|\psi\rangle \langle a_n| \psi\rangle = \frac{a_n}{P_n|\psi\rangle} \implies P_n|\psi\rangle = \frac{a_n}{\sqrt{P_n|\psi\rangle}}
$$

which is the post-measurement state.

**Remark:** The measurement of an observable is an irreversible process, i.e., we cannot get the original state back from the measurement results. In this process, we acquire information from the system through measurement. The wave package collapse is a truly random process, and we would lose information, namely $P_n$ would kill the information encoded in the state $|a_k\rangle$ with $k \neq n$, in the same time.

| (1) Information acquirement: $P_n|\psi\rangle$ |
| (2) Information loss // wavepacket collapse: $P_m|\psi\rangle$ killed ($m \neq n$) |
| (3) True randomness: $P_n|\psi\rangle$ with probability |
| (4) Irreversible, non-unitary process: $|\psi\rangle \rightarrow P_n|\psi\rangle, P_n|\psi\rangle \rightarrow |\psi\rangle$ |

### 3.1.4 Schrödinger equation

**Axiom 3.1.4.** *The time evolution of a closed state is unitary. Equivalently, the time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:*

$$
i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H(t) |\psi(t)\rangle,
$$

where $H(t)$ is the Hamiltonian operator of the physical system.

From the Schrödinger equation we can see that the time evolution operator $\mathcal{U}(t + \Delta t, t)$ should be

$$
\mathcal{U}(t + dt, t) = 1 - \frac{i H dt}{\hbar}.
$$

As we shall see that

$$
\mathcal{U}^\dagger(t + dt, t) \mathcal{U}(t + dt, t) = \mathcal{U}(t + dt, t) \mathcal{U}^\dagger(t + dt, t) = I + O(dt^2).
$$

which means that $\mathcal{U}(t + dt, t)$ is unitary up to the second order of $dt$. And we can infer that

$$
|\psi(t)\rangle = \mathcal{U}(t, t - dt) \mathcal{U}(t - dt, t - 2dt) \cdots \mathcal{U}(dt, 0) |\psi(0)\rangle.
$$

In the special case of time-independent Hamiltonian $H$, we can write down the explicit expression of the time evolution operator $\mathcal{U}(t, 0)$, for simplicity which we denote as $\mathcal{U}(t)$,

$$
\mathcal{U}(t) = e^{-i H t / \hbar}.
$$
In the case of time-dependent Hamiltonian $H(t)$, we have to use the Dyson’s formula

$$\mathcal{U}(t) = I + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_0^t dt_n \int_0^t dt_{n-1} \cdots \int_0^t dt_1 H(t_n) H(t_{n-1}) \cdots H(t_1),$$

(3.1.18)

see Google & Wiki for more details.

**Remark:** Why is the Schrödinger equation linear? Why is the time evolution unitary, but different with measurement which is a non-unitary process? Why we have two distinct evolutions in quantum mechanics?

### 3.1.5 Composite system

For system A, $|\psi\rangle_A \in \mathcal{H}_A$, and system B, $|\varphi\rangle_B \in \mathcal{H}_B$, the composite system of system A and system B is described by the tensor product of Hilbert spaces, i.e., $|\psi\rangle_A \otimes |\varphi\rangle_B \in \mathcal{H}_A \otimes \mathcal{H}_B$. 
Chapter 4

Qubit, Quantum Gates and Bell States

References:

- [Preskill] Chapter 2: Foundations I: states and ensembles;
- [Preskill] Chapter 4: Quantum entanglement;
- [Nielsen & Chuang] Chapter 1: Introduction and overview;

4.1 Overview

Classical computation vs. Quantum Computation

<table>
<thead>
<tr>
<th>Information unit</th>
<th>Classical Computation</th>
<th>Quantum Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation</td>
<td>gate</td>
<td>quantum gate</td>
</tr>
<tr>
<td>bit</td>
<td>gate</td>
<td>qubit</td>
</tr>
</tbody>
</table>

For Classical Computation, the basic units that store the information and are manipulated are the bits. The “tool” that can manipulate bits are the so-called classical logic gate.

- bit: the short name of “binary digit”, which can only take the value of 0 or 1.
- gate:

\[
gate \{ \begin{align*}
\text{one-bit gate} & : \text{NOT} \\
\text{two-bit gate} & : \text{AND, OR, } \ldots
\end{align*}
\]

NOT gate: $\overline{a}$ in mod 2. AND gate: $a \land b \equiv a \cdot b$, OR gate: $a \lor b \equiv a \oplus b$ in mod 2.

While for Quantum computation, the counterpart for bit should be qubit, and for classical logic gates are quantum gates.

- qubit: the short name of “quantum bit”, which can be considered as a state vector in the two-dimensional Hilbert space $\mathcal{H}_2$:

\[
\begin{pmatrix}
\alpha \\
\beta
\end{pmatrix} = \alpha |0\rangle + \beta |1\rangle, \quad \text{with} \quad |\alpha|^2 + |\beta|^2 = 1, \quad \alpha, \beta \in \mathbb{C}, \quad (4.1.1)
\]
where $|0\rangle$ could mean “spin up” and for $|1\rangle$ mean “spin down”, i.e., this is a two-level physical system, and the spin-$\frac{1}{2}$ system is a typical example.

- quantum gate:

$$\begin{aligned}
\text{quantum gate} & = \left\{ \begin{array}{ll}
\text{single-qubit gate:} & \left( \begin{array}{c} \alpha \\ \beta \end{array} \right) \xrightarrow{SU(2)} \left( \begin{array}{c} \alpha' \\ \beta' \end{array} \right), \\
\text{two-qubit gate:} & \left( \begin{array}{c} \alpha_1 \\ \beta_1 \end{array} \right) \otimes \left( \begin{array}{c} \alpha_2 \\ \beta_2 \end{array} \right) \xrightarrow{SU(4)} \left( \begin{array}{c} \alpha'_1 \beta'_1 \\ \alpha'_2 \beta'_2 \end{array} \right).
\end{array} \right.
\end{aligned}$$

### 4.2 Pure state formalism of a qubit

The Hilbert space $\mathcal{H}_2$ can be spanned by basis $\{ |0\rangle, |1\rangle \}$, i.e.,

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \forall |\psi\rangle \in \mathcal{H}_2, \quad \text{with } \alpha, \beta \in \mathbb{C} \text{ and } |\alpha|^2 + |\beta|^2 = 1.$$

We may notice that $\alpha$ and $\beta$ being complex numbers means four real numbers (four degrees of freedom). With the constraint $|\alpha|^2 + |\beta|^2 = 1$, we can cut down the degrees of freedom into three. And, if we ignore the global phase of the state vector (qubit), which has no physical meaning, then we can reduce the degrees of freedom to be two. Therefore, the state vector $|\psi\rangle$ can be described by two real numbers $(\theta, \varphi)$, for example

$$|\psi_+(\theta, \varphi)\rangle := \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\varphi} \sin \frac{\theta}{2} \end{pmatrix}, \quad \text{with } 0 \leq \theta < \pi, \text{ and } 0 \leq \varphi < 2\pi. \quad (4.2.1)$$

The two real variables $(\theta, \varphi)$ can actually determine a unit vector in the three-dimensional Euclidean $\mathbb{R}^3$, namely $\hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$, as shown in Figure 4.1. And we shall also see that with $\hat{n}$ pointing along different directions, namely different values of $\theta$ and $\varphi$, the state vectors $|\psi_+(\theta, \varphi)\rangle$ are different:

1. $\hat{n} = \hat{e}_x = (1, 0, 0)$, $\theta = \frac{\pi}{2}, \varphi = 0$, $\Rightarrow |\psi_+(\frac{\pi}{2}, 0)\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)$;
2. $\hat{n} = \hat{e}_y = (0, 1, 0)$, $\theta = \frac{\pi}{2}, \varphi = \frac{\pi}{2}$, $\Rightarrow |\psi_+(\frac{\pi}{2}, \frac{\pi}{2})\rangle = \frac{1}{\sqrt{2}} (|0\rangle + i |1\rangle)$;
3. $\hat{n} = \hat{e}_z = (0, 0, 1)$, $\theta = 0, \varphi \in [0, 2\pi)$, $\Rightarrow |\psi_+(0, \varphi)\rangle = |0\rangle$.

The vector $\hat{n}$ in three dimensional space is called Bloch vector, and the unit boundary of the vector set is called Bloch sphere, i.e., $|\hat{n}| = 1$. 

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{bloch_sphere.png}
\caption{Bloch sphere}
\end{figure}
4.2.1 Single-qubit gate $SU(2)$ in the spin-1/2 case

4.2.1.1 Spin-1/2 operator and Pauli matrices

The spin-1/2 operator can be expressed as

$$\hat{J} = \frac{1}{2} \hbar \hat{\sigma}$$

or

$$\hat{J} = \frac{1}{2} \hat{\sigma},$$  \hspace{1cm} (4.2.2)

if we set the reduced Planck constant $\hbar$ to be one. And the $\hat{\sigma}$ is the so-called Pauli vector,

$$\hat{\sigma} := \sigma_x \hat{e}_x + \sigma_y \hat{e}_y + \sigma_z \hat{e}_z \hspace{1cm} (4.2.3)$$

where $\sigma_x$, $\sigma_y$ and $\sigma_z$ are the Pauli matrices, which have the standard form

$$\begin{align*}
\sigma_x &:= \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\
\sigma_x &:= \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\
\sigma_z &:= \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{align*}$$

The Pauli matrices satisfy the following properties,

- the anticommutative relation
  \begin{align*}
  \{ \sigma_i, \sigma_j \} &= 2 \delta_{ij}; \hspace{1cm} (4.2.4)
  \end{align*}
- the commutative relation
  \begin{align*}
  [ \sigma_i, \sigma_j ] &= 2i \varepsilon_{ijk} \sigma_k, \quad \text{with } i, j, k \in \{1, 2, 3\} \text{ and } k \notin \{i, j\}. \hspace{1cm} (4.2.5)
  \end{align*}

4.2.1.2 Spinor representation of $SU(2)$ group

We can define a unitary operator $\mathcal{D}_{\frac{1}{2}}(\theta, \hat{n})$ for the spin-1/2 system which is induced by

the rotation in the Euclidean space $\mathbb{R}^3$ along the direction $\hat{n}$ through an angle of $\theta$:

$$\mathcal{D}_{\frac{1}{2}}(\theta, \hat{n}) := \exp(-i \theta \hat{n} \cdot \hat{J})$$  \hspace{1cm} (4.2.6)

which is equivalent to

$$\mathcal{D}_{\frac{1}{2}}(\theta, \hat{n}) = \exp\left(-i \frac{\theta}{2} \hat{n} \cdot \hat{\sigma} \right) = \cos \frac{\theta}{2} - i \hat{n} \cdot \hat{\sigma} \sin \frac{\theta}{2}. \hspace{1cm} (4.2.7)$$

$\mathcal{D}_{\frac{1}{2}}(\theta, \hat{n})$ is a single-qubit gate. And it has some interesting features, for instance

$$\mathcal{D}_{\frac{1}{2}}(2\pi, \hat{n}) = -1, \quad \mathcal{D}_{\frac{1}{2}}(4\pi, \hat{n}) = 1. \hspace{1cm} (4.2.8)$$

Though this would be meaningless if it only gives the global phase, there can be significant effect if the relative phase is changed because of this.
4.2.2 Properties

**Thm 4.2.2.1.** If we define $\sigma_n = \tilde{\sigma} \cdot \tilde{n}$, then

$$\sigma_n |\psi_+(\theta, \varphi)\rangle = |\psi_+(\theta, \varphi)\rangle, \text{ with } \tilde{n} := (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta).$$

(4.2.9)

**Proof.** Let’s firstly express $\sigma_n$ with the Pauli matrices,

$$\sigma_n = \sin \theta \cos \varphi \sigma_1 + \sin \theta \sin \varphi \sigma_2 + \cos \theta \sigma_3$$

$$= \sin \theta \cos \varphi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin \theta \sin \varphi \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \cos \theta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos \theta & \sin \theta \cos \varphi - i \sin \theta \sin \varphi \\ \sin \theta \cos \varphi + i \sin \theta \sin \varphi & -\cos \theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos \theta & \sin \theta (\cos \varphi - i \sin \varphi) \\ \sin \theta (\cos \varphi + i \sin \varphi) & -\cos \theta \end{pmatrix},$$

namely

$$\sigma_n = \begin{pmatrix} \cos \theta & \sin \theta e^{-i\varphi} \\ \sin \theta e^{i\varphi} & -\cos \theta \end{pmatrix}. \quad (4.2.10)$$

Therefore, we can calculate $\sigma_n |\psi_+(\theta, \varphi)\rangle$ in the following way

$$\sigma_n |\psi_+(\theta, \varphi)\rangle = \begin{pmatrix} \cos \theta & \sin \theta e^{-i\varphi} \\ \sin \theta e^{i\varphi} & -\cos \theta \end{pmatrix}\begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\varphi} \sin \frac{\theta}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \cos \theta \cos \frac{\theta}{2} + \sin \theta e^{-i\varphi} e^{i\varphi} \sin \frac{\theta}{2} \\ \sin \theta e^{i\varphi} \cos \frac{\theta}{2} - \cos \theta e^{i\varphi} \sin \frac{\theta}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\varphi} \sin \frac{\theta}{2} \end{pmatrix}$$

$$= |\psi_+(\theta, \varphi)\rangle.$$  

(4.2.11)

There we get E.Q. (4.2.9) proved.

**Thm 4.2.2.2.**

$$\{\psi_+(\theta, \varphi)| \tilde{\sigma} \cdot \tilde{m} |\psi_+(\theta, \varphi)\rangle\} = \tilde{n} \cdot \tilde{m},$$

(4.2.12)

with $\tilde{n} := (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ and $\tilde{m} \in \mathbb{R}^3$, $\|\tilde{m}\| = 1$.

**Proof.** In analogy to E.Q. (4.2.10), we can get the expression for $\tilde{\sigma} \cdot \tilde{m}$:

$$\tilde{\sigma} \cdot \tilde{m} = \begin{pmatrix} \cos \theta' & \sin \theta' e^{-i\varphi'} \\ \sin \theta' e^{i\varphi'} & -\cos \theta' \end{pmatrix}, \text{ with } \tilde{m} := (\sin \theta' \cos \varphi', \sin \theta' \sin \varphi', \cos \theta').$$

(4.2.13)
Therefore, we can get
\[
\begin{align*}
\langle \psi_+(\theta, \varphi) | \hat{\sigma} \cdot \hat{m} | \psi_+(\theta, \varphi) \rangle &= \left( \cos \frac{\theta}{2} e^{-i \varphi} \sin \frac{\theta}{2} \right) \left( \cos \theta' \sin \theta' e^{-i \varphi'} - \cos \theta' \right) \left( \cos \frac{\theta}{2} e^{i \varphi} \right) \\
&= \left( \cos \frac{\theta}{2} e^{-i \varphi} \sin \frac{\theta}{2} \right) \left( \cos \frac{\theta}{2} \cos \theta' + \sin \frac{\theta}{2} \sin \theta' e^{i (\varphi - \varphi')} \right) \\
&= \cos \theta \left( \cos \frac{\theta}{2} \cos \theta' + \sin \frac{\theta}{2} \sin \theta' \right) + e^{-i \varphi} \sin \frac{\theta}{2} \left( \cos \frac{\theta}{2} \sin \theta' e^{i (\varphi - \varphi')} \right) \\
&= \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos (\varphi - \varphi'). 
\end{align*}
\]
where we have used the fact that
\[
\begin{align*}
\hat{n} \cdot \hat{m} &= \left( \sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta \right) \left( \sin \theta' \cos \varphi', \sin \theta' \sin \varphi', \cos \theta' \right)^T \\
&= \sin \theta \sin \theta' \cos \varphi \cos \varphi' + \sin \theta \sin \theta' \sin \varphi \sin \varphi' + \cos \theta \cos \theta' \\
&= \sin \theta \sin \theta' \left( \cos \varphi \cos \varphi' + \sin \varphi \sin \varphi' \right) + \cos \theta \cos \theta' \\
&= \sin \theta \sin \theta' \cos (\varphi - \varphi') + \cos \theta \cos \theta'. 
\end{align*}
\]
(4.2.15)

There, E.Q. (4.2.12) is verified, too. \(\square\)

**Remark:** For the expression in terms of density matrix, we have
\[
\text{tr}\left( \rho(\hat{\sigma} \cdot \hat{m}) \right) = \hat{n} \cdot \hat{m},
\]
(4.2.16)
where \(\rho = |\psi_+(\theta, \varphi)\rangle \langle \psi_+(\theta, \varphi)|\).

**Thm 4.2.2.3.** As we shall know from the definition of \(|\psi_+(\theta, \varphi)\rangle\) that
\[
|0\rangle = |\uparrow_z\rangle = |\psi_+(0, \varphi)\rangle,
\]
(4.2.17)
then
\[
|\psi_+(\theta, \varphi)\rangle = \mathcal{D}(\hat{e}_z \rightarrow \hat{n}) |0\rangle, \text{ with } \hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),
\]
(4.2.18)
where
\[
\mathcal{D}(\hat{e}_z \rightarrow \hat{n}) := \left( \begin{array}{cc}
\cos \frac{\theta}{2} & -e^{-i \varphi} \sin \frac{\theta}{2} \\
\cos \frac{\theta}{2} & \cos \frac{\theta}{2}
\end{array} \right).
\]
(4.2.19)

**Proof.** Firstly, we should realize that
\[
\mathcal{D}_z \left( \hat{e}_z \rightarrow \hat{n} \right) = \mathcal{D}_z \left( \hat{n}', \hat{n}' \right) = \exp \left( -i \frac{\theta}{2} \hat{\sigma} \cdot \hat{n}' \right),
\]
(4.2.20)
where we define
\[
\begin{align*}
\hat{n}_{xy} &:= (\cos \varphi, \sin \varphi, 0), \\
\hat{n}_{xy}' &:= \left( \cos (\varphi + \frac{\pi}{2}), \sin (\varphi + \frac{\pi}{2}), 0 \right) = (- \sin \varphi, \cos \varphi, 0),
\end{align*}
\]
(4.2.21)
\[ \mathbf{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \]

\[ \mathbf{n}'_{xy} = (-\sin \varphi, \cos \varphi, 0) \]

\[ \mathbf{n}_{xy} = (\cos \varphi, \sin \varphi, 0) \]

Figure 4.2: Rotation in the Euclidean space

i.e., \( \mathbf{n}'_{xy} \) is a unit vector that is orthogonal to both \( \mathbf{n} \) and \( \hat{e}_z \), and they can form a right-hand coordinate system, see Figure 4.2, which means that the corresponding rotation taken place in the Euclidean space should be a rotation around \( \mathbf{n}'_{xy} \) through an angle \( \theta \).

And, now it would be easy to show that E.Q. (4.2.18) is right.

E.Q. (4.2.20) can be rewritten as

\[
\mathcal{D}_2 (\hat{e}_z \to \mathbf{n}) = \cos \frac{\theta}{2} - \frac{i}{2} \mathbf{n} \cdot \mathbf{n}_{xy} \sin \frac{\theta}{2} \\
= \cos \frac{\theta}{2} - i \sin \frac{\theta}{2} (\sin \varphi \mathbf{\sigma}_1 + \cos \varphi \mathbf{\sigma}_2) \\
= \begin{pmatrix} \cos \frac{\theta}{2} & 0 \\ 0 & \cos \frac{\theta}{2} \end{pmatrix} + \sin \frac{\theta}{2} \begin{pmatrix} 0 & -\cos \varphi + i \sin \varphi \\ \cos \varphi + i \sin \varphi & 0 \end{pmatrix} \\
= \begin{pmatrix} \cos \frac{\theta}{2} - \sin \frac{\theta}{2} e^{-i\varphi} \\ \sin \frac{\theta}{2} e^{i\varphi} \end{pmatrix}.
\]

(4.2.22)

And we can verify E.Q. (4.2.18) directly,

\[
\mathcal{D}_2 (\hat{e}_z \to \mathbf{n}) |0\rangle = \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} e^{-i\varphi} \\ \sin \frac{\theta}{2} e^{i\varphi} & \cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\varphi} \end{pmatrix} = |\psi_+ (\theta, \varphi)\rangle.
\]

(4.2.23)

Remarks:

(1) As we have shown above

\[
\begin{cases} 
\mathbf{\sigma} \cdot \mathbf{n} |\psi_+ (\theta, \varphi)\rangle = |\psi_+ (\theta, \varphi)\rangle, \\
|\psi_+ (\theta, \varphi)\rangle = \mathcal{D}_2 (\hat{e}_z \to \mathbf{n}) |0\rangle.
\end{cases}
\]

(4.2.24)

If we define

|\psi_- (\theta, \varphi)\rangle = \mathcal{D}_2 (\hat{e}_z \to \mathbf{n}) |1\rangle,

(4.2.25)

then, we could verify that

\[
\mathbf{\sigma} \cdot \mathbf{n} |\psi_- (\theta, \varphi)\rangle = -|\psi_- (\theta, \varphi)\rangle.
\]

(4.2.26)

Firstly,

\[
|\psi_- (\theta, \varphi)\rangle = \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} e^{-i\varphi} \\ \sin \frac{\theta}{2} e^{i\varphi} & \cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\varphi} \\ \cos \frac{\theta}{2} \end{pmatrix}.
\]

(4.2.27)
Then,

$$\sigma \cdot \hat{n} |\psi_-(\theta, \varphi)\rangle = \left( \begin{array}{cc} \cos \theta & \sin \theta e^{-i\varphi} \\ -\sin \theta e^{i\varphi} & -\cos \theta \end{array} \right) \left( \begin{array}{c} -\sin \frac{\theta}{2} e^{-i\varphi} \\ \cos \frac{\theta}{2} \end{array} \right)$$

$$= \left( \begin{array}{cc} -\sin \frac{\theta}{2} e^{-i\varphi} \cos \theta + \sin \frac{\theta}{2} \sin \theta e^{-i\varphi} \\ -\sin \frac{\theta}{2} e^{i\varphi} \sin \theta e^{i\varphi} - \cos \frac{\theta}{2} \cos \theta \end{array} \right)$$

$$= \left( \begin{array}{c} \sin \frac{\theta}{2} e^{-i\varphi} \\ -\cos \frac{\theta}{2} \end{array} \right)$$

$$= -|\psi_-(\theta, \varphi)\rangle. \quad (4.2.28)$$

(2) The spinor representation of the $SO(3)$ group, $D(R)$, satisfies

$$D(R) \vec{x} \cdot \sigma \sigma^\dagger (R) = \vec{x}' \cdot \sigma,$$  \hspace{1cm} (4.2.29)

in which the vector $\vec{x}$ is rotated under $\vec{x}' = R \vec{x}$, i.e., $x_j' = R_{ij} x_i$.

Example: $D(R) \equiv \mathcal{D}_z(\hat{e}_z \rightarrow \hat{n})$ gives

$$D(R) \sigma_3 D^\dagger(R) = \sigma \cdot \hat{n}. \quad (4.2.30)$$

It implies that

$$\sigma \cdot \hat{n} |\bar{n}\rangle = D(R) \sigma_3 D^\dagger(R) D(R) |\bar{e}_z\rangle$$

$$= D(R) \sigma_3 |0\rangle$$

$$= |\bar{n}\rangle. \quad (4.2.31)$$

4.2.3 Physical realization of qubit

<table>
<thead>
<tr>
<th></th>
<th>Electron</th>
<th>Photon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spin</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>Mass</td>
<td>0.5MeV</td>
<td>0</td>
</tr>
<tr>
<td>Qubit</td>
<td>Spin-state</td>
<td>Photon-polarization</td>
</tr>
</tbody>
</table>

Note: Though the two-level quantum system is equivalent with the $\frac{1}{2}$-spin system, not every two-level system, like photon-polarization state, is transformed as a spinor, due to the fact that the photon has spin 1.

4.3 Bell states

- Bell states are maximally entangled two-qubit pure states, also named as EPR pair states.
- Bell states are widely used in quantum information and computation: Bell’s inequalities, dense coding, teleportation, cryptography, etc.
4.3.1 Notation

Pauli matrices are unitary matrices defined in spin-$\frac{1}{2}$ space:

$$
\begin{align*}
\sigma_x & := X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \\
\sigma_y & := -iY := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \\
\sigma_z & := Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{align*}
$$

There we get three quantum gates: quantum gate $X$, quantum gate $Z$ and quantum gate $Y = ZX$.

Def 4.3.1 (Bell states). The following four double-qubit states

$$
\begin{align*}
|\phi^+\rangle & := |\psi(0,0)\rangle := \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right), \\
|\phi^-\rangle & := |\psi(0,1)\rangle := \frac{1}{\sqrt{2}} \left( |00\rangle - |11\rangle \right), \\
|\psi^+\rangle & := |\psi(1,0)\rangle := \frac{1}{\sqrt{2}} \left( |01\rangle + |10\rangle \right), \\
|\psi^-\rangle & := |\psi(1,1)\rangle := \frac{1}{\sqrt{2}} \left( |01\rangle - |10\rangle \right),
\end{align*}
$$

are the so called Bell states.

Lemma 4.3.1.1. All the four Bell states defined in E.Q. (4.3.2) can be expressed as

$$
|\psi(i,j)\rangle = \left( I_2 \otimes X^i Z^j \right) |\psi(0,0)\rangle, \quad \text{with } i, j = 0, 1.
$$

Proof. With the definition of the Bell states (4.3.2), we can check E.Q. (4.3.3) one by one for all the four cases of $i, j = 0, 1$.

(1) For the case of $i = j = 0$, E.Q (4.3.3) is absolutely right.

(2) For the case of $i = 0, j = 1$, the right-hand-side (RHS) of E.Q. (4.3.3) should be

$$
\text{RHS} = \left( I_2 \otimes Z \right) |\psi(0,0)\rangle \\
= \left( I_2 \otimes Z \right) \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right) \\
= \frac{1}{\sqrt{2}} \left( |00\rangle - |11\rangle \right) \\
= |\psi(0,1)\rangle,
$$

which means E.Q (4.3.3) is correct in this case.
(3) For the case of $i = 1$, $j = 0$, we evaluate the right-hand-side (RHS) of Eq. 4.3.3

$$\text{RHS} = \left( I_2 \otimes X \right) |\psi(0, 0)\rangle$$
$$= \left( I_2 \otimes X \right) \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$$
$$= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle)$$
$$= |\psi(1, 0)\rangle,$$

(4.3.5)

which also shows the legitimation of Eq. 4.3.3 in this circumstance.

(4) For the case of $i = j = 1$, we can get the right-hand-side (RHS) of Eq. 4.3.3

$$\text{RHS} = \left( I_2 \otimes XZ \right) |\psi(0, 0)\rangle$$
$$= \left( I_2 \otimes XZ \right) \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$$
$$= \left( I_2 \otimes X \right) \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle)$$
$$= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)$$
$$= |\psi(1, 1)\rangle,$$

(4.3.6)

which says Eq. 4.3.3 for the last situation.

Now, we can conclude that for all the four cases $i, j = 0, 1$, Eq. 4.3.3 is valid.

\[\square\]

Remark: For four Bell states (4.3.2), we have the following geometric representations,

\[|\psi(00)\rangle = |\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) = \]

(4.3.7)

\[|\psi(10)\rangle = |\psi^+\rangle = \left( I_2 \otimes X \right) |\psi(00)\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) = \]

(4.3.8)

\[|\psi(01)\rangle = |\phi^-\rangle = \left( I_2 \otimes Z \right) |\psi(00)\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) = \]

(4.3.9)

\[|\psi(11)\rangle = |\psi^-\rangle = \left( I_2 \otimes XZ \right) |\psi(00)\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) = \]

(4.3.10)

The vertical line denotes one Hilbert space $H_2$.

Lemma 4.3.1.2. The four Bell states can also be expressed as

$$|\psi(i, j)\rangle = \frac{1}{\sqrt{2}} \left( |0i\rangle + (-1)^{\tilde{i}} |1i\rangle \right),$$

(4.3.11)

where $\tilde{i} = (i + 1) \mod 2$ and $i = 0, 1$.

We can show in the following that Eq. 4.3.11 is consistent with Eq. 4.3.2.
Proof. From E.Q. (4.3.3) we can get
\[ |\psi(i, j)\rangle = \frac{1}{\sqrt{2}} \left( I_2 \otimes X^i \right) \left( |00\rangle + |11\rangle \right) \]
\[ = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes X^j |0\rangle + |1\rangle \otimes X^j |1\rangle \right) \]
\[ = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes X^j |0\rangle + (-1)^j |1\rangle \otimes X^j |1\rangle \right) \]
\[ = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes |i\rangle + (-1)^j |1\rangle \otimes |\bar{i}\rangle \right), \]
which is what exactly E.Q. (4.3.11) shows. \( \square \)

Remark: All the four Bell states defined in E.Q. (4.3.2) are normalized:

(i) \( \langle \psi(0, 0) | \psi(0, 0) \rangle = 1 \), since
\[ \frac{1}{\sqrt{2}} \left( \langle 00 \rangle + \langle 11 \rangle \right) \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right) = \frac{1}{2} \left( \langle 00 | 00 \rangle + \langle 00 | 11 \rangle + \langle 11 | 00 \rangle + \langle 11 | 11 \rangle \right) \]
\[ = \frac{1}{2} \left( 1 + 0 + 0 + 1 \right) \]
\[ = 1. \]

(ii) \( \langle \psi(i, j) | \psi(i, j) \rangle = 1 \), because
\[ \langle \psi(i, j) | \psi(i, j) \rangle = \left( \langle \psi(0, 0) | I_2 \otimes Z^j X^i \right) \left( I_2 \otimes X^i Z^j | \psi(0, 0) \rangle \right) \]
\[ = \langle \psi(0, 0) | I_2 \otimes X^j X^i Z^j | \psi(0, 0) \rangle \]
\[ = \langle \psi(0, 0) | \psi(0, 0) \rangle \]
\[ = 1. \]

Lemma 4.3.1.3. Let’s \( M \) denotes an arbitrary \( SU(2) \) matrix, namely single-qubit gate, and \( M^T \) is the transpose of \( M \). Then
\[ \left( I_2 \otimes M \right) |\psi(0, 0)\rangle = \left( M^T \otimes I_2 \right) |\psi(0, 0)\rangle, \]
with the diagrammatical representation
\[ \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
| & \downarrow \bar{M} = M^T \downarrow |
\end{array}
\end{array}
\end{array}
\end{array} \]
Proof. Firstly, evaluate the left-hand-side ($LHS$) of E.Q. \[4.3.12\]

\[
LHS = \left(I_2 \otimes M\right) |\psi(0, 0)\rangle
\]

\[
= \frac{I_2 \otimes M}{\sqrt{2}} \sum_{i=0}^{1} |i\rangle
\]

\[
= \frac{1}{\sqrt{2}} \sum_{i=0}^{1} |i\rangle \otimes M |i\rangle
\]

\[
= \frac{1}{\sqrt{2}} \left( |0\rangle \otimes (M_{00}|0\rangle + M_{10}|1\rangle) + |1\rangle \otimes (M_{01}|0\rangle + M_{11}|1\rangle) \right)
\]

\[
= \frac{1}{\sqrt{2}} \left( (M_{00}|0\rangle + M_{01}|1\rangle) \otimes |0\rangle + (M_{10}|0\rangle + M_{11}|1\rangle) \otimes |1\rangle \right)
\]

\[
= \frac{1}{\sqrt{2}} \left( M^T |0\rangle \otimes |0\rangle + M^T |1\rangle \otimes |1\rangle \right)
\]

\[
= \frac{M^T \otimes I_2}{\sqrt{2}} \left( |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \right),
\]

i.e.,

\[
LHS = \left(M^T \otimes I_2\right) |\psi(00)\rangle,
\]

which happens to be equal to the right-hand-side of E.Q. \[4.3.12\]. In the derivation we have assumed that $M$ has the matrix form

\[
M := \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix},
\]

namely

\[
M^T = \begin{pmatrix} M_{00} & M_{10} \\ M_{01} & M_{11} \end{pmatrix}.
\]

We may notice that

\[
\begin{cases}
X^T = X, \\
Z^T = Z, \\
(XZ)^T = ZX.
\end{cases}
\]

Therefore, we can get the conclusion that

\[
\begin{cases}
|\psi^+\rangle = \left(I_2 \otimes X\right) |\psi(00)\rangle = \left(X \otimes I_2\right) |\psi(00)\rangle, \\
|\phi^-\rangle = \left(I_2 \otimes Z\right) |\psi(00)\rangle = \left(Z \otimes I_2\right) |\psi(00)\rangle, \\
|\psi^-\rangle = \left(I_2 \otimes XZ\right) |\psi(00)\rangle = \left(XZ \otimes I_2\right) |\psi(00)\rangle.
\end{cases}
\]

\[
|\psi(10)\rangle = |\psi^+\rangle = x = x
\]

\[
|\psi(01)\rangle = |\phi^-\rangle = z = z
\]

\[
|\psi(11)\rangle = |\psi^-\rangle = xz = zx
\]
4.3.2 Parity-bit \((i)\) and Phase-bit \((j)\)

- **Parity-bit \((i)\)**
  - \(i = 0\), two spins are aligned, denoted by “\(\phi\)”;  
  - \(i = 1\), two spins are anti-aligned, denoted by “\(\psi\)”.

- **Phase-bit \((j)\)**
  - \(j = 0\), superposition with “\(+\)”, i.e., with equal phase;  
  - \(j = 1\), superposition with “\(-\)”, i.e., with opposite phase.

\[
\begin{array}{c|cc}
\hline
i & j & 0 & 1 \\
\hline
0 & |\phi^+\rangle & |\phi^−\rangle \\
1 & |\psi^+\rangle & |\psi^−\rangle \\
\hline
\end{array}
\]

**Lemma 4.3.2.1.** Bell states are eigenstates of the two commutative operators:

1. **parity-bit operator:** \(Z \otimes Z\),
2. **phase-bit operator:** \(X \otimes X\),

namely

\[
\begin{align*}
\left( Z \otimes Z \right) |\psi(i,j)\rangle &= (-1)^i |\psi(i,j)\rangle, \\
\left( X \otimes X \right) |\psi(i,j)\rangle &= (-1)^j |\psi(i,j)\rangle.
\end{align*}
\]

(4.3.20)

**Proof.** The parity-bit operator and the phase-bit operator are commutative, because

\[
\begin{align*}
(X \otimes X)(Z \otimes Z) &= (XZ) \otimes (XZ) \\
&= (Z \otimes Z)(X \otimes X) \\
&= (Z \otimes Z)(X \otimes X),
\end{align*}
\]

(4.3.21)

namely

\[
[X \otimes X, Z \otimes Z] = 0.
\]

(4.3.22)

We may examine the two operators separately.
(a) Parity-bit operator $Z \otimes Z$.

\[
(Z \otimes Z) |\psi(i,j)\rangle = (Z \otimes Z) \frac{1}{\sqrt{2}} \left[ |0i⟩ + (-1)^j |1i⟩ \right] = \frac{1}{\sqrt{2}} \left[ Z |0⟩ \otimes Z |i⟩ + (-1)^j Z |1⟩ \otimes Z |i⟩ \right] = \frac{1}{\sqrt{2}} \left[ |0⟩ \otimes (-1)^i |i⟩ + (-1)^j (-1)^i |1⟩ \otimes (-1)^{i-1} |i⟩ \right] = (1)^i \frac{1}{\sqrt{2}} \left[ |0⟩ \otimes |i⟩ + (-1)^j |1⟩ \otimes |i⟩ \right],
\]

i.e.,

\[ (Z \otimes Z) |\psi(i,j)\rangle = (-1)^i |\psi(i,j)\rangle. \quad (4.3.23) \]

In the derivation, we have used the facts that

\[
\begin{align*}
Z |i⟩ &= (-1)^i |i⟩, \\
Z |i⟩ &= (-1)^{i-1} |i⟩. \quad (4.3.24)
\end{align*}
\]

(b) Phase-bit operator $X \otimes X$.

\[
(X \otimes X) |\psi(i,j)\rangle = (X \otimes X) \frac{1}{\sqrt{2}} \left[ |0i⟩ + (-1)^j |1i⟩ \right] = \frac{1}{\sqrt{2}} \left[ X |0⟩ \otimes X |i⟩ + (-1)^j X |1⟩ \otimes X |i⟩ \right] = \frac{1}{\sqrt{2}} \left[ |1⟩ \otimes |i⟩ + (-1)^j |0⟩ \otimes |i⟩ \right] = (1)^j \frac{1}{\sqrt{2}} \left[ |0⟩ \otimes |i⟩ + (-1)^i |1⟩ \otimes |i⟩ \right],
\]

namely

\[ (X \otimes X) |\psi(i,j)\rangle = (-1)^j |\psi(i,j)\rangle. \quad (4.3.25) \]

And we have used the facts in the following to go through the above derivation,

\[
\begin{align*}
X |i⟩ &= |i⟩, \\
X |i⟩ &= |i⟩. \quad (4.3.26)
\end{align*}
\]

With E.Q. (4.3.23) and E.Q. (4.3.25), we get E.Q. (4.3.20) proved.

4.3.3 Orthonormal basis of the two-qubit Hilbert space

Thm 4.3.3.1. Bell states form an orthonormal basis of two-qubit Hilbert space, namely

\[ \{ |\phi^+⟩, |\psi^+⟩, |\phi^-⟩, |\psi^-⟩ \} \]

is an orthonormal basis of two-qubit Hilbert space.
Proof. Firstly, we would prove that all the Bell states are mutually orthogonal and normalized. Secondly, the completeness of the set of the four Bell state vectors would be verified.

(a) The set of the four Bell state vectors make an orthonormal vector set. From E.Q. (4.3.3), we can derive that

\[
\langle \psi(i,j) | \psi(i',j') \rangle = \langle \psi(0,0) | (I_2 \otimes Z^i X^j) (I_2 \otimes X^{i'} Z^{j'}) | \psi(0,0) \rangle
\]

\[
= \frac{1}{2} \langle \langle 00 | + \langle 11 | \rangle (I_2 \otimes Z^i X^{i+1} Z^{j'}) \rangle (|00 | + |11 \rangle)
\]

\[
= \frac{1}{2} \sum_{k,\ell=0}^1 \langle k | k | \phi \rangle \langle \phi | Z^i X^{i+1} Z^{j'} | \ell \rangle
\]

\[
= \frac{1}{2} \sum_{k,\ell=0}^1 \delta_{k\ell} \langle k | Z^i X^{i+1} Z^{j'} | \ell \rangle
\]

\[
= \frac{1}{2} \sum_{k=0}^1 \langle k | Z^i X^{i+1} Z^{j'} | k \rangle
\]

\[
= \frac{1}{2} \text{tr}(Z^j X^{i+1} Z^{j'})
\]

\[
= \frac{1}{2} \text{tr}(Z^{j+i} X^{i+1} Z^i).
\]

Because

\[
Z^2 = X^2 = I_2,
\]

and \(i, j, i', j' \in \{0, 1\}\), then we can obtain

\[
\langle \psi(i,j) | \psi(i',j') \rangle = \delta_{jj'} \delta_{ii'}, \tag{4.3.27}
\]

which means that the four Bell states are mutually orthogonal and are all of unit length.

In diagrammatical representation, we use the cup configuration for ket state, and cap configuration for bra state, shown as

\[
| \psi(i' j') \rangle = \bigcup x' z' \quad \langle \psi(ij) | = \bigcap z' x'
\]

(4.3.28)

Therefore, the orthonormal relation has the diagrammatical representation

\[
\langle \psi(ij)| \psi(i'j') \rangle = \quad \tag{4.3.29}
\]

From the diagrammatical rules, we would have the normalized trace of the single-qubit gates on the loop,

\[
\langle \psi(ij)| \psi(i'j') \rangle = \frac{1}{2} \text{tr}(Z^j X^i X^{i'} Z^{j'})
\]

\[
= \frac{1}{2} \text{tr}(Z^{j+i} X^{i+i'})
\]

\[
= \delta_{jj'} \delta_{ii'}, \tag{4.3.30}
\]
(b) The vector set consisted of the four Bell states is complete. By utilizing E.Q. (4.3.11), we can get

\[
\sum_{i,j=0}^{1} |\psi(i, j)\rangle \langle \psi(i, j)|
\]

\[
= \frac{1}{\sqrt{2}} \sum_{i,j=0}^{1} \left( |0i\rangle + (-1)^j |1\bar{i}\rangle \right) \left( |0i\rangle + (-1)^j |1\bar{i}\rangle \right) \frac{1}{\sqrt{2}}
\]

\[
= \frac{1}{2} \sum_{i,j=0}^{1} \left( |0i\rangle \langle 0i| + (-1)^j |1\bar{i}\rangle \langle 0i| + (-1)^j |0i\rangle \langle 1\bar{i}| + |1\bar{i}\rangle \langle 1\bar{i}| \right).
\]

As we see that

\[
\sum_{i,j=0}^{1} (-1)^j |1\bar{i}\rangle \langle 0i| = 0, \quad \sum_{i,j=0}^{1} (-1)^j |0i\rangle \langle 1\bar{i}| = 0,
\]

(4.3.31)

therefore

\[
\sum_{i,j=0}^{1} |\psi(i, j)\rangle \langle \psi(i, j)| = \frac{1}{2} \sum_{i,j=0}^{1} \left( |0i\rangle \langle 0i| + |1\bar{i}\rangle \langle 1\bar{i}| \right)
\]

\[
= \sum_{i=0}^{1} \left( |0i\rangle \langle 0i| + |1i\rangle \langle 1i| \right)
\]

\[
= \sum_{i,j=0}^{1} |ji\rangle \langle ji|
\]

\[
= I_4,
\]

i.e.,

\[
\sum_{i,j=0}^{1} |\psi(i, j)\rangle \langle \psi(i, j)| = I_4.
\]

(4.3.32)

\[\square\]

**Remark:**

- \{ |\psi(i, j)\rangle | i, j = 0, 1 \} is the Bell basis of \( \mathcal{H}_2 \otimes \mathcal{H}_2 \);
- \{ |i, j\rangle | i, j = 0, 1 \} is the product basis of \( \mathcal{H}_2 \otimes \mathcal{H}_2 \).

### 4.3.4 How to distinguish (create) \( |\psi(i, j)\rangle \).

There are two different cases

(i) Alice and Bob are in the same lab, which means that the distance between them is very close. Therefore, they can do the measurement jointly, e.g. \( X_A \otimes X_B \) and \( Z_A \otimes Z_B \).

(ii) Alice and Bob are far away from each other, which leaves them two choices:

- perform local measurement, e.g. \( X_A \otimes I_B \), \( Z_A \otimes I_B \), \( I_A \otimes X_B \), and \( I_A \otimes Z_B \).
- classical communication (phone call).

But, with only local operation (LO) and classical communication (CC), it is impossible to distinguish/create \( |\psi_{i,j}\rangle \). The reason is that local operation changes \( |\psi(i, j)\rangle \), since

\[
\left\{ \begin{array}{c}
[X \otimes I_2, Z \otimes Z] \neq 0, \\
[Z \otimes I_2, X \otimes X] \neq 0.
\end{array} \right.
\]

(4.3.33)
**Remark:** Quantum entanglement cannot be created by remote pairs by using local operation and classical communication.

### 4.4 Quantum circuit model of Bell states

#### 4.4.1 Quantum circuit model

- A Quantum Circuit model is composed of a series of single-qubit-gates $SU(2)/U(2)$ and two-qubit-gates $SU(4)/U(4)$.
- Diagrammatic representation as shown in Figure 4.3.

![](image)

**Figure 4.3:** Quantum Circuit Model of single-qubit, two-qubit, and $n$-qubit gates

#### 4.4.2 Hadamard gate: $H$

**Def 4.4.1** (Hadamard gate).

\[
H := \frac{1}{\sqrt{2}} (X + Z) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},
\]

(4.4.1)

is named as Hadamard gate.

We can easily verify the following relations by utilizing the properties of the Pauli matrices,

\[
\begin{align*}
HXH & = Z; \quad \text{(4.4.2)} \\
HZH & = X; \quad \text{(4.4.3)} \\
HYH & = -Y; \quad \text{(4.4.4)} \\
H^\dagger & = H; \quad \text{(4.4.5)} \\
H^2 & = I_2. \quad \text{(4.4.6)}
\end{align*}
\]

As we can see that Hadamard gate $H$ is not only a unitary transformation but also self-adjoint.

**Proof:** From the definition (4.4.1) of the Hadamard gate $H$, and the properties of the Pauli matrices, we can derive the properties (4.4.2)~(4.4.6) of the Hadamard gate.
(1) For E.Q. (4.4.2), we can evaluate its left-hand-side,

\[ HXH = \frac{1}{\sqrt{2}}(X + Z)X \frac{1}{\sqrt{2}}(X + Z) \]
\[ = \frac{1}{2}(X + Z)X(X + Z) \]
\[ = \frac{1}{2}(XXX + ZXX + XXZ + ZZZ) \]
\[ = \frac{1}{2}(X + Z + Z - XXZ) \]
\[ = \frac{1}{2}(X + Z + Z - X) \]
\[ = Z, \]

which is equivalent to the corresponding right-hand-side of E.Q. (4.4.2).

(2) As for E.Q. (4.4.3), we can get in the following manner

\[ HZH = \frac{1}{\sqrt{2}}(X + Z)Z \frac{1}{\sqrt{2}}(X + Z) \]
\[ = \frac{1}{2}(X + Z)Z(X + Z) \]
\[ = \frac{1}{2}(XZ + ZZ + XXZ + ZZZ) \]
\[ = \frac{1}{2}(-ZZ + X + X + Z) \]
\[ = \frac{1}{2}(-Z + X + X + Z) \]
\[ = X. \]

(3) We do the same trick as in the former two cases to derive E.Q. (4.4.4),

\[ HYH = \frac{1}{\sqrt{2}}(X + Z)Y \frac{1}{\sqrt{2}}(X + Z) \]
\[ = \frac{1}{2}(X + Z)Y(X + Z) \]
\[ = \frac{1}{2}(XY + ZY + XYZ + ZYZ) \]
\[ = \frac{1}{2}(-Y - i + i - Y) \]
\[ = -Y. \]

(4) Because the Pauli matrices \( X \) and \( Z \) are self-adjoint, therefore should be the Hadamard operator \( H \),

\[ H^\dagger = \frac{1}{\sqrt{2}}(X + Z)^\dagger \]
\[ = \frac{1}{\sqrt{2}}(X^\dagger + Z^\dagger) \]
\[ = \frac{1}{\sqrt{2}}(X + Z) \]
\[ = H. \]
(5) Still from the definition (4.4.1) of the Hadamard gate $H$, we can calculate the square of the Hadamard operator $H$.

\[
H^2 = \frac{1}{2} (X + Z)^2 \\
= \frac{1}{2} \left( X^2 + Z^2 + \{X, Z\} \right) \\
= \frac{1}{2} (I_2 + I_2 + 0) \\
= I_2.
\]

Therefore, we have shown that $H$ gate is a self-adjoint unitary $2 \times 2$ matrix. 

In the viewpoint of rotation operator, we have

\[
H = \frac{1}{\sqrt{2}} (\sigma_x + \sigma_z) \\
= \frac{1}{\sqrt{2}} (\hat{e}_x + \hat{e}_z) \cdot \hat{\sigma},
\]

namely

\[
H = \hat{n} \cdot \hat{\sigma}, \quad \text{with} \quad \hat{n} := \frac{1}{\sqrt{2}} (\hat{e}_x + \hat{e}_y) \in \mathbb{R}^3.
\]

Thus,

\[
iH = \exp \left( -i \frac{\theta}{2} \hat{n} \cdot \hat{\sigma} \right) \bigg|_{\theta = \pi},
\]

which we can put in another way

\[
iH = \mathcal{D}(\hat{n}, \pi),
\]

namely a unitary transformation induced by the rotation along $\hat{n}$ through angle $\pi$ in the three-dimensional Euclidean space. As a matter of fact, the Hadamard gate $H$ can make a unitary basis transformations between $Z$ basis and $X$ basis:

\[
\begin{cases}
H |0\rangle = |+\rangle, \\
H |1\rangle = |--\rangle,
\end{cases}
\]

and

\[
\begin{cases}
H |+\rangle = |0\rangle, \\
H |--\rangle = |1\rangle.
\end{cases}
\]

E.Q. (4.4.9) can be easily derived by utilizing the definition (4.4.1) of the Hadamard gate $H$ as shown in the following

\[
H |0\rangle = \frac{1}{\sqrt{2}} (X + Z) |0\rangle \\
= \frac{1}{\sqrt{2}} (|1\rangle + |0\rangle) \\
= |+\rangle,
\]

and

\[
H |1\rangle = \frac{1}{\sqrt{2}} (X + Z) |1\rangle \\
= \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \\
= |--\rangle.
\]
While E.Q. (4.4.10) can be deduced by applying $H$ on both sides of E.Q. (4.4.9) and utilizing E.Q. (4.4.6).

Furthermore, we can reformulate E.Q. (4.4.9) into one index form, expressed as

$$H|i\rangle = \frac{1}{\sqrt{2}} \left( (-1)^i |i\rangle + |\bar{i}\rangle \right),$$

(4.4.11)

where $i = 0, 1$.

On the other hand we can also get

$$\begin{cases} 
H|\uparrow_y\rangle = \frac{1+i}{2} |\downarrow_y\rangle, \\
H|\downarrow_y\rangle = \frac{1-i}{2} |\uparrow_y\rangle,
\end{cases}$$

(4.4.12)

also as a consequence of E.Q. (4.4.9)

$$\frac{1}{\sqrt{2}} H \left( |0\rangle \pm i |1\rangle \right) = \frac{1}{\sqrt{2}} \left( H |0\rangle \pm i H |1\rangle \right) = \frac{1}{\sqrt{2}} \left( |+\rangle \pm i |-\rangle \right) = \frac{1}{2} \left( |0\rangle + |1\rangle \pm i |0\rangle \mp i |1\rangle \right) = \frac{1}{2} \left( (1 \pm i) |0\rangle + (1 \mp i) |1\rangle \right) = \frac{1 \pm i}{2} \left( |0\rangle \mp i |1\rangle \right).$$

4.4.3 CNOT gate

Def 4.4.2 (CNOT). The Controlled-Not gate, or CNOT gate for short, is defined as

$$\text{CNOT}: \ |a\rangle \otimes |b\rangle = |a, b\rangle \mapsto |a, a \oplus b\rangle, \ a \oplus b = (a + b) \mod 2,$$

(4.4.13)

where we call $|a\rangle$ as the controlled qubit, $|b\rangle$ as the target qubit. The CNOT gate is usually represented with the diagram

$$\text{CNOT}_{ab} = \begin{array}{c}
|a\rangle \\
|b\rangle
\end{array} \begin{array}{c}
\downarrow \\
\downarrow
\end{array} \begin{array}{c}
|a\rangle \\
|b\oplus a\rangle
\end{array}.$$

Remarks:

(1) With $a = 0$, then

$$\text{CNOT}_{0b} = |0\rangle \otimes |b\rangle = |0\rangle \otimes |b\rangle.$$

(4.4.14)

With $a = 1$, then

$$\text{CNOT}_{1b} = |1\rangle \otimes |b\rangle = |1\rangle \otimes |\bar{b}\rangle.$$

(4.4.15)

It is the reason why the CNOT gate is called controlled not gate, namely if target bit is 1, we do the NOT operation.
(2) The controlled-$U$ gate is defined as

\[ CU \ket{c} \ket{t} := \ket{c} U^c \ket{t}. \quad (4.4.16) \]

\[ CU = \begin{array}{c}
\end{array} \]

See more details and properties of the controlled-$U$ gate in following section 8.12.

(3) From the definition of the CNOT gate, we can verify that

\[ (\text{CNOT})^2 = I_4, \quad (4.4.17) \]

since

\[ (\text{CNOT})^2 \ket{a, b} = \text{CNOT} \ket{a, b \oplus a} = \ket{a, (b \oplus a) \oplus a} = \ket{a, b}, \]

with $a, b = 0, 1$.

(4) The CNOT gate can be expressed as

\[ \text{CNOT} = \ket{0} \bra{0} \otimes I_2 + \ket{1} \bra{1} \otimes X. \quad (4.4.18) \]

by which we can get the matrix formalism,

\[ \text{CNOT} = \begin{pmatrix}
1 & 0 \\
0 & 0 \\
1 & 0 \\
0 & 1
\end{pmatrix} \]

namely

\[ \text{CNOT} = \begin{pmatrix}
1 \\
1 \\
1 \\
1
\end{pmatrix}. \quad (4.4.19) \]

(5) CNOT gate is the quantum analogy of the classical gate XOR (the Exclusive OR gate):

\[ \quad \begin{array}{ll}
\text{XOR}: & (a, b) \rightarrow a \oplus b, \quad \text{irreversible}, \\
\text{CNOT}: & (\ket{a}, \ket{b}) \rightarrow (\ket{a}, \ket{a \oplus b}), \quad \text{reversible}.
\end{array} \]

(6) CNOT$ \in SU(4)$. Since we have derived the matrix formalism, we can see

\[ \det \text{CNOT} = 1, \quad (4.4.20) \]

and

\[ (\text{CNOT})^\dagger = \text{CNOT}. \quad (4.4.21) \]

On the other hand, we know that $(\text{CNOT})^2 = I_4$, thus

\[ \text{CNOT}(\text{CNOT})^\dagger = \text{CNOT}(\text{CNOT})^\dagger = (\text{CNOT})^2 = I_4. \]

Therefore, we can infer that CNOT$ \in SU(4)$. 

29
Lemma 4.4.3.1. The CNOT gate has the following properties:

\[
\begin{align*}
(X \otimes X) \text{CNOT} &= \text{CNOT}(X \otimes I_2), \quad (4.4.22a) \\
(X \otimes I_2) \text{CNOT} &= \text{CNOT}(X \otimes X), \quad (4.4.22b) \\
(I_2 \otimes X) \text{CNOT} &= \text{CNOT}(I_2 \otimes X), \quad (4.4.22c) \\
(Z \otimes Z) \text{CNOT} &= \text{CNOT}(I_2 \otimes Z), \quad (4.4.22d) \\
(I_2 \otimes Z) \text{CNOT} &= \text{CNOT}(Z \otimes Z), \quad (4.4.22e) \\
(Z \otimes I_2) \text{CNOT} &= \text{CNOT}(Z \otimes I_2). \quad (4.4.22f)
\end{align*}
\]

Proof. Now, we are going to verify E.Q. (4.4.22a) - E.Q. (4.4.22f), one by one, from the expression (4.4.18) of CNOT gate.

(a) \((X \otimes X) \text{CNOT} = \text{CNOT}(X \otimes I_2)\)

\[
\begin{align*}
(X \otimes X) \text{CNOT} &= (X \otimes X) \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) \\
&= X \left( |0\rangle \langle 0| \otimes X + X |1\rangle \langle 1| \right) \otimes X^2 \\
&= |1\rangle \langle 1| X \otimes XI_2 + |0\rangle \langle 0| X \otimes I_2 I_2 \\
&= \left( |1\rangle \langle 1| \otimes X + |0\rangle \langle 0| \otimes I_2 \right) \left( X \otimes I_2 \right) \\
&= \text{CNOT}(X \otimes I_2). \quad (4.4.23)
\end{align*}
\]

(b) \((X \otimes I_2) \text{CNOT} = \text{CNOT}(X \otimes X)\)

\[
\begin{align*}
(X \otimes I_2) \text{CNOT} &= (X \otimes I_2) \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) \\
&= X \left( |0\rangle \langle 0| \otimes I_2 + X |1\rangle \langle 1| \right) \otimes X \\
&= |1\rangle \langle 1| X \otimes XI_2 + |0\rangle \langle 0| X \otimes I_2 X \\
&= \left( |1\rangle \langle 1| \otimes X + |0\rangle \langle 0| \otimes I_2 \right) (X \otimes X) \\
&= \text{CNOT}(X \otimes X). \quad (4.4.24)
\end{align*}
\]

(c) \((I_2 \otimes X) \text{CNOT} = \text{CNOT}(I_2 \otimes X)\)

\[
\begin{align*}
(I_2 \otimes X) \text{CNOT} &= (I_2 \otimes X) \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) \\
&= |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X^2 \\
&= |0\rangle \langle 0| \otimes I_2 X + |1\rangle \langle 1| \otimes XX \\
&= \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) (I_2 \otimes X) \\
&= \text{CNOT}(I_2 \otimes X). \quad (4.4.25)
\end{align*}
\]

(d) \((Z \otimes Z) \text{CNOT} = \text{CNOT}(I_2 \otimes Z)\)

\[
\begin{align*}
(Z \otimes Z) \text{CNOT} &= (Z \otimes Z) \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) \\
&= Z |0\rangle \langle 0| \otimes Z + Z |1\rangle \langle 1| \otimes ZX \\
&= |0\rangle \langle 0| \otimes Z - |1\rangle \langle 1| \otimes (-XZ) \\
&= |0\rangle \langle 0| \otimes I_2 Z + |1\rangle \langle 1| \otimes XZ \\
&= \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right) (I_2 \otimes Z) \\
&= \text{CNOT}(I_2 \otimes Z). \quad (4.4.26)
\end{align*}
\]
(e) \((I_2 \otimes Z)\text{CNOT} = \text{CNOT}(Z \otimes Z)\)

\[
(I_2 \otimes Z)\text{CNOT} = (I_2 \otimes Z)\left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right)
= |0\rangle \langle 0| \otimes Z + |1\rangle \langle 1| \otimes X Z
= |0\rangle \langle 0| Z \otimes I_2 Z + |1\rangle \langle 1| Z \otimes X I_2
= \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right)(Z \otimes Z)
= \text{CNOT}(Z \otimes Z).
\]

(f) \((Z \otimes I_2)\text{CNOT} = \text{CNOT}(Z \otimes I_2)\)

\[
(Z \otimes I_2)\text{CNOT} = (Z \otimes I_2)\left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right)
= Z |0\rangle \langle 0| \otimes I_2 + Z |1\rangle \langle 1| \otimes X
= |0\rangle \langle 0| Z \otimes I_2^2 + |1\rangle \langle 1| Z \otimes X I_2
= \left( |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X \right)(Z \otimes I_2)
= \text{CNOT}(Z \otimes I_2).
\]

There we get all the six relations \((4.4.22a)\sim(4.4.22f)\) verified. □

On the other hand, we can present these properties of the CNOT gate, namely the six relations \((4.4.22a)\sim(4.4.22f)\), in the form of quantum circuit diagram, which is shown in Figure 4.4.

![CNOT gate properties](image)

(a) \((X \otimes X)\text{CNOT} = \text{CNOT}(X \otimes I_2)\)
(b) \((X \otimes I_2)\text{CNOT} = \text{CNOT}(X \otimes X)\)

(c) \((I_2 \otimes X)\text{CNOT} = \text{CNOT}(I_2 \otimes X)\)
(d) \((Z \otimes Z)\text{CNOT} = \text{CNOT}(I_2 \otimes Z)\)

(e) \((I_2 \otimes Z)\text{CNOT} = \text{CNOT}(Z \otimes Z)\)
(f) \((Z \otimes I_2)\text{CNOT} = \text{CNOT}(Z \otimes I_2)\)

Figure 4.4: Properties of the CNOT gate \((4.4.22a)\sim(4.4.22f)\).

**Lemma 4.4.3.2.** The Hadmard gate and CNOT gate have the following connection

\[
(H \otimes H)\text{CNOT}_{12}(H \otimes H) = \text{CNOT}_{21},
\]

with CNOT\(_{12}\) and CNOT\(_{21}\) defined as

\[
\begin{align*}
\text{CNOT}_{12} & : |a, b\rangle \mapsto |a, a \oplus b\rangle, \\
\text{CNOT}_{21} & : |a, b\rangle \mapsto |a \oplus b, b\rangle.
\end{align*}
\]
Proof. The prove is very simple. We can utilize the properties of the Hadmard gate as expressed in E.Q. (4.4.2)−(4.4.6) in the following way

\[(H \otimes H)\text{CNOT}_{12}(H \otimes H) \]
\[= (H \otimes H)(|0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X)(H \otimes H) \]
\[= H|0\rangle \langle 0|H \otimes H^2 + H|1\rangle \langle 1|H \otimes HXH \]
\[= |+\rangle \langle +| \otimes (|0\rangle \langle 0| + |1\rangle \langle 1|) \]
\[= |+\rangle \langle +| \otimes (|0\rangle \langle 0| - |1\rangle \langle 1|) \]
\[= (|0\rangle \langle 0| + X \otimes |1\rangle \langle 1|) \]
\[= \text{CNOT}_{21}. \quad (4.4.31)\]

In this derivation, we have also used the following facts

\[
\begin{align*}
\text{CNOT}_{12} &= |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes X, \\
\text{CNOT}_{12} &= I_2 \otimes |0\rangle \langle 0| + X \otimes |1\rangle \langle 1|,
\end{align*}
\]

and

\[
\begin{align*}
I_2 &= |0\rangle \langle 0| + |1\rangle \langle 1|, \\
I_2 &= |+\rangle \langle +| + |\rangle \langle |, \\
Z &= |0\rangle \langle 0| - |1\rangle \langle 1|, \\
X &= |+\rangle \langle +| - |\rangle \langle |. \\
\end{align*}
\]

And we can also represent the relation (4.4.29) between the Hadmard gate and CNOT gate in the diagram formalism, see Figure 4.5.

\[
\begin{align*}
\begin{array}{c}
\text{H} \\
\text{H} \\
\text{H} \\
\text{H}
\end{array}
\end{align*}
\]

\[
\begin{align*}
\end{align*}
\]

Figure 4.5: \((H \otimes H)\text{CNOT}_{12}(H \otimes H) = \text{CNOT}_{21}\)

4.4.4 Quantum circuit model of Bell states

Thm 4.4.4.1. With the help of the Hadmard gate and CNOT gate, we can construct the Bell state \(|\psi(i, j)\rangle\) from the product state \(|j, i\rangle\)

\[|\psi(i, j)\rangle = \text{CNOT}(H \otimes I_2)|j, i\rangle, \quad \text{with } i, j = 0, 1, \quad (4.4.32)\]

which can also be expressed in the Quantum circuit model

\[
\begin{align*}
\text{phase-bit } |j\rangle & \quad \text{H} \\
\text{parity-bit } |i\rangle
\end{align*}
\]

Here we present two types of proofs.
Proof. The first type of proof.
From the index formulation of the Hadamard gate E.Q. (4.4.11),
\[ (H \otimes I_2) |j\rangle |i\rangle = \frac{1}{\sqrt{2}} \left( (-1)^j |j\rangle |i\rangle + |j\rangle |i\rangle \right). \] (4.4.33)

Following the CNOT gate, we get
\[ \text{CNOT}(H \otimes I_2) |j\rangle |i\rangle = \frac{1}{\sqrt{2}} \left( (-1)^j |j\rangle |i \oplus j\rangle + |j\rangle |i \oplus j\rangle \right). \] (4.4.34)

And considering the case \( j = 0 \) and \( j = 1 \),
\[
\begin{aligned}
  j = 0: & \quad \frac{1}{\sqrt{2}} (|0\rangle |i\rangle + |1\rangle |i\rangle) = |\psi(i, 0)\rangle \\
  j = 1: & \quad \frac{1}{\sqrt{2}} (|0\rangle |i\rangle - |1\rangle |i\rangle) = |\psi(i, 1)\rangle
\end{aligned}
\] (4.4.35)

Conclusively, we have
\[ \text{CNOT}(H \otimes I_2) |j\rangle |i\rangle = |\psi(i, j)\rangle. \] (4.4.36)

Proof. The second type of proof.
To prove E.Q. (4.4.32), we may make use of the definition (4.3.2) of the Bell states and E.Q. (4.3.3). We are going to reach our destination by two main steps.

**Step 1:** We will prove that \( |\psi(0, 0)\rangle = \text{CNOT}(H \otimes I_2) |00\rangle \). As we can show that
\[
\text{CNOT}(H \otimes I_2) |00\rangle = \text{CNOT}(H |0\rangle \otimes |0\rangle)
\]
\[
= \frac{1}{\sqrt{2}} \text{CNOT}(|+\rangle \otimes |0\rangle)
\]
\[
= \frac{1}{\sqrt{2}} \text{CNOT}(|00\rangle + |10\rangle)
\]
\[
= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)
\]
\[
= |\psi(0, 0)\rangle. \] (4.4.37)

**Step 2:** With the equation \( |\psi(i, j)\rangle = (I_2 \otimes X^t Z^j) |\psi(0, 0)\rangle \), we can get
\[
|\psi(i, j)\rangle = (I_2 \otimes X^t Z^j) |\psi(0, 0)\rangle
\]
\[
= (I_2 \otimes X^t Z^j) \left( \text{CNOT}(H \otimes I_2) |00\rangle \right)
\]
\[
= (I_2 \otimes X)^t (I_2 \otimes Z)^j \text{CNOT}(H \otimes I_2 |00\rangle)
\]
\[
= (I_2 \otimes X)^t \text{CNOT}(Z \otimes Z)^j (H \otimes I_2 |00\rangle)
\]
\[
= \text{CNOT}(I_2 \otimes X)^t (Z \otimes Z)^j (H \otimes I_2) |00\rangle,
\]
in the last two jumps, we have used the relations (4.4.22c) and (4.4.22c), which we can verified in the following context:
\[
(I_2 \otimes X)^t \text{CNOT} = (I_2 \otimes X)^{t-1} \text{CNOT}(I_2 \otimes X)
\]
\[
= \text{CNOT}(I_2 \otimes X)^t,
\]

33
and

\((I_2 \otimes Z)^j \text{CNOT} = (I_2 \otimes Z)^{j-1} \text{CNOT}(Z \otimes Z)\)

\(= \text{CNOT}(Z \otimes Z)^j.\)

And we should notice that

\((I_2 \otimes X)^j (Z^i \otimes Z)^j = Z^i \otimes X^i Z^j\)
\(= (-1)^j Z^i \otimes X^{i-1} Z^j X\)
\(= (-1)^j Z^i \otimes Z^j X^i\)
\(= (-1)^j (Z^i \otimes Z^j)(I_2 \otimes X^i). \quad (4.4.38)\)

Thus, we can infer that

\(|\psi(i, j)\rangle = (-1)^j \text{CNOT}(Z^i \otimes Z^j)(I_2 \otimes X^i)(H \otimes I_2) |00\rangle\)
\(= (-1)^j \text{CNOT}(Z^i \otimes Z^j)(H \otimes X^i) |00\rangle\)
\(= (-1)^j \text{CNOT}(Z^i \otimes Z^j)(H \otimes I_2)(I_2 \otimes X^i) |00\rangle\)
\(= (-1)^j \text{CNOT}(Z^i \otimes Z^j)(H \otimes I_2)(0i)\).

On the other hand, if we’ve noticed that

\[
\begin{align*}
H \text{Z} \text{H} & = X \\
H^2 & = I_2
\end{align*}
\]

\(\Rightarrow \text{ZH} = \text{HX}, \quad (4.4.39)\)

then

\[
\begin{align*}
Z^j \text{H} & = Z^{j-1} \text{HX} \\
& = \text{HX}^j.
\end{align*}
\]

Hence,

\(|\psi(i, j)\rangle = (-1)^j \text{CNOT}(Z^j \text{H} \otimes Z^j) |0i\rangle\)
\(= (-1)^j \text{CNOT}(HX^j \otimes Z^j) |0i\rangle\)
\(= (-1)^j \text{CNOT}(H \otimes I_2)(X^j \otimes Z^j) |0i\rangle\)
\(= (-1)^j \text{CNOT}(H \otimes I_2)(X^j |0 \otimes Z^j |i\rangle)\)
\(= (-1)^j \text{CNOT}(H \otimes I_2)(|j \otimes (-1)^j |i\rangle)\)
\(= \text{CNOT}(H \otimes I_2) |ji\rangle,\)

which is equivalent to E.Q. (4.4.32). \(\square\)
Chapter 5

No-Cloning, Dense Coding, Teleportation and Cryptography

Reference:
- [Preskill] Chapter 4: Quantum entanglement.
- [Nielsen & Chuang] Chapter 12: quantum information theory.

5.1 No-cloning theorem

Def 5.1.1 (Cloning Machine). The cloning machine is a unitary transformation $U$, which satisfies

$$U(\phi \otimes 0) = \phi \otimes \phi \quad (5.1.1)$$

for arbitrary state $\phi$. It can be represented in the diagram shown in Figure 5.1.

![Figure 5.1: Copy machine in quantum mechanics](image)

Thm 5.1.0.2 (No-cloning theorem 1). The cloning machine doesn’t exist (in Quantum Mechanics).

Proof. For simplicity, we deal with the two-dimensional Hilbert space. As the definition (5.1.1) shows, we choose the blank object to be $|0\rangle$ and the target object to be be $|0\rangle$ or $|1\rangle$, and the copy machine satisfies the following equations:

$$\begin{align*}
U|0\rangle \otimes |0\rangle &= |0\rangle \otimes |0\rangle, \\
U|1\rangle \otimes 0 &= |1\rangle \otimes |1\rangle.
\end{align*}$$

(5.1.2)
For example, the most popular two-qubit quantum gate in quantum computation is the CNOT gate, which has the property,

\[
\begin{align*}
\text{CNOT} |0\rangle \otimes |0\rangle &= |0\rangle \otimes |0\rangle, \\
\text{CNOT} |1\rangle \otimes |0\rangle &= |1\rangle \otimes |1\rangle.
\end{align*}
\] (5.1.3)

For \(\forall |\phi\rangle \in \mathcal{H}_2\), which has the form of

\[|\phi\rangle = a|0\rangle + b|1\rangle,\]

with \(|a|^2 + |b|^2 = 1\),

we obtain

\[
U |\phi\rangle \otimes |0\rangle = \begin{pmatrix} a |0\rangle \otimes |0\rangle \\
U \left( a |0\rangle + b |1\rangle \right) \otimes |0\rangle \\
a |0\rangle \otimes |0\rangle + b |1\rangle \otimes |1\rangle \end{pmatrix}.
\] (5.1.5)

However, the copy machine (5.1.1) tells us another thing:

\[
U |\phi\rangle \otimes |0\rangle = (a |0\rangle + b |1\rangle) \otimes (a |0\rangle + b |1\rangle)
\] (5.1.6)

In general, E.Q. (5.1.5) and E.Q. (5.1.6) are not consistent with one another. Therefore, the cloning machine is not available for \(\forall |\phi\rangle \in \mathcal{H}_2\).

Remarks:

- The copy machine is a non-linear process, but Quantum Mechanics respects linear super-position principle.
- The no-cloning theorem is compatible with the Heisenberg uncertainty relation. If a state can be exactly copied, then it can be exactly measured, which violates the Heisenberg uncertainty relation.

**Thm 5.1.0.3** (No-cloning theorem 2). *There is no unitary transformation (cloning machine) which can make copies on distinct non-orthogonal states.*

**Proof.** We firstly consider that case that if such a unitary transformation \(U\) exists, what we can obtain. And choose two unital state vectors \(|\phi\rangle\) and \(|\psi\rangle\): being distinct means that

\[\langle \phi | \psi \rangle \neq 1.\] (5.1.7)

While, being non-orthogonal means that

\[\langle \phi | \psi \rangle \neq 0.\] (5.1.8)

Because the copy machine \(U\) satisfies E.Q. (5.1.1), we find out

\[
\langle \phi | \psi \rangle = \langle \phi | \psi \rangle U^\dagger U |\psi\rangle |0\rangle = \langle \phi | \psi \rangle |0\rangle.
\]
thus
\[ \langle \phi | \psi \rangle = |\phi \rangle |\psi \rangle^2. \]  \hfill (5.1.9)

It's clear that the E.Q. \((5.1.9)\) will either violate the relation \((5.1.7)\) or \((5.1.8)\). Therefore, the assumption is invalid, i.e., such a unitary transformation \(U\) does not exist. \(\square\)

**Remark:** Two orthogonal states can be exactly copied:

\[
\begin{align*}
U(|0\rangle \otimes |0\rangle) &= |00\rangle, \\
U(|1\rangle \otimes |0\rangle) &= |11\rangle,
\end{align*}
\]  \hfill (5.1.10)

where \(U\) can be the CNOT gate in quantum computation.

**Thm 5.1.0.4** (No-cloning theorem 3). *There is no unitary transformation to distinguish two non-orthogonal states without disturbing them.*

This is actually the third version of the no-cloning theorem.

**Proof.** We denote two arbitrary distinct non-orthogonal normalized states with \(|\phi\rangle\) and \(|\psi\rangle\),

\[
\begin{align*}
\{ \langle \phi | \psi \rangle \neq 1, \\
\langle \phi | \psi \rangle \neq 0. 
\end{align*}
\]  \hfill (5.1.11)

Assume that the unitary transformation \(U\) can distinguish \(|\phi\rangle\) and \(|\psi\rangle\), without disturbing them, namely

\[
\begin{align*}
U |\phi\rangle \otimes |0\rangle &= |\phi\rangle \otimes |e\rangle, \\
U |\psi\rangle \otimes |0\rangle &= |\psi\rangle \otimes |f\rangle,
\end{align*}
\]  \hfill (5.1.12)

where \(|e\rangle \neq |f\rangle\), means \(|\phi\rangle\) can be distinguished from \(|\psi\rangle\). Because

\[
\begin{align*}
\left( \langle \phi \otimes |0\rangle \right) \left( |\psi \rangle \otimes |0\rangle \right) &= \left( \langle \phi \otimes |0\rangle \right) U^\dagger U \left( |\psi \rangle \otimes |0\rangle \right) \\
&= \left( \langle \phi \otimes |e\rangle \right) \left( |\psi \rangle \otimes |f\rangle \right) \\
&= \langle \phi | \psi \rangle \langle e | f \rangle,
\end{align*}
\]

which is

\[ \langle \phi | \psi \rangle = \langle \phi | \psi \rangle \langle e | f \rangle. \]  \hfill (5.1.13)

From E.Q. \((5.1.13)\) we know that either

\[ \langle \phi | \psi \rangle = 0 \]  \hfill (5.1.14)

or

\[ \langle e | f \rangle = 1 \]  \hfill (5.1.15)

must be true. Therefore we are able to distinguish two orthogonal states without disturbing them, or we cannot distinguish two non-orthogonal states without disturbing them. \(\square\)

**Remarks:**

- The no-cloning theorem means that no quantum cloning machine exists, which may be bad news to quantum mechanics, but is good to quantum information security or quantum cryptography.
- The no-cloning theorem denies the possibility for a third party to extract information from the communication between the other two parties, without being noticed (without disturbance on the communication), if they make use of the resource of non-orthogonal states. Therefore, the security of information can be ensured.
5.2 Dense coding

Dense coding is that, with the entangled resource, Alice sends Bob two classical bits of information by transmitting a qubit to Bob. Dense coding can be executed in the following manner step by step:

**Step 1: Experiment setup.**
Alice and Bob share a maximally entangled state, e.g.

\[
|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)
\]

which is the cup representation for the Bell state \(|\phi^+\rangle\) defined in (4.3.7).

**Step 2: Local unitary transformation.**
Alice chooses one of the four unitary transformation \(\{I_2, X, Z, ZX\}\) and performs it on her qubit.

\[
|\psi(0, 0)\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)
\]

\[
|\psi(00)\rangle_{AB} = \begin{cases} I_2 \end{cases}
\]

\[
|\psi(0, 1)\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)
\]

\[
|\psi(01)\rangle_{AB} = \begin{cases} Z \end{cases}
\]

\[
|\psi(1, 0)\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)
\]

\[
|\psi(1)\rangle_{AB} = \begin{cases} I_2 \otimes X \end{cases}
\]

\[
|\psi(1)\rangle_{AB} = \begin{cases} X \otimes I_2 \end{cases}
\]

\footnote{Note: The Holevo bound (Old Chapter 5.4.1, page 36, John Preskill’s lecture notes) says that without entanglement at most a classical bit information can be transmitted via sending a qubit.}
\[ |\psi(10)\rangle_{AB} = \begin{bmatrix} X \\ \hline \end{bmatrix} = \begin{bmatrix} X \\ \hline \end{bmatrix} \]

\begin{align*}
|\psi(1,1)\rangle &= \frac{1}{\sqrt{2}} \left( |01\rangle - |10\rangle \right) \\
&= I_2 \otimes XZ |\psi(0,0)\rangle_{Bob} \\
&= ZX \otimes I_2 |\psi(0,0)\rangle_{Alice}
\end{align*}

\[ |\psi(11)\rangle_{AB} = \begin{bmatrix} ZX \\ \hline \end{bmatrix} = \begin{bmatrix} XZ \\ \hline \end{bmatrix} \]

The texts “Bob” and “Alice” appearing in the above equations mean that the corresponding systems, that are “Bob” and “Alice”, on which the associated nontrivial local unitary transformations are to be performed to obtain the rightful target states.

**Step 3: Qubit transmission.**

Alice sends her qubit to Bob.

\[ \begin{bmatrix} X'Z' \\ \hline \end{bmatrix} \rightarrow \begin{bmatrix} X'Z' \\ \hline \end{bmatrix} \]

**Step 4: Bell measurements.**

Bob performs the Bell measurement:

\[ \left( X \otimes X \right) |\psi(ij)\rangle = (-1)^j |\psi(ij)\rangle, \]  

which gives the parity bit \( i \), and

\[ \left( Z \otimes Z \right) |\psi(ij)\rangle = (-1)^i |\psi(ij)\rangle \]

from which gives the phase bit \( j \). With \((i, j)\), Bob get two bits of information.

As we see that, for each two-bit \((i, j)\) that Alice wants to send to Bob, she just modifies the qubit in her system with the corresponding local unitary transformation by utilizing the protocol as illustrated in Table 5.1, and then transmits such the qubit to Bob.

**Remarks:**

- On the one hand, the word “dense” in dense coding means that sending one qubit is to transmit two classical bits.

- On the other hand, we still have that sending two qubits is to transmit two classical bits, if we think about it the following way: Alice prepares the entangled state \(|\phi^+\rangle\) and then sends one qubit to Bob, so Alice sends two qubits to Bob in the entire procedure.
Table 5.1: Dense coding

<table>
<thead>
<tr>
<th>Local unitary transf.</th>
<th>Final state</th>
<th>Two bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice</td>
<td>Bob</td>
<td>Bob</td>
</tr>
<tr>
<td>$I_2$</td>
<td>$</td>
<td>\phi^+\rangle =</td>
</tr>
<tr>
<td>$X$</td>
<td>$</td>
<td>\psi^+\rangle =</td>
</tr>
<tr>
<td>$Z$</td>
<td>$</td>
<td>\phi^-\rangle =</td>
</tr>
<tr>
<td>$ZX$</td>
<td>$</td>
<td>\psi^-\rangle =</td>
</tr>
</tbody>
</table>

5.3 Quantum teleportation


In some sense, Quantum Teleportation is a kind of inverse process of dense coding (see Table 5.2).

Alice sends two classical bits to Bob

(1st qubit) Alice $\sim\sim\sim\sim|\phi^+\rangle \sim\sim\sim\sim$ Bob (2nd qubit)

Bob gets one qubit from Alice

Table 5.2: Dense coding vs. Quantum teleportation

<table>
<thead>
<tr>
<th></th>
<th>Resource</th>
<th>Send</th>
<th>transmit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dense coding</td>
<td>$</td>
<td>\phi^+\rangle$</td>
<td>1 qubit</td>
</tr>
<tr>
<td>Quantum Teleportation</td>
<td>2 bits</td>
<td>1 qubit</td>
<td></td>
</tr>
</tbody>
</table>

**Task:** Alice wants to send an unknown qubit to Bob, who is far away from her.

**Lemma 5.3.0.1.**

$$|\psi\rangle \otimes |\phi^+\rangle = \frac{1}{2} \left( |\phi^+\rangle \otimes |\psi\rangle + (X \otimes I_2 |\phi^+\rangle) \otimes X |\psi\rangle + (Z \otimes I_2 |\phi^+\rangle) \otimes Z |\psi\rangle + (ZX \otimes I_2 |\phi^+\rangle) \otimes XZ |\psi\rangle \right)$$

(5.3.1)
which can be represented in the diagram formulism also,

\[ \begin{array}{c}
\psi \downarrow |\phi^+\rangle \\
|\psi\rangle = \frac{1}{2} \left\{ \left( \begin{array}{c}X \downarrow \not{X} + Z \downarrow \not{Z} + ZX \downarrow \not{ZX} \end{array} \right) \right\}.
\end{array} \] (5.3.2)

**Proof.** We can give an expression to unknown state \( |\psi\rangle \)

\[ |\psi\rangle := a |0\rangle + b |1\rangle, \quad \text{with } a, b \in \mathbb{C} \text{ and } a^2 + b^2 = 1. \] (5.3.3)

From the definition of the four Bell states (4.3.2), we can get

\[
\begin{align*}
|00\rangle &= \frac{1}{\sqrt{2}} \left( |\phi^+\rangle + |\phi^-\rangle \right), \\
|01\rangle &= \frac{1}{\sqrt{2}} \left( |\psi^+\rangle + |\psi^-\rangle \right), \\
|10\rangle &= \frac{1}{\sqrt{2}} \left( |\psi^+\rangle - |\psi^-\rangle \right), \\
|11\rangle &= \frac{1}{\sqrt{2}} \left( |\phi^+\rangle - |\phi^-\rangle \right).
\end{align*}
\] (5.3.4)

With these materials we can make the following derivation

\[
\begin{align*}
|\psi\rangle |\phi^+\rangle &= \frac{1}{\sqrt{2}} \left( a \langle 0 | + b \langle 1 | \right) \left( |00\rangle + |11\rangle \right) \\
&= \frac{1}{\sqrt{2}} \left[ a \langle 0 | (|00\rangle + |11\rangle) + b \langle 1 | (|00\rangle + |11\rangle) \right] \\
&= \frac{1}{\sqrt{2}} \left[ a \left( \langle 00 \rangle + \langle 01 \rangle \right) + b \left( \langle 10 \rangle + \langle 11 \rangle \right) \right] \\
&= \frac{1}{2} \left[ a \left( |\phi^+\rangle + |\phi^-\rangle \right) |0\rangle + b \left( |\psi^+\rangle - |\psi^-\rangle \right) |0\rangle \\
&\quad + a \left( |\psi^+\rangle + |\psi^-\rangle \right) |1\rangle + b \left( |\phi^+\rangle - |\phi^-\rangle \right) |1\rangle \right] \\
&= \frac{1}{2} \left[ |\phi^+\rangle \left( a \langle 0 | + b \langle 1 | \right) + |\phi^-\rangle \left( a \langle 0 | - b \langle 1 | \right) \right] \\
&\quad + |\psi^+\rangle \left( a \langle 1 | + b \langle 0 | \right) + |\psi^-\rangle \left( a \langle 1 | - b \langle 0 | \right) \right] \\
&= \frac{1}{2} \left[ |\phi^+\rangle \otimes \langle \psi \rangle + \left( Z \otimes I_2 \right) |\phi^+\rangle \otimes \langle \psi \rangle + \left( X \otimes I_2 \right) |\phi^+\rangle \otimes \langle \psi \rangle \\
&\quad + \left( Z \otimes I_2 \right) |\phi^+\rangle \otimes \langle \psi \rangle \right].
\end{align*}
\]

which is equivalent to E.Q. (5.3.1). Therefore, Lemma 5.3.0.1 has been proved. \(\square\)

**Remarks:** Magic of QM.

- With the superposition principle in QM, for one state, it can be realized by the superposition of many other states, namely Lemma 5.3.0.1 (1 \(\rightarrow\) N)
• In QM, wave function collapse due to the measurement process, i.e., measurement can extract one state from the superposition of many other states. ($N \rightarrow 1$)

The Quantum Teleportation can be accomplished with the following steps:

**Step 1: State preparation.**
Alice and Bob share the entangled state $|\phi^+\rangle$. And Alice has the unknown qubit $|\varphi\rangle_A$ in hand. This can also be represented in the form of diagram:

![Diagram](image)

**Step 2: Bell measurement by Alice.**
Alice makes joint measurement for the observables $X\otimes X$ and $Z\otimes Z$, on the composite of the subsystem $A$ and the unknown particle that Alice wants to send to Bob. The measurement results and the two-bit information associated with the measurement datum, along with post measurement states, are listed in Table 5.3.

<table>
<thead>
<tr>
<th>post-measurement state</th>
<th>$Z\otimes Z$</th>
<th>$X\otimes X$</th>
<th>two-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\phi^+\rangle$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$</td>
<td>\phi^-\rangle$</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$</td>
<td>\psi^+\rangle$</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$</td>
<td>\psi^-\rangle$</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

It is transparent to that the measurement of the observables $X\otimes X$ and $Z\otimes Z$ are equivalent to the projection measurements for the Bell measurement, defined as

$$
E_{00} := |\phi^+\rangle\langle \phi^+|, \\
E_{01} := |\phi^-\rangle\langle \phi^-|, \\
E_{10} := |\psi^+\rangle\langle \psi^+|, \\
E_{11} := |\psi^-\rangle\langle \psi^-|,
$$

which can be represented in a diagrammatic formalism shown as

$$
E_{00} = \text{Diagram}, \quad E_{01} = \text{Diagram}, \\
E_{10} = \text{Diagram}, \quad E_{11} = \text{Diagram}.
$$

42
We now utilize Lemma 5.3.0.1 which actually means

\[
|\psi\rangle \otimes |\phi^+\rangle = \frac{1}{2}\left(|\phi^+\rangle \otimes |\psi\rangle + |\phi^-\rangle \otimes (Z |\psi\rangle) + |\psi^+\rangle \otimes (X |\psi\rangle) + |\psi^-\rangle \otimes (XZ |\psi\rangle)\right).
\]  

(5.3.6)

Therefore, after the Bell measurement carried out on the qubit of Alice and the unknown state, we obtain

\[
\begin{align*}
\left( |\phi^+\rangle \langle \phi^+| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) &= \frac{1}{2} |\phi^+\rangle \otimes |\psi\rangle, \\
\left( |\phi^-\rangle \langle \phi^-| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) &= \frac{1}{2} |\phi^-\rangle \otimes Z |\psi\rangle, \\
\left( |\psi^+\rangle \langle \psi^+| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) &= \frac{1}{2} |\psi^+\rangle \otimes X |\psi\rangle, \\
\left( |\psi^-\rangle \langle \psi^-| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) &= \frac{1}{2} |\psi^-\rangle \otimes XZ |\psi\rangle.
\end{align*}
\]  

(5.3.7)

E.Q. (5.3.7) can also be rewritten in the diagram language shown in the following context:

\begin{itemize}
  \item \( \left( |\phi^+\rangle \langle \phi^+| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) = \frac{1}{2} |\phi^+\rangle \otimes |\psi\rangle \)

\begin{figure}[h]
\centering
\begin{tikzpicture}
  \node (psi) at (0,0) {$\psi$};
  \node (phi) at (0,-1) {$\phi^+$};
  \node (i2) at (1,0) {$I_2$};
  \draw[->] (psi) -- (phi);
  \draw[->] (phi) -- (i2);
  \draw[->] (i2) -- (psi);
\end{tikzpicture}
\end{figure}

(5.3.8)

\item \( \left( |\phi^-\rangle \langle \phi^-| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) = \frac{1}{2} |\phi^-\rangle \otimes Z |\psi\rangle \)

\begin{figure}[h]
\centering
\begin{tikzpicture}
  \node (psi) at (0,0) {$\psi$};
  \node (phi) at (0,-1) {$\phi^+$};
  \node (i2) at (1,0) {$I_2$};
  \draw[->] (psi) -- (phi);
  \draw[->] (phi) -- (i2);
  \draw[->] (i2) -- (psi);
\end{tikzpicture}
\end{figure}

(5.3.9)

\item \( \left( |\psi^+\rangle \langle \psi^+| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) = \frac{1}{2} |\psi^+\rangle \otimes X |\psi\rangle \)

\begin{figure}[h]
\centering
\begin{tikzpicture}
  \node (psi) at (0,0) {$\psi$};
  \node (phi) at (0,-1) {$\phi^+$};
  \node (i2) at (1,0) {$I_2$};
  \draw[->] (psi) -- (phi);
  \draw[->] (phi) -- (i2);
  \draw[->] (i2) -- (psi);
\end{tikzpicture}
\end{figure}

(5.3.10)

\item \( \left( |\psi^-\rangle \langle \psi^-| \otimes I_2 \right) \left( |\psi\rangle \otimes |\phi^+\rangle \right) = \frac{1}{2} |\psi^-\rangle \otimes XZ |\psi\rangle \)

\begin{figure}[h]
\centering
\begin{tikzpicture}
  \node (psi) at (0,0) {$\psi$};
  \node (phi) at (0,-1) {$\phi^+$};
  \node (i2) at (1,0) {$I_2$};
  \draw[->] (psi) -- (phi);
  \draw[->] (phi) -- (i2);
  \draw[->] (i2) -- (psi);
\end{tikzpicture}
\end{figure}

(5.3.11)
\end{itemize}
Step 3: Classical communication between Alice and Bob.
Alice informs her results \((i, j)\) (two-bit of information as shown in Table 5.3) to Bob.

Step 4: Unitary correction by Bob.
Bob performs unitary corrections on his particle to obtain \(|\psi\rangle\). The protocol between Alice’s measurement results and Bob’s unitary corrections shows in Table 5.4.

Table 5.4: Quantum Teleportation Protocol

<table>
<thead>
<tr>
<th>Bell measurement Alice</th>
<th>Classical communication two-bit</th>
<th>Unitary correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>\phi^+\rangle)</td>
<td>((0, 0))</td>
</tr>
<tr>
<td>(</td>
<td>\phi^-\rangle)</td>
<td>((0, 1))</td>
</tr>
<tr>
<td>(</td>
<td>\psi^+\rangle)</td>
<td>((1, 0))</td>
</tr>
<tr>
<td>(</td>
<td>\psi^-\rangle)</td>
<td>((1, 1))</td>
</tr>
</tbody>
</table>

Remarks:
- The state \(|\psi\rangle\), which Alice means to transmit to Bob, is unknown to Alice.
- No-cloning theorem [5.1.0.2] is consistent with teleportation process, since once Bob gets the state \(|\psi\rangle_B\), Alice’s state \(|\psi\rangle_A\) has been destroyed by measurement.

Quantum copy machine:
\[
U(|\phi\rangle \otimes |0\rangle) = |\phi\rangle \otimes |\phi\rangle; \quad (5.3.12)
\]
Quantum teleportation:
\[
T(|\phi\rangle \otimes |0\rangle) = |X\rangle \otimes |\phi\rangle. \quad (5.3.13)
\]

- Quantum teleportation has the interpretation of space-time topology.

5.4 The quantum teleportation using continuous variables

Problem description

One complete orthonormal basis for the Hilbert space of two particles on the real line is the (separable) position eigenstate basis \(\{|q_1\rangle \otimes |q_2\rangle\}\). Another is the entangled basis \(\{|Q, P\rangle\}\), where
\[
|Q, P\rangle = \frac{1}{\sqrt{2\pi}} \int dq \, e^{ip\cdot q}|q\rangle \otimes |q + Q\rangle; \quad (5.4.1)
\]
these are the simultaneous eigenstates of the relative position operator \(\mathbf{Q} = \mathbf{q}_2 - \mathbf{q}_1\) and the total momentum operator \(\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2\).

(a) Verify that
\[
\langle Q', P'|Q, P\rangle = \delta(Q' - Q)\delta(P' - P); \quad (5.4.2)
\]

\(^2\)Originated from the exercise 4.3, revised Chapter 4 of John Preskill’s online lecture notes.
(b) Since the states $|Q, p\rangle$ are a basis, we can expand a position eigenstate as

$$|q_1\rangle \otimes |q_2\rangle = \int dQdP |Q, P\rangle \langle Q, P| (|q_1\rangle \otimes |q_2\rangle).$$  \hfill (5.4.3)

Evaluate the coefficients $\langle Q, P| (|q_1\rangle \otimes |q_2\rangle)$.

(c) Alice and Bob have prepared the entangled state of two particles $A$ and $B$; Alice has kept particle $A$ and Bob has particle $B$; Now Alice has received an unknown single particle wavepacket $|\psi\rangle_C = \int dq|q\rangle_C C\langle q|\psi\rangle_C$ that she intends to teleport to Bob. Design a protocol that they can execute to achieve the teleportation. What should Alice measure? What information should she send to Bob? What should Bob do when he receives this information, so that particle $B$ will be prepared in the state $|\psi\rangle_B$?

1°) **Background**: Quantum teleportation is a quantum information protocol in which Alice and Bob are space-like separated, but Alice can send Bob a qubit based on the application of both quantum entanglement and quantum measurement. This protocol usually consists of the four steps:

i) State preparation

ii) Bell measurement

iii) Classical communication

iv) Unitary correction

2°) **Notation**:

$$|\Omega\rangle = \int dq|qq\rangle = |Q = 0, P = 0\rangle = \int \delta(q, 0) \delta(p, 0) dq dp,$$  \hfill (5.4.4)

Define the $U(1)$ phase operator as

$$U_P = e^{-iPq} = \int dq' e^{-iPq'} |q'\rangle \langle q'|,$$  \hfill (5.4.5)

$$U_{-P}|q\rangle = e^{iPq}|q\rangle.$$  \hfill (5.4.6)

The unitary formalism shows as

$$U_P^\dagger = U_{-P}, \quad \langle q|U_P^\dagger = \langle q|U_{-P} = \langle q|e^{iPq}.$$  \hfill (5.4.7)

Define the translation operator as

$$T_Q = e^{-iPq} = \int dq'|q' + Q\rangle \langle q'|,$$  \hfill (5.4.8)

$$T_Q|q\rangle = |q + Q\rangle.$$  \hfill (5.4.9)

The unitary formalism shows as

$$T_Q^\dagger = T_{-Q}, \quad \langle q|T_Q^\dagger = \langle q|T_{-Q} = \langle q + Q|,$$  \hfill (5.4.10)

$$\langle q| \int dq'|q' - Q\rangle \langle q'| = \int dq' \delta_{q', q + Q}\langle q'| = \langle q + Q|.$$  \hfill (5.4.11)

Note we have the relations

$$U_P T_Q|q\rangle = U_P|q + Q\rangle = e^{-iP(q + Q)}|q + Q\rangle,$$  \hfill (5.4.12)
\[ T_Q U_p |q\rangle = T_Q e^{-iP \cdot q} |q\rangle = e^{-iP \cdot q} |q + Q\rangle, \quad (5.4.13) \]

therefore
\[ U_p T_Q = e^{-iP \cdot Q} T_Q U_p. \quad (5.4.14) \]

The entangled basis is defined as
\[ |Q, P\rangle = (U_{-P} \otimes T_Q) |\Omega\rangle = \int dq e^{iP \cdot q} |q, q + Q\rangle \quad (5.4.15) \]

with the cup configuration
\[ |Q, P\rangle = \begin{array}{c}
- P \\
Q
\end{array} \quad (5.4.16) \]

And the cap configuration expresses as
\[ \langle Q, P \rangle = P \begin{array}{c}
- Q
\end{array} \quad (5.4.17) \]

The normalization relation in diagrammatical language:
\[ \langle Q', P' |Q, P\rangle = \begin{array}{c}
P' \\
Q'
\end{array} = \delta(P - P') \delta(Q - Q'). \quad (5.4.18) \]

Other diagrammatical rules:
\[ \begin{array}{c}
- P \\
Q
\end{array} = \begin{array}{c}
- P \\
Q
\end{array} = \begin{array}{c}
- Q
\end{array} = \begin{array}{c}
Q
\end{array} = \begin{array}{c}
- P \\
Q
\end{array} \quad (5.4.19) \]

The product state has the diagrammatical representation:
\[ |q_1, q_2\rangle = \begin{array}{c}
q_1 \\
q_2
\end{array} \]

The inner product of product state with the entangled state \( |Q, P\rangle \) has the expression
\[ \langle Q, P |q_1, q_2\rangle = \begin{array}{c}
P \\
q_1
\end{array} = \frac{1}{\sqrt{2\pi}} e^{-iP \cdot q_1} \delta(Q - (q_2 - q_1)). \]

Note that \( Q \) stands for the relative position of two particles, namely \( Q = q_2 - q_1 \).

3°) Continuous teleportation:
i) State preparation

\[ | \psi \rangle_C = \int dq(q) | q \rangle_C = \int dq \psi(q) | q \rangle. \quad (5.4.22) \]

where \( | \psi \rangle_C \) is the unknown state, expressed as

\[
| \psi \rangle_C = \int dq | q \rangle_C | \psi \rangle_C = \int dq \psi(q) | q \rangle.
\]

ii) Bell measurement

\[ | \psi \rangle_B = T_Q U_{-P} T_Q U_{P^*} | \psi \rangle_B. \quad (5.4.24) \]

Therefore, he is required to perform the unitary correction

\[ U^\dagger = U_{-P} T_{-Q} U_P T_{Q^*} = e^{-iP \cdot Q'} U_{-P^*} T_{-Q - Q'}, \quad (5.4.25) \]

where the commutative relation \((5.4.14)\) has been applied.

5.5 Quantum cryptography (information security)

5.5.1 Classical cryptography

Alice and Bob want to communicate (transmit information) with each other.

1. To ensure information security, they choose to get an encryption key, e.g.

\[ K := (1 \ 1 \ 1 \ 1 \ 1). \]

2. Now, Alice wants to send Bob some information, e.g.

\[ A := (0 \ 1 \ 0 \ 0 \ 0). \]

For information security consideration, she would encrypt her information before sending out, i.e.,

\[ A + K = (1 \ 0 \ 1 \ 1 \ 1). \]
3. Bob then receives the message, namely

\[ B := A + K. \]

To read the message, Bob will have to decrypt it, i.e.,

\[ B + K = A + K + K = A. \]

**Remark:** The encryption key is the most important thing for information security. In classical physics, the key can be copied without Alice and Bob’s notice, by a third party different from Alice and Bob, who we can name “Eve”.

### 5.5.2 Quantum key distribution (QKD)

In Quantum physics, if the encryption keys are encoded in non-orthogonal states, for example, \(|\uparrow_x\rangle \) and \(|\uparrow_z\rangle\), it cannot be copied without disturbing Alice and Bob, which is ensured by the non-cloning theorem. The process to conduct a Quantum key distribution is shown step by step in the following:

**Step 1: State preparation.**

Alice and Bob share the Bell state.

\[
|\phi^+\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow_z\rangle |\uparrow_z\rangle + |\downarrow_z\rangle |\downarrow_z\rangle \right)
\]

\[
= \frac{1}{2\sqrt{2}} \left( \left( |\uparrow_x\rangle + |\downarrow_x\rangle \right) \otimes \left( |\uparrow_x\rangle + |\downarrow_x\rangle \right) \\
+ \left( |\uparrow_x\rangle - |\downarrow_x\rangle \right) \otimes \left( |\uparrow_x\rangle - |\downarrow_x\rangle \right) \right)
\]

\[
= \frac{1}{\sqrt{2}} \left( |\uparrow_x\rangle \otimes |\uparrow_x\rangle + |\downarrow_x\rangle \otimes |\downarrow_x\rangle \right),
\]

i.e.,

\[
|\phi^+\rangle = \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right) = \frac{1}{\sqrt{2}} \left( |++\rangle + |--\rangle \right). \tag{5.5.1}
\]

**Step 2: Random local measurements by Alice.**

Alice makes purely random (Prob = 1/2) choices from \(\{\sigma_z, \sigma_x\}\) to conduct measurements on her qubit with the chosen observables (see Table 5.5).

<table>
<thead>
<tr>
<th>Alice</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>random prob= 1/2</td>
<td>Z</td>
<td>Z</td>
<td>X</td>
<td>X</td>
<td>Z</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>(\sigma_z^A)</td>
<td>(\sigma_z^A)</td>
<td>(\sigma_x^A)</td>
<td>(\sigma_x^A)</td>
<td>(\sigma_z^A)</td>
<td>(\sigma_x^A)</td>
</tr>
<tr>
<td>eigenstate</td>
<td>(</td>
<td>\uparrow_z\rangle)</td>
<td>(</td>
<td>\uparrow_z\rangle)</td>
<td>(</td>
<td>\uparrow_x\rangle)</td>
</tr>
<tr>
<td>eigenvalue</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
</tr>
</tbody>
</table>

**Step 3: Random local measurements by Bob.**

At the same time, Bob does the same thing on his particle as Alice does to hers system (see Table 5.6).
Step 4: Inform each other the experiments.
Alice and Bob inform each other which observable they have chosen in every experiment, but do not mention the measurement results. In this circumstance, the third party, i.e., Eve, can get the observables without notice, but cannot get the measurement results.

Step 5: Get the Key.
Alice and Bob choose the same results, which are shown in Table 5.7.

Table 5.7: The same observables that Alice and Bob share

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>measurement</td>
<td>$\sigma_z$</td>
<td>$\sigma_x$</td>
<td>$\sigma_x$</td>
</tr>
<tr>
<td>eigenvalue</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
</tr>
</tbody>
</table>

And then encode them as

$$
\begin{cases}
|\uparrow_x\rangle, |\uparrow_z\rangle \rightarrow 0 \\
|\downarrow_x\rangle, |\downarrow_z\rangle \rightarrow 1.
\end{cases}
$$

Finally, we can get the encryption key, i.e.,

$$K = (0 \; 1 \; 1)$$

Step 6: For practical consideration, we may have to develop from this row key to a practical key.

Remark: If Eve wants to know the key $K$, then he (she) must detect the states, but disturb Alice’s and Bob’s systems at the same time.

5.5.3 BB84 quantum key distribution
The BB84 Quantum key distribution was firstly presented in 1984.

Principle 5.5.3.1. Non-orthogonal state (e.g. $|\uparrow_x\rangle$, $|\uparrow_z\rangle$), cannot be distinguished without any changes, so Quantum information can be securely encoded in non-orthogonal states.

To obtain the BB84 Quantum key distribution, we have the following steps:
**Step1:** Alice randomly prepares one of the four states

\[ |\uparrow_z\rangle, |\downarrow_z\rangle, |\uparrow_x\rangle, |\downarrow_x\rangle, \]

and labels

\[ |\uparrow_z\rangle, |\downarrow_z\rangle \]

with its observable \( Z \), while

\[ |\uparrow_x\rangle, |\downarrow_x\rangle \]

with \( X \).

**Step2:** Alice sends her state to Bob. Bob then makes a random measurement on observable \( X \) or \( Z \).

**Step3:** Alice and Bob inform each other their observables, not including measurement outcomes.

**Step4:** Alice and Bob keep states in which the same observable are exploited, and keep the remaining outcomes as the raw key.

\[
\begin{cases}
|\uparrow_x\rangle, |\uparrow_z\rangle \rightarrow 0 \\
|\downarrow_x\rangle, |\downarrow_z\rangle \rightarrow 1
\end{cases}
\]

**Step5:** Develop the row key to a practical key.
Chapter 6

Bell Inequalities

...what is proved by impossibility proofs is lack of imagination.

—John Bell

The true logic of this world is in the calculus of probabilities.

—James Clerk Maxwell

Reference:

- [Preskill] Chapter 4: Quantum entanglement;

6.1 Einstein’s quantum mechanics: local hidden variable theory (LHV)

6.1.1 What hidden variable (HV) theory?

In Einstein’s opinion, quantum theory is not complete and the reason is that a complete theory should be deterministic. He explained further that quantum randomness is a result of our ignorance of local hidden variables, and the local hidden variable theory is complete. In other words, quoting the famous statement by Einstein, “God does not play dice.” We may show it in the table 6.1. But, Bohr didn’t agree with him, and argued that the quantum theory is complete, and the measurement output has to be probabilistic, which is the intrinsic character of quantum mechanics. They were the two greatest minds in the 20th century, but held opposite opinions concerning quantum mechanics. Whose opinion is right? It should be answered by the experiment, and there is no other way. Up to now, experiments always agree with Bohr’s viewpoint.

Table 6.1: QM and LHV

<table>
<thead>
<tr>
<th>Quantum theory</th>
<th>Hidden variable theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\psi\rangle = \alpha</td>
</tr>
</tbody>
</table>
6.1.2 What local theory?

As we know that one of the two axioms of Special Relativity is that there is no faster-than-light communication, which ensures the causality. We call this the “relativistic locality”. From that, we can infer that two events at space-like separated regions can not have any causal connection. Einstein thought that if two subsystems A and B are space-like separated, measurements on subsystem A cannot modify subsystem B, neither measurements on subsystem B can modify subsystem A. We call this principle as Einstein’s locality.

Einstein’s locality can be violated in quantum mechanics, under a given circumstance. For example, when the two subsystems A and B share the Bell state $|\phi^+\rangle$ \[4.3.2\], Alice measures her subsystem A along the $z$-axis, and Bob measures his subsystem B soon after Alice’s measurements, and Bob’s results will be the same as Alice’s results. Afterwards, Alice measures her subsystem A along the $x$-axis, and Bob measures his subsystem B soon after Alice’s measurements, and Bob’s results will be still the same as Alice’s results. Hence, Bob’s measurement results can be modified by Alice’s measurements. With the GHJW theorem \[14.4.2.1\] we know that if two subsystems A and B share an entangled state in the composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, local measurements on subsystem (e.g., B) can lead to different state ensembles for another subsystem (e.g., A), even if the two subsystems are space-like separated, but, which does not cause the faster-than-light information communication. Note that “information is physical”. If there are no classical communication between Alice and Bob, then Alice’s measurements do not modify the ensemble description for subsystem B. Therefore, the relativistic causality survives quantum mechanics but Einstein’s locality does not.

6.1.3 The rule to justify the rightful theory

The local hidden variable theory (LHV) contradicts with quantum mechanics, but which of these two theories is right? Of course, it should be eventually answered by experiments. How can an experiment tell us which one is right, or what kind of experiments can distinguish these two theories? Bell’s inequality is derived for this purpose, and we can directly check the Bell’s inequality in our experiment, from the result we will know which theory is telling the truth.

6.2 Bell’s inequality in the local hidden variable theory

Let’s consider the following experiment:

1. Alice has three coins, each with head or tail face. We can label the three coins, which Alice holds, with $1_A$, $2_A$ and $3_A$ respectively. For each coin we assign a specific variable to describe the state of the coin, i.e., $x$ for coin $1_A$, $y$ for coin $2_A$ and $z$ for coin $3_A$, and

$$x, y, z \in \{H, T\}, \text{ with “H” for Head, and “T” for Tail.}$$

2. The local hidden theory allows us to assign the probability distribution of the three coins faces, denoted as $\text{Prob}(x, y, z)$, with $x, y, z \in \{H, T\}$. And we shall know that the total probability should be

$$\sum_{x, y, z \in \{H, T\}} \text{Prob}(x, y, z) = 1.$$
The probability that the $i$-th coin and the $j$-th coin have the same value can be denoted as $P_{\text{same}}(i,j)$, e.g.

$$P_{\text{same}}(1,2) := \text{Prob}(H,H,H) + \text{Prob}(H,H,T) + \text{Prob}(T,T,H) + \text{Prob}(T,T,T).$$  \hspace{1cm} (6.2.1)

Now, we can define

$$\text{BI} := P_{\text{same}}(1,2) + P_{\text{same}}(1,3) + P_{\text{same}}(2,3),$$  \hspace{1cm} (6.2.2)

and evaluate it in the following way

$$\text{BI} = \text{Prob}(H,H,H) + \text{Prob}(H,H,T) + \text{Prob}(T,T,H) + \text{Prob}(T,T,T)$$
$$+ \text{Prob}(H,H,H) + \text{Prob}(H,T,H) + \text{Prob}(T,H,T) + \text{Prob}(T,T,T)$$
$$+ \text{Prob}(H,H,H) + \text{Prob}(T,H,H) + \text{Prob}(H,T,T) + \text{Prob}(T,T,T)$$
$$= 1 + 2 \left( \text{Prob}(H,H,H) + \text{Prob}(T,T,T) \right)$$  \hspace{1cm} (6.2.3)

i.e.,

$$\text{BI} = P_{\text{same}}(1,2) + P_{\text{same}}(1,3) + P_{\text{same}}(2,3) \geq 1,$$  \hspace{1cm} (6.2.4)

which is the so-called Bell’s inequality.

In the logical of probabilities distribution, we have the following Venn diagram to describe $\text{Prob}(x,y,z)$, shown in Figure 6.1. For another method of calculation BI, with the help of the diagram, we obtain

$$\text{BI} := P_{\text{same}}(1,2) + P_{\text{same}}(1,3) + P_{\text{same}}(2,3)$$
$$= (A + C) + (B + C) + (C + D)$$
$$= 1 + 2C$$
$$= 1 + 2 \left( \text{Prob}(H,H,H) + \text{Prob}(T,T,T) \right)$$
$$\geq 1,$$  \hspace{1cm} (6.2.5)

where, through derivation, we have applied the relation $A + B + C + D = 1$, due to every time at least two coins share the same face value.

Note: The Bell’s inequality actually says one simple thing, i.e., for three coins placed on the table, there would be at least two coins shown the same face up or down, in any circumstances.

(3) Mutually Exclusive Experiments (MEE) or Complementary experiment: Each time Alice is allowed to measure just one coin, namely, the other two coins are forbidden to measure. Then $x$, $y$, and $z$ are called complementary variables, which means you cannot have two values of them simultaneously.

Remark: LHV agrees with MEE.

(4) Perfect correlation.

- Alice and Bob are space-like separated.
- Bob also has three coins, which is numbered with $1_B$, $2_B$, and $3_B$.
- Alice and Bob obtain the same results when measuring their coins with the same label, i.e., when Alice measures the coin $i_A$ and Bob measures the coin $i_B$, with $i = 1, 2, 3$, then their results are always coincide.

53
Figure 6.1: Venn diagram of probabilities distribution Prob\((x, y, z)\). The circles labeled by 1, 2 and 3 represent the corresponding probabilities distribution of coin 1, 2 and 3. The overlap regions labeled by \(A, B, C, D\) stand for the probability of the coin \(i\) and \(j\) have the same face value, e.g., region \(A\) tells us that coin 1 and 2 share the same face value, whereas coin 1 and 3, coin 2 and 3, have the different values.

- This correlation can be proved by classical communications between Alice and Bob.

**Remark:** LHV agrees with the perfection correlation because Alice and Bob are space-like separated.

(5) With the perfect correlation, for a set of complementary variables, the Bell’s inequalities can be checked, and so the local hidden variable theory.

### 6.3 Bell’s inequality in quantum mechanics

(1) Experiment setup: Alice and Bob share the spin singlet state \(|\psi^\rangle\),

\[
|\psi^\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)
\]  

(6.3.1)

Here, we replaced a coin with a spin polarization direction. For example, we can choose the three orientations \(\hat{a}_i, i = 1, 2, 3\), for Alice’s system, and \(\hat{b}_i, i = 1, 2, 3\), for Bob’s system, with

\[
|\hat{a}_i| = |\hat{b}_i| = 1, \quad \text{and} \quad \hat{a}_i \neq \hat{a}_j, \quad \hat{b}_i \neq \hat{b}_j, \quad \text{if} \ i \neq j.
\]

A spin pointing along the positive (negative) direction of an orientation corresponds to a coin with “Head” (“Tail”) face.

(2) Mutually Exclusive Experiment (MEE).

Because

\[
\begin{align*}
[\vec{\sigma} \cdot \hat{a}_i, \vec{\sigma} \cdot \hat{a}_j] &\neq 0, \quad \text{if} \ i \neq j, \\
[\vec{\sigma} \cdot \hat{b}_i, \vec{\sigma} \cdot \hat{b}_j] &\neq 0, \quad \text{if} \ i \neq j,
\end{align*}
\]

which means that \(\vec{\sigma} \cdot \hat{a}_i\) and \(\vec{\sigma} \cdot \hat{a}_j\) have no simultaneous eigenstates, neither \(\vec{\sigma} \cdot \hat{a}_i\) and \(\vec{\sigma} \cdot \hat{b}_i\). As a result, there comes some differences between the quantum mechanics and the local hidden variable theory:
• **QM**: Complementary variables cannot be assigned values simultaneously, because of uncertainty relation.

• **LHV**: Complementary variables can be assigned values simultaneously, and then the classical logic can be applied, because LHV supports pre-existing variables before measurement.

(3) Perfect correlation.

• **QM**: Perfect correlation exists and agrees with the Special Relativity causality.

• **LHV**: No information transformation in space-like regions, namely no classical communication in space-like regions.

(4) Calculation on successive measurements

$$
\begin{align*}
E_A(\hat{a}) & := |\hat{a}\rangle \langle \hat{a}| \otimes I_2 := \frac{1}{2} \left( 1 + \hat{a} \cdot \sigma_A \right), \\
E_B(\hat{b}) & := I_2 \otimes |\hat{b}\rangle \langle \hat{b}| := \frac{1}{2} \left( 1 + \hat{b} \cdot \sigma_B \right), \\
\text{with } |\hat{a}| &= |\hat{b}| = 1, \text{ and } \hat{a}, \hat{b} \in \mathbb{R}^3.
\end{align*}
$$

The density matrix of the composite system is expressed as

$$\rho_{AB} := |\psi^\rangle_{AB} \langle \psi^|.$$  \hfill (6.3.3)

The probability, that Alice’s spin is pointing along the direction $\hat{a}$ and Bob’s spin along the orientation of $\hat{b}$ when we do the measurements, is calculated by

$$
\begin{align*}
\text{Prob}(\hat{a}, \hat{b}) &= AB \langle \psi^| \left( E_A(\hat{a}) E_B(\hat{b}) \right) |\psi^\rangle_{AB} \\
&= \frac{1}{4} AB \langle \psi^| \left( 1 + \hat{a} \cdot \sigma_A \right) \left( 1 + \hat{b} \cdot \sigma_B \right) |\psi^\rangle_{AB} \\
&= \frac{1}{4} AB \langle \psi^| \left( 1 + \hat{a} \cdot \sigma_A + \hat{b} \cdot \sigma_B + (\hat{a} \cdot \sigma_A)(\hat{b} \cdot \sigma_B) \right) |\psi^\rangle_{AB}. \hfill (6.3.4)
\end{align*}
$$

And we can calculate the two parts of expression separately. Firstly, for **part I**, \[part I = 1 + AB \langle \psi^| \hat{a} \cdot \sigma_A |\psi^\rangle_{AB} + AB \langle \psi^| \hat{b} \cdot \sigma_B |\psi^\rangle_{AB},\]

where

$$
\begin{align*}
&AB \langle \psi^| \hat{a} \cdot \sigma_A |\psi^\rangle_{AB} \\
&= \frac{1}{2} \left( AB \langle 01| - AB \langle 10| \right) \hat{a} \cdot \sigma_A \left( |01\rangle_{AB} - |10\rangle_{AB} \right) \\
&= \frac{1}{2} \left( A \langle 0| \hat{a} \cdot \sigma_A |0\rangle_A + A \langle 1| \hat{a} \cdot \sigma_A |1\rangle_A \right) \\
&= \frac{1}{2} \text{tr}_A (\hat{a} \cdot \sigma_A) \\
&= 0. \hfill (6.3.5)
\end{align*}
$$
because Pauli matrices are traceless. Therefore,

\[ \text{part I} = 1. \tag{6.3.7} \]

Next, we come to **part II**,\n
\[ \text{part II} = \langle \psi^- | \left( b \hat{\sigma}_A \right) \left( b \hat{\sigma}_B \right) | \psi^- \rangle_{AB} \]

\[ = \frac{1}{2} \left\langle AB \langle 01 | - AB \langle 10 | \rightangle \left( a \hat{\sigma}_A \right) \left( b \hat{\sigma}_B \right) \left( | 01 \rangle_{AB} - | 10 \rangle_{AB} \right) \]

\[ = \frac{1}{2} \left( A \langle 0 | a \hat{\sigma}_A | 0 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 1 \rangle_{B} - A \langle 0 | a \hat{\sigma}_A | 1 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 0 \rangle_{B} \right) \]

\[ = \frac{1}{2} \left( - a_{3b_{3}} - \left( a_{1} - i a_{2} \right) \left( b_{1} + i b_{2} \right) - \left( a_{1} + i a_{2} \right) \left( b_{1} - i b_{2} \right) - a_{3b_{3}} \right) \]

\[ = \frac{1}{2} \left( - a_{3b_{3}} - 2a_{1b_{1}} - 2a_{2b_{2}} - a_{3b_{3}} \right) \]

\[ = - \tilde{a} \tilde{b}. \tag{6.3.8} \]

In another method on the calculation of **part II**, from the physical point that Bell state \( | \psi^- \rangle_{AB} \) is a singlet state, namely spin-less state, we get

\[ \left( \hat{\sigma}_A + \hat{\sigma}_B \right) | \psi^- \rangle_{AB} = 0. \tag{6.3.9} \]

therefore the **part II** can be rewritten as

\[ \text{part II} = \langle \psi^- | \left( a \hat{\sigma}_A \right) \left( b \hat{\sigma}_B \right) | \psi^- \rangle_{AB} \]

\[ = - \langle AB \langle 01 | - AB \langle 10 | \rightangle \left( a \hat{\sigma}_A \right) \left( b \hat{\sigma}_B \right) \left( | 01 \rangle_{AB} - | 10 \rangle_{AB} \right) \]

\[ = - \langle AB \langle 0 | a \hat{\sigma}_A | 0 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 1 \rangle_{B} - A \langle 0 | a \hat{\sigma}_A | 1 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 0 \rangle_{B} \rangle \]

\[ = - \langle AB \langle 0 | a \hat{\sigma}_A | 0 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 1 \rangle_{B} + A \langle 0 | a \hat{\sigma}_A | 1 \rangle_{AB} \langle 1 | b \hat{\sigma}_B | 0 \rangle_{B} \rangle \]

\[ = \left( a_{3b_{3}} - \left( a_{1} - i a_{2} \right) \left( b_{1} + i b_{2} \right) - \left( a_{1} + i a_{2} \right) \left( b_{1} - i b_{2} \right) - a_{3b_{3}} \right) \]

\[ = \frac{1}{2} \left( - a_{3b_{3}} - 2a_{1b_{1}} - 2a_{2b_{2}} - a_{3b_{3}} \right) \]

\[ = - \tilde{a} \tilde{b}. \tag{6.3.10} \]

where the repeated indexes imply sum and the relation of the Pauli matrices is applied,

\[ \sigma_i \sigma_j = \delta_{ij} I_2 + i \varepsilon_{ijk} \sigma_k. \tag{6.3.11} \]

Hence, integrating the calculation of **part I** and **part II**, we get

\[ \text{Prob}(\tilde{a}, \tilde{b}) = \frac{1}{4} - \frac{1}{4} \tilde{a} \tilde{b} \]

\[ = \frac{1}{4} \left[ 1 - \cos(\tilde{a}, \tilde{b}) \right]. \tag{6.3.12} \]
Following, we obtain the probability that Alice’s spin points along \( \hat{a}_i \) and Bob’s spin points along \( \hat{b}_j \) or Alice’s spin points along \( -\hat{a}_i \) and Bob’s spin \( -\hat{b}_j \), namely measurement results are same,

\[
P_{\text{same}}^{ij} := \text{Prob}(\hat{a}, \hat{b}) + \text{Prob}(-\hat{a}, -\hat{b}) = \frac{1}{2}[1 - \cos(\hat{a}_i, \hat{b}_j)]. \tag{6.3.13}
\]

There, we can see that with \( \hat{b}_j = -\hat{a}_i \), E.Q. (6.3.13) should be

\[
P_{\text{same}}^{ij} = 1,
\]

i.e., physical systems Alice and Bob are perfectly anticorrelated. Now, we can calculate the quantity BI in Quantum Mechanics view point,

\[
BI = P_{\text{same}}^{12} + P_{\text{same}}^{13} + P_{\text{same}}^{23}, \quad \text{with } \hat{b}_i = -\hat{a}_i, \ i = 1, 2, 3. \tag{6.3.14}
\]

For the explanation of E.Q. (6.3.14), we may have some discussions.

- E.Q. (6.3.14) is derived from the Mutually Exclusive Experiments, due to the fact, in Quantum Mechanics, that we cannot get any two components of the angular momentum in different directions, simultaneously.
- With two subsystems, we still cannot find out the values of two complementary variables out of three in Mutually Exclusive Experiments, but with the help from subsystem Bob, we can obtain the inequality (6.2.4) and relation (6.3.14) from the view point of LHV and QM respectively. But, of course, when we consider all the matters in the frame of LHV or Quantum Mechanics, the results are different.
- The correlation between Alice and Bob is specified by the maximal entangled bipartite state, namely \( |\psi^-\rangle \). The reason why we have to set the constraint \( \hat{b}_i = -\hat{a}_i, \ i = 1, 2, 3 \) is to ensure that, when Alice and Bob measure the same axis, or same coin, their results are perfect correlated.

Now, we can calculate out some special cases by specifically setting the orientations of \( \hat{a}_i, \ i = 1, 2, 3 \).

**Case 1:** \( \hat{a}_i \) with different \( i \in \{1, 2, 3\} \) are mutually orthogonal, which is shown in Figure 6.2a. Therefore, we can see

\[
\cos(\hat{a}_1, \hat{b}_2) = \cos(\hat{a}_1, \hat{b}_3) = \cos(\hat{a}_2, \hat{b}_3) = 0. \tag{6.3.15}
\]

Following, for BI,

\[
BI = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2},
\]

i.e.,

\[
BI \geq 1. \tag{6.3.16}
\]

In this case, the Bell inequality (6.2.4) is correct, and it means QM and LHV are consistent with each other.
Case 2: \{\hat{a}_i, \ i = 1, 2, 3\} are placed in a counterclockwise way on a plane, as is shown in Figure 6.2.

\[
\cos(\hat{a}_1, \hat{b}_2) = \cos(\hat{a}_1, \hat{b}_3) = \cos(\hat{a}_2, \hat{b}_3) = \frac{1}{2},
\]

\[
\begin{align*}
\text{BI} & = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \\
& = \frac{3}{4},
\end{align*}
\]

namely

\[
\text{BI} \leq 1,
\]

which violates the inequality (6.2.4), which implies QM and LHV disagrees with each other at this time. The Bell’s inequality is violated in QM.

6.4 The CHSH inequality

The CHSH inequality, instead of the Bell’s inequality, is often tested in experiment. In experiments, it often uses light polarization instead of spin polarization to stand for qubit.

(1) Experimental setup, see Table 6.2

<table>
<thead>
<tr>
<th>Alice observables</th>
<th>Bob observables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\psi^\rightarrow\rangle$</td>
</tr>
<tr>
<td>$A, B$</td>
<td>$C, D$</td>
</tr>
<tr>
<td>$A = \pm 1, B = \pm 1$</td>
<td>$C = \pm 1, D = \pm 1$</td>
</tr>
</tbody>
</table>

(2) Local Hidden Variable theory

- Observables $A, B, C, D$ can be assigned values simultaneously, because they are pre-determined or pre-existing.

\[
\begin{align*}
A + B &= 0 \quad \rightarrow \quad B = -A \\
A - B &= 0 \quad \rightarrow \quad B = A \\
A - B &= 2A \quad \rightarrow \quad A + B &= 2A \quad \rightarrow \quad A + B &= \pm 2.
\end{align*}
\]

Define the new observable, $M$

\[
M := (A + B)C + (A - B)D = \pm 2.
\]
• Since HV is unknown, we only have probability distribution.

\[ M = \pm 2 \Rightarrow |\langle M \rangle| \leq |\langle M \rangle| = 2, \quad (6.4.2) \]

i.e.,

\[ |\langle M \rangle| = |\langle AC \rangle + \langle BC \rangle + \langle AD \rangle - \langle BD \rangle| \leq 2, \quad (6.4.3) \]

which is called CHSH inequality.

(3) Quantum Mechanics

We can set these four observable \( A, B, C, D \) as

\[
\begin{align*}
A &= \hat{a} \hat{\sigma}_A, \\
B &= \hat{a}' \hat{\sigma}_A, \\
C &= \hat{b} \hat{\sigma}_B, \\
D &= \hat{b}' \hat{\sigma}_B.
\end{align*}
\]

(6.4.4)

The relative positions of \( \hat{a} \) to \( \hat{a}' \) and that of \( \hat{b} \) to \( \hat{b}' \) are shown in Figure 6.3.

Figure 6.3: Choice of the operators

With the help of E.Q. (6.3.8), we can evaluate the left-hand-side of the CHSH inequality,

\[ |\langle M \rangle| = |- \cos(\hat{a}, \hat{b}) - \cos(\hat{a}', \hat{b}) - \cos(\hat{a}, \hat{b}') + \cos(\hat{a}', \hat{b}')|. \quad (6.4.5) \]

Now considering the special case depicted in Figure 6.4. We can derive

\[ \langle AC \rangle = \langle AD \rangle = \langle BC \rangle = - \langle BD \rangle = - \frac{\sqrt{2}}{2}. \quad (6.4.6) \]

Therefore,

\[ |\langle M \rangle| = 4 \times \frac{\sqrt{2}}{2} = 2 \sqrt{2} \geq 2. \quad (6.4.7) \]

It violates the CHSH inequality (6.4.3).

Remarks:

• The CHSH inequality is violated for all entangled pure states, see Chapter 4.3.4, John Preskill’s lecture notes.

59
• What is involved in our calculation is spin-$\frac{1}{2}$ system, but usually it is the photon-polarization represented for qubit utilized in experiments.

• There are disagreements on experiment tests of the CHSH inequality, see Chapter 4.3.6, John Preskill’s lecture notes.

### 6.5 Hints for the violation of the Bell inequality

The violation of the Bell inequality denies the local hidden variable theory (LHV) for Quantum Mechanics, but does favor the following two theories:

- Copenhagen’s quantum mechanics: local non-hidden variable theory. (Standard quantum mechanics in textbook)
- De Broglie-Bohm’s quantum mechanics: non-local hidden variable theory. Particle’s trajectories are influenced with each other non-locally.

### 6.6 Hardy’s theorem

**Problem description[1]**

Bob (in Boston) and Claire (in Chicago) share many identically prepared copies of the two-qubit state

$$|\psi\rangle = \sqrt{1-2x}|00\rangle + \sqrt{x}|01\rangle + \sqrt{x}|10\rangle.$$  

(6.6.1)

where $x$ is a real number between 0 and 1/2. They conduct many trials in which each measures his/her qubit in the basis ${|0\rangle, |1\rangle}$, and they learn that if Bob’s outcomes is 1 then Claire’s is always 0, and if Claire’s outcome is 1 then Bob’s is always 0.

Bob and Claire conduct further experiments in which Bob measures in the basis ${|0\rangle, |1\rangle}$ and they learn that if Bob’s outcome is 1 then Claire’s is always 0, and if Claire’s outcomes is 1 then Bob’s is always 0.

Bob and Claire conduct further experiments in which Bob measures in the basis ${|0\rangle, |1\rangle}$ and Claire measures in the orthonormal basis ${|\varphi\rangle, |\varphi^\perp\rangle}$. They discover that if Bob’s outcome is 0, then Claire’s outcome is always $\varphi$ and never $\varphi^\perp$. Similarly, if Claire

---

[1] Originated from the exercise 4.1, revised Chapter 4 of John Preskill’s online lecture notes.
measures in the basis \{|0\rangle, |1\rangle\} and Bob measures in the basis \{|\varphi\rangle, |\varphi^+\rangle\}, then if Claire’s outcome is 0, Bob’s outcome is always \varphi and never \varphi^+.

Bob and Claire now wonder what will happen if they both measure in the basis \{|\varphi\rangle, |\varphi^+\rangle\}. Their friend Albert, a firm believer in local realism predicts that it is impossible for both to obtain the outcome \varphi^+ (a prediction knows as Hardy’s theorem). Albert argues as follows:

When both Bob and Claire measures in the basis \{|\varphi\rangle, |\varphi^+\rangle\}, it is reasonable to consider what might have happened if one or the other had measured in the basis \{|0\rangle, |1\rangle\} instead.

So suppose that Bob and Claire both measure in the basis \{|\varphi\rangle, |\varphi^+\rangle\}, and that they both obtain the outcome \varphi^+. Now if Bob had measured in the basis \{|0\rangle, |1\rangle\} instead, we can be certain that his outcome would have been 1, since experiment has shown that if Bob had obtain 0 then Claire could not have obtained \varphi^+. Similarly, if Claire had measured in the basis \{|0\rangle, |1\rangle\}, then she certainly would have obtained the outcome 1. We conclude that if Bob and Claire both measured in the basis \{|0\rangle, |1\rangle\}, both would have obtained the outcome 1. But this a contradiction, for experiment has shown that it is not possible for both Bob and Claire to obtain the outcome 1 if they both measure in the basis \{|0\rangle, |1\rangle\}.

We are therefore forced to conclude that if Bob and Claire both measure in the basis \{|\varphi\rangle, |\varphi^+\rangle\}, it is impossible for both to obtain the outcome \varphi^+.

Though impressed by Albert’s reasoning, Bob and Claire decide to investigate what prediction can be inferred from quantum mechanics.

(a) Express the basis \{|\varphi\rangle, |\varphi^+\rangle\} in terms of the basis \{|0\rangle, |1\rangle\}.

(b) If Bob and Claire both measure in the basis \{|\varphi\rangle, |\varphi^+\rangle\}, what is the quantum-mechanical prediction for the probability \(P(x)\) that both obtain the outcome \varphi^+?

(c) Find the ”maximal violation” of Hardy’s theorem: show that the maximal value of \(P(x)\) is \(P[(3 - \sqrt{5})/2] = (5\sqrt{5} - 11)/2 \approx 0.0902\).

(d) Bob and Claire conduct an experiment that confirms the prediction of quantum mechanics. What was wrong with Albert’s reasoning?

1°) Motivation: Hardy’s theorem is able to make a suitable judgement on the conflict between the local hidden variable model and standard quantum mechanics, as well as the Bell inequalities do.

2°) In local hidden variable model, Bob and Charlie are space-like separated and perfect correlated.

The Box represents for the hidden variable that correlates Bob and Charlie.

- The observables \(u_b\) and \(u_c\) satisfy the experimental fact:

\[u_b u_c = 0,\]  \hspace{1cm} (6.6.2)

where \(u_b, u_c = 0, 1\).
• The observables $w_b$ and $w_c$ satisfy the experimental facts:

$$u_b = 0 \implies w_c = 0; \quad (6.6.3)$$
$$u_c = 0 \implies w_b = 0. \quad (6.6.4)$$

Conclusion:

$$u_b u_c = 0 \implies u_b = 0 \text{ or } u_c = 0 \implies w_b = 0 \text{ or } w_c = 0. \quad (6.6.5)$$

Therefore,

$$\text{Prob}(w_b \neq 0 \& w_c \neq 0 | u_b u_c = 0)_{LHV} = 0, \quad (6.6.6)$$

namely

$$\text{Prob}(w_b = 1 \& w_c = 1 | u_b u_c = 0)_{LHV} = 0. \quad (6.6.7)$$

In following, we will calculate it in quantum mechanics, and verify

$$\text{Prob}(w_b = 1 \& w_c = 1 | u_b u_c = 0)_{QM} \approx 0. \quad (6.6.8)$$

3°) In quantum mechanics, Bob and Charlie are correlated by entangled state $|\psi\rangle_{BC}$.

Bob  \hspace{1cm} |\psi\rangle_{BC} \hspace{1cm} \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \sim \si
After Bob’s measurement $U_b$, he obtains the state $|0\rangle_B$, and with certainty, Charlie find his state in $|\varphi\rangle_C$. Namely, observable value $u_b = 0$ leads to $w_c = 0$.

After Charlie’s measurement $U_c$, he obtains the state $|0\rangle_C$, and with certainty, Bob find his state in $|\varphi\rangle_B$. Namely, observable value $u_c = 0$ leads to $w_b = 0$.

4°) The task is to calculate the probability

$$\text{Prob}(w_b = 1 \& w_c = 1 | u_b u_c = 0)_{QM} = |BC(\tilde{\varphi}^i, \tilde{\psi}^i | \psi\rangle_{BC}|^2. \quad (6.6.16)$$

From

$$|0\rangle_B (0\rangle_A |\psi\rangle_{AB} = |0\rangle_B |\varphi\rangle_C, \quad (6.6.17)$$

we know the state $|\varphi\rangle$ can be expressed as

$$|\varphi\rangle = \sqrt{1 - 2x}|0\rangle + \sqrt{x}|1\rangle, \quad (6.6.18)$$

and the normalized form

$$|\tilde{\varphi}\rangle = \frac{1}{\sqrt{1 - x}} (\sqrt{1 - 2x}|0\rangle + \sqrt{x}|1\rangle). \quad (6.6.19)$$

Find the orthogonal state of $|\tilde{\varphi}\rangle$, denoted as $|\tilde{\varphi}^i\rangle$,

$$\begin{align*}
|\tilde{\varphi}\rangle &= A|0\rangle + B|1\rangle, \\
|\tilde{\varphi}^i\rangle &= B|0\rangle - A|1\rangle,
\end{align*} \quad (6.6.20)$$

where

$$A = \frac{\sqrt{1 - 2x}}{\sqrt{1 - x}}, \quad B = \frac{\sqrt{x}}{\sqrt{1 - x}}. \quad (6.6.21)$$

Rewrite it into matrix formalism

$$\begin{pmatrix} |\tilde{\varphi}\rangle \\ |\tilde{\varphi}^i\rangle \end{pmatrix} = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix}, \quad (6.6.22)$$

63
where the transformation matrix is a real orthogonal matrix, namely
\[
\begin{pmatrix}
A & B \\
B & -A
\end{pmatrix}
\begin{pmatrix}
A & B \\
B & -A
\end{pmatrix}
= \begin{pmatrix}
A^2 + B^2 & 0 \\
0 & B^2 + A^2
\end{pmatrix}
= \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
\] (6.6.23)

Therefore,
\[
\begin{pmatrix}
|0\rangle \\
|1\rangle
\end{pmatrix}
= \begin{pmatrix}
A & B \\
B & -A
\end{pmatrix}
\begin{pmatrix}
|\tilde{\varphi}\rangle \\
|\tilde{\varphi}^\dagger\rangle
\end{pmatrix},
\] (6.6.24)
\[
\begin{pmatrix}
|0\rangle \\
|1\rangle
\end{pmatrix}
= A|\tilde{\varphi}\rangle + B|\tilde{\varphi}^\dagger\rangle,
\] (6.6.25)

Consider the term \(|\tilde{\varphi}^\dagger\rangle_A|\tilde{\varphi}^\dagger\rangle_B\) solely in state \(|\psi\rangle_{AB}\):
\[
|00\rangle_{AB} \propto B^2|\tilde{\varphi}^\dagger\rangle_A|\tilde{\varphi}^\dagger\rangle_B; \quad (6.6.26)
\]
\[
|01\rangle_{AB} \propto -AB|\tilde{\varphi}^\dagger\rangle_A|\tilde{\varphi}^\dagger\rangle_B; \quad (6.6.27)
\]
\[
|10\rangle_{AB} \propto -AB|\tilde{\varphi}^\dagger\rangle_A|\tilde{\varphi}^\dagger\rangle_B. \quad (6.6.28)
\]

Hence
\[
|\psi\rangle_{AB} \propto (\sqrt{1-2x}B^2 - 2\sqrt{x}AB)|\tilde{\varphi}^\dagger\rangle_B|\tilde{\varphi}^\dagger\rangle_B = -\sqrt{1-2x} \frac{x}{1-x} |\tilde{\varphi}^\dagger\rangle_B|\tilde{\varphi}^\dagger\rangle_B. \quad (6.6.29)
\]

\[
\text{Prob}(w_b = 1 & w_c = 1|u_bu_c = 0)_{QM} = \left|BC(\tilde{\varphi}^\dagger|\tilde{\varphi}^\dagger\rangle_{BC}\right|^2
= (1 - 2x) \frac{x^2}{(1-x)^2}
= p(x). \quad (6.6.30)
\]

Find the maximal value of function \(p(x)\):
\[
p'(x) = \frac{2x}{(1-x)^3}(x^2 - 3x + 1) = 0. \quad (6.6.31)
\]

And
\[
x_\pm = \frac{3 \pm \sqrt{5}}{2}, \quad x_0 = 0. \quad (6.6.32)
\]

Due to \(0 < x < \frac{1}{2}\), we have
\[
p_{\text{max}}(x_-) = \frac{5\sqrt{5} - 11}{2} \approx 0.0902 \neq 0. \quad (6.6.33)
\]

The result is conflicted with local hidden variable theory.

5°) Reason for conflict between quantum mechanics and local hidden variable theory:
\[
[U_b, W_b] \neq 0, \quad [U_c, W_c] \neq 0. \quad (6.6.34)
\]

Due to uncertainty relation, \(u_b, w_b\) and \(u_c, w_c\) can not be determined simultaneously.

Note:
\[
[X, Z] \neq 0 \Rightarrow [X \otimes X, Z \otimes Z] = 0. \quad (6.6.35)
\]
Table 6.3: A comparison between Bell’s inequalities, Hardy’s theorem and GHZ theorem

<table>
<thead>
<tr>
<th>Parties</th>
<th>Formalism</th>
<th>Correlation</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell inequalities</td>
<td>2</td>
<td>Inequality</td>
<td>Statistical</td>
</tr>
<tr>
<td>Hardy’s theorem</td>
<td>2</td>
<td>Equality</td>
<td>Statistical</td>
</tr>
<tr>
<td>GHZ theorem</td>
<td>3</td>
<td>Equality</td>
<td>Perfect</td>
</tr>
</tbody>
</table>

6.7 The GHZ theorem

1°) A comparison between Bell’s inequalities, Hardy’s theorem and GHZ theorem is shown in Table 6.3

2°) GHZ theorem (the 3-qubit system)

**Step I:** Alice, Bob and Charlie are space-like from each other and share the GHZ state, i.e.,

\[ |\text{GHZ}\rangle_3 = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle), \]  

(6.7.1)

as shown in the diagram formalism in Figure 6.5.

![Diagram](Figure 6.5: 3-qubit GHZ state.)

**Step II:** Chose four observables with the form of

\[ \begin{align*}
    A &:= \sigma_x^{(1)} \sigma_y^{(2)} \sigma_y^{(3)}; \\
    B &:= \sigma_y^{(1)} \sigma_x^{(2)} \sigma_y^{(3)}; \\
    C &:= \sigma_y^{(1)} \sigma_y^{(2)} \sigma_x^{(3)}; \\
    D &:= \sigma_x^{(1)} \sigma_x^{(2)} \sigma_x^{(3)}. 
\end{align*} \]

(6.7.2a,b,c,d)

For these four observables, we may notice the following things:

- \( A^2 = B^2 = C^2 = D^2 = I_8; \)
- \( A, B, C, D \) are mutually commutative;
- \( ABCD = -I_8; \)
- \( D = X_1X_2X_3, \) i.e. phase-bit operator,
  \(-AD = -\langle \sigma_x \rangle^2 \otimes (\sigma_y \sigma_x) \otimes (\sigma_y \sigma_x) = I_2 \otimes Z \otimes Z, \) i.e. the second parity-bit operator,
  \(-CD = -\langle \sigma_y \sigma_x \rangle \otimes (\sigma_y \sigma_x) \otimes (\sigma_x)^2 = Z \otimes Z \otimes I_2, \) i.e. the first parity-bit operator;
\[
\begin{align*}
A|\text{GHZ}\rangle_3 &= -|\text{GHZ}\rangle_3, \\
B|\text{GHZ}\rangle_3 &= -|\text{GHZ}\rangle_3, \\
C|\text{GHZ}\rangle_3 &= -|\text{GHZ}\rangle_3, \\
D|\text{GHZ}\rangle_3 &= |\text{GHZ}\rangle_3,
\end{align*}
\] (6.7.3a)

since
\[
\begin{align*}
A &= X_1 \otimes (-iZ_2 X_2) \otimes (-iZ_3 X_3) = -X_1 \otimes (Z_2 X_2) \otimes (Z_3 X_3), \\
B &= (-iZ_1 X_1) \otimes X_2 \otimes (-iZ_3 X_3) = -(Z_1 X_1) \otimes X_2 \otimes (Z_3 X_3), \\
C &= (-iZ_1 X_1) \otimes (-iZ_2 X_2) \otimes X_3 = -(Z_1 X_1) \otimes (Z_2 X_2) \otimes X_3, \\
D &= X_1 \otimes X_2 \otimes X_3.
\end{align*}
\]

**Step III:** For LHV hypothesis, \(A, B, C, D\) can be exactly determined, i.e.
\[
\begin{align*}
A &= m_x^{(1)} m_y^{(2)} m_y^{(3)}; \quad (6.7.4a) \\
B &= m_y^{(1)} m_x^{(2)} m_y^{(3)}; \quad (6.7.4b) \\
C &= m_y^{(1)} m_y^{(2)} m_x^{(3)}; \quad (6.7.4c) \\
D &= m_x^{(1)} m_x^{(2)} m_x^{(3)}; \quad (6.7.4d)
\end{align*}
\]

with
\[
m_x^{(i)}, m_y^{(j)} = \pm 1, \quad i, j = 1, 2, 3. \quad (6.7.5)
\]

Thus,
\[
ABCD = \left( m_x^{(1)} m_y^{(1)} m_x^{(2)} m_y^{(2)} m_x^{(2)} m_x^{(2)} m_y^{(3)} m_x^{(3)} \right)^2 = 1. \quad (6.7.6)
\]

**Step IV:** While for QM, we can get from E.Q. (6.7.3a) \(\sim\) (6.7.3d) that
\[
ABCD|\text{GHZ}\rangle = -|\text{GHZ}\rangle, \quad (6.7.7)
\]
i.e.
\[
ABCD = -1. \quad (6.7.8)
\]

**Step V:** Experiment support E.Q. (6.7.8), which means LHV is denied.

**Remark:** In QM, \(\sigma^{(1)} \sigma^{(3)} \neq 0\), with \(i = 1, 2, 3\), thus \(m_x^{(i)}\) and \(m_y^{(i)}\) cannot be determined simultaneously.
Part II

Quantum Computing and Quantum Algorithm
Chapter 7

Classical Circuit Model

Computers are physical objects, and computations are physical processes. What computers can or can not compute is determined by the laws of physics alone, and not by pure mathematics.

—David Deutsch

By raising these issues we wish to introduce the question of the completeness of the quantum circuit model, and reemphasize the fundamental point that information is physical.

—Nielsen & Chuang

Whether physically reasonable models of computation exist, which beyond the quantum circuit model is a fascinating question which we leave open for you.

—Nielsen & Chuang

A detailed examination and attempted justification of the physics underlying the quantum circuit model is outside of the scope of the present discussions, and indeed outside the scope of the present knowledge.

—Nielsen & Chuang

In our attempts to formulate the models of information processing, we should always attempt to go back to fundamental physical laws.

—Nielsen & Chuang

Reference:

- [Preskill] New Chapter 5: Classical circuit and quantum circuit;
- [Nielsen & Chuang] Chapter 3: Introduction to computer science;
- [Nielsen & Chuang] Chapter 4: Quantum circuits.

7.1 Classical circuit

Def 7.1.1. Classical circuit, a circuit model of classical computation, is a finite sequence of elementary gates applied to a finite string of input bits.
7.1.1 Elementary logical gates

(1) NOT gate:

\[
\text{NOT: } x \mapsto \bar{x} = 1 - x, \quad (7.1.1)
\]

from which we shall see that

\[
\text{NOT} \circ \text{NOT} = \text{Id}, \quad (7.1.2)
\]

i.e. “NOT” gate is a reversible gate. The truth table of the NOT gate is shown in Table[7.1]

<table>
<thead>
<tr>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.1: Truth table of NOT gate.

(2) AND gate:

\[
\text{AND: } (x, y) \mapsto (x \cdot y) \mod 2 = x \land y. \quad (7.1.3)
\]

The “AND” gate is irreversible. Its truth table is Table[7.2]

<table>
<thead>
<tr>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0</td>
<td>0</td>
</tr>
<tr>
<td>0 1</td>
<td>0</td>
</tr>
<tr>
<td>1 0</td>
<td>0</td>
</tr>
<tr>
<td>1 1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.2: Truth table of AND gate.

(3) OR gate:

\[
\text{OR: } (x, y) \mapsto x + y - xy = x \lor y. \quad (7.1.4)
\]

As we can see that “OR” gate is irreversible. The corresponding truth table is Table[7.3]

<table>
<thead>
<tr>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0</td>
<td>0</td>
</tr>
<tr>
<td>0 1</td>
<td>1</td>
</tr>
<tr>
<td>1 0</td>
<td>1</td>
</tr>
<tr>
<td>1 1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.3: Truth table of OR gate.
7.1.2 Universal gate set

**Def 7.1.2.** An universal gate is a finite set of elementary gates which suffice to evaluate any function of a finite number of input bits.

**Note:** For a general function:

\[
F : \{0, 1\}^n \rightarrow \{0, 1\}^n \quad \Rightarrow \quad F = (f_1, f_2, \ldots, f_n),
\]

i.e., the general function can be represented by typical functions.

**Examples:**

- Thm1: \{NOT, AND\} is an universal gate set;
- Thm2: \{NOT, OR\} is an universal gate set;
- Thm3: \{NAND, COPY\} is an universal gate set;

**Remark:**

\[
\text{NAND}:= \text{NOT} \circ \text{AND}, \quad \text{(7.1.5)}
\]

and

\[
\text{COPY} : x \mapsto (x, x). \quad \text{(7.1.6)}
\]

Since

\[
\text{NOT}(x) = 1 - x = 1 - x^2 = 1 - \text{AND}(x) = \text{NOT} \circ \text{AND}(x, x), \quad \text{(7.1.7)}
\]

thus

\[
\text{NOT}(x) = \text{NAND}(x, x) = \text{NAND} \circ \text{COPY}(x). \quad \text{(7.1.8)}
\]

- Thm4: \{NOR, COPY\} is an universal gate set.

\[
\text{NOR}:= \text{NOT} \circ \text{OR}. \quad \text{(7.1.9)}
\]

7.2 Reversible classical computation

7.2.1 Irreversible computation

Classical computation in terms of irreversible gates, such as AND, OR, NAND, NOR gates, is irreversible.

**Thm 7.2.1.1** (Landauer’s principle). *Irreversible logic gates erase information with irreducible expenditure of power.*

This means that if we want to erase information, we have to pay for it with work (energy), i.e., energy loss. On the other hand, the reversible gates would ensure that there is no expenditure of power, i.e., there is no energy loss.

Quantum Computer is reversible, which is one of the reasons people prefer Quantum Computer.
Table 7.4: Irreversible and reversible computation

<table>
<thead>
<tr>
<th>Computer</th>
<th>Gates</th>
<th>Erasure of information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irreversible</td>
<td>Irreversible gates</td>
<td>Irreducible expenditure</td>
</tr>
<tr>
<td>Reversible</td>
<td>Reversible gates</td>
<td>No expenditure</td>
</tr>
</tbody>
</table>

7.2.2 Classical reversible gate

**One-bit reversible gates:**

\[
\begin{align*}
\text{NOT}(x) &= 1 - x = \bar{x}, \quad \text{NOT gate}, \\
\text{Id}(x) &= x, \quad \text{Identity gate}.
\end{align*}
\]

**Two-bit reversible gate:** XOR, i.e., Exclusive OR gate. The quantum analogy of XOR is the CNOT gate.

\[
\text{XOR}: \quad (x, y) \mapsto (x, x \oplus y). \quad (7.2.1)
\]

With three XOR gate, we can form a swap gate, which is defined as

\[
\text{SWAP}: \quad (x, y) \mapsto (y, x). \quad (7.2.2)
\]

And the construction should be

\[
\text{SWAP} = \text{XOR}_{12} \circ \text{XOR}_{21} \circ \text{XOR}_{12}. \quad (7.2.3)
\]

And, we can verify that,

\[
\begin{align*}
\text{XOR}_{12} \circ \text{XOR}_{21} \circ \text{XOR}_{12}(x, y) &= \text{XOR}_{12}(x, x \oplus y) \\
&= \text{XOR}_{12}(x, x \oplus y) \\
&= (x \oplus x \oplus y, x \oplus y) \\
&= (y, x).
\end{align*}
\]

Similarly, the quantum analogy should be

\[
\text{SWAP} = \text{CNOT}_{12} \circ \text{CNOT}_{21} \circ \text{CNOT}_{12}, \quad (7.2.4)
\]

with \(\text{SWAP}\) defined as

\[
\text{SWAP}: \quad |\psi\rangle \otimes |\phi\rangle \mapsto |\phi\rangle \otimes |\psi\rangle. \quad (7.2.5)
\]

And in quantum circuit model, it has the diagrammatical expression

\[
\begin{aligned}
|\psi\rangle \quad \quad \quad \quad \quad |\phi\rangle \\
\downarrow \quad \quad \quad \quad \quad \downarrow \\
|\phi\rangle \quad \quad \quad \quad \quad |\psi\rangle
\end{aligned}
\]

**Thm 7.2.2.1.** One-bit gates and two-bit gates cannot suffice universal classical reversible computation.
7.2.3 Three-bit Toffoli gate

Def 7.2.1 (Three-bit Toffoli gate).

Toffoli gate $\theta^{(3)}$: $(x, y, z) \rightarrow (x, y, z \oplus xy)$.  \(\text{(7.2.6)}\)

which can also be shown in the diagram formalism,

\[
\begin{array}{c}
\text{x} \\
\text{y} \\
\text{z} \\
\end{array}
\begin{array}{c}
\oplus \\
\oplus \\
\end{array}
\begin{array}{c}
x \\
y \\
z \oplus xy.
\end{array}
\]

As we can see from the definition of the 3-bit Toffoli gate, it works as the controlled-controlled NOT gate with two control-bit and one target bit, besides it is a nonlinear gate.

Thm 7.2.3.1. The 3-bit Toffoli gate with constant bits is universal for classical reversible computation.

Proof. The strategy is by setting constant bits for the Toffli gate, we can obtain an universal gate set made up of two-bit gates and one-bit gates.

1°) Toffli gate with $z = 1$:

\[
\begin{align*}
\theta^{(3)}(x, y, 1) \\
= (x, y, 1 - xy) \\
= (x, y, \text{NAND}(x, y))
\end{align*}
\]

here we get the NAND gate;

2°) Toffli gate with $x = 1, z = 0$:

\[
\begin{align*}
\theta^{(3)}(1, y, 0) \\
= (1, y, 0 + y) \\
= (1, y, y) \\
= (1, \text{COPY}(y))
\end{align*}
\]

the COPY gate for this time;

3°) Toffli gate with $x = 1$:

\[
\begin{align*}
\theta^{(3)}(1, y, z) \\
= (1, y, z \oplus y) \\
= (1, \text{XOR}(y, z))
\end{align*}
\]

and we get the XOR gate.

Actually, the NAND gate and COPY gate together can make an universal gate set, hence they can construct a universal reversible computer, i.e., the 3-bit Toffoli gate with constant bits is universal for classical reversible computation, since the 3-bit Toffoli gate is reversible:

\[
\left[\theta^{(3)}\right]^2(x, y, z)
\]

\[
= \theta^{(3)}(x, y, z \oplus xy)
\]

\[
= (x, y, z \oplus xy \oplus xy)
\]

\[
= (x, y, z),
\]
namely
\[
\left[ \theta^{(3)} \right]^2 = \text{Id.} \tag{7.2.10}
\]

**Remark**: To construct given classical gate, one may need the exponential number of \(\theta^{(3)}\) gates, which is terrible in practice.

### 7.2.4 Three-bit Fredkin gate

**Def 7.2.2** (Three-bit Fredkin gate). *The three bit Fredkin gate is defined as*

\[
\text{Fredkin gate: } (x, y, z) \mapsto (x, xz + \overline{xy}, xy + \overline{z}) \tag{7.2.11}
\]

This definition shows that the three-bit Fredkin gate is nonlinear, too.

**Thm 7.2.4.1.** *The three-bit Fredkin gate with constant bits is universal for classical reversible computation.*

**Proof.** The strategy is the same as the proof that three-bit Toffoli gate is universal for the classical reversible computation.

1°) Fredkin gate with \(z = 0\):

\[
\text{Fredkin}(x, y, 0) = (x, \overline{xy}, xy) = (x, \overline{xy}, \text{AND}(x, y)), \tag{7.2.12}
\]

here we get the AND gate;

2°) Fredkin gate with \(y = 0, z = 1\):

\[
\text{Fredkin}(x, 0, 1) = (x, x, \overline{x}) = (\text{COPY}(x), \text{NOT}(x)), \tag{7.2.13}
\]

and we obtain here the COPY gate and NOT gate;

3°) Fredkin gate with \(x = 1\):

\[
\text{Fredkin}(1, y, z) = (1, z, y) = (1, \text{SWAP}(y, z)), \tag{7.2.14}
\]

and we get the SWAP gate.

Because the NAND gate can be constructed by using the NOT gate and the AND gate, and the NAND gate and COPY gate can make an universal gate set. Therefore, the AND
gate, the NOT gate and the COPY gate together make an universal gate set. On the other hand, the Fredkin gate is reversible too, since

\[
\text{Fredkin}^2(x, y, z) = \text{Fredkin}(x, xz + \bar{y}, xy + \bar{z}) \\
= (x, x(xy + \bar{z}) + \bar{x}(xz + \bar{y}), x(xz + \bar{y}) + \bar{x}(xy + \bar{z})) \\
= (x, x^2y + x\bar{z}z + \bar{x}xz + \bar{x}^2y, x^2z + x\bar{y} + \bar{x}zy + \bar{x}^2z) \\
= (x, xz + 0 + 0 + \bar{x}y, xz + 0 + 0 + \bar{x}z) \\
= (x, (x + \bar{x})y, (x + \bar{x})z) \\
= (x, y, z),
\]
i.e.,

\[
\text{Fredkin}^2 = \text{Id}.
\]

(7.2.15)

Hence, the three-bit Fredkin gate with constant bits is universal for classical reversible computation.

7.3 The construction of an \(n\)-bit Toffoli gate using the 3-bit Toffoli gate

\textbf{Def 7.3.1 (n-bit Toffoli gate).} 

\[
\theta^{(n)}(x_1, x_2, \cdots, x_{n-1}, y) = (x_1, x_2, \cdots, x_{n-1}, y \oplus x_1x_2\cdots x_{n-1})
\]

which can also be shown in the diagram formalism,

\[
\begin{array}{cccccccc}
  & x_1 & \quad & & \quad & x_1 \\
  & & \quad & \quad & \quad & \quad & \\
  & x_2 & \quad & \quad & \quad & x_2 \\
  & & \quad & \quad & \quad & \quad & \\
  & & \quad & \quad & \quad & \\
  & \vdots & \quad & \quad & \quad & \vdots \\
  & x_{n-1} & \quad & \quad & \quad & x_{n-1} \\
  & & \quad & \quad & \quad & \quad & \\
  & x_n & \quad & \quad & \quad & y \oplus x_1x_2\cdots x_{n-1}
\end{array}
\]

which has \(n-1\) control bits and one target bit.

\textbf{Thm 7.3.0.2.} With only one bit of scratch space, performing \(\theta^{(n)}\) gate needs at least number of \(2^{n-3} + 2^{n-2} - 2\) of \(\theta^{(3)}\) gates.

\textbf{Remark:} With more bit of scratch space, the number of \(\theta^{(3)}\) needed to perform \(\theta^{(n)}\) can be reduced in polynomial scale.

\textbf{Thm 7.3.0.3.} With \(n-3\) scratch bits returning the initial value 0 at the end of computation, the number of \(\theta^{(3)}\) gates needed to perform \(\theta^{(n)}\) gate is \(2n - 5\).

\textbf{Proof.} The proof is done by induction. When \(n = 4\), we need 1 scratch bit and \((2n - 5) |_4 = 3\) number of \(\theta^{(3)}\) gates.
To prevent possible notation confusion and to maintain the notation consistence in following, in above quantum circuit diagram, we apply the notation of $\theta(3)$ gate as a three-qubit gate. And the new constructed gate functional as the $\theta(4)$ gate, we will denote it as $\theta'(4)$. Note gate $\theta'(4)$ is a 5-bit gate with scratch bit.

Assume we can construct $\theta(n)$ gate with $n - 3$ scratch bits and $(2n - 5)$ number of $\theta(3)$ gates. The gate we have constructed to perform $n$-bit Toffoli gate is denoted as $\theta(n)$. In the manner of recursive construction, to construct $\theta(n + 1)$ gate, we have

$$
\begin{array}{ccc}
\quad & \theta(3) & \quad \\
\quad & x_1 & \quad \\
\quad & x_2 & \quad \\
\quad & 0 & \quad \\
\quad & x_3 & \quad \\
\quad & 0 & \quad \\
\quad & x_4 & \quad \\
\quad & 0 & \quad \\
\quad & x_5 & \quad \\
\quad & \vdots & \quad \\
x_n & \theta'(n) & x_n \\
y & \quad & y \oplus x_1 x_2 x_3 \ldots x_n
\end{array}
$$

Therefore, the number of scratch space needed to construct $\theta(n + 1)$ gate is $(n - 3) + 1 = (n + 1) - 3$. And the number of $\theta(3)$ gates is $2n - 5 + 2 = 2(n + 1) - 5$. \hfill \Box

**Thm 7.3.0.4.** With $n - 3$ scratch bits returning the initial value at the end of computation, the number of $\theta(3)$ gates needed to perform $\theta(n)$ gate is $4n - 12$.

**Proof.** When $n = 4$, we need 1 scratch bit and $(4n - 12) |_4 = 4$ number of $\theta(3)$ gates.

After the first $\theta'(4)$ gate, we have

$$
\begin{align*}
y' &= y \oplus (x_1 x_2 \oplus s_1) x_3 = y \oplus x_1 x_2 x_3 \oplus s_1 x_3. 
\end{align*}
$$

After the second $\theta(3)$ gate, we obtain

$$
\begin{align*}
y'' &= y \oplus x_1 x_2 x_3. 
\end{align*}
$$

In the case of construction $\theta(n)$ gate, we can find we need one $\theta'(n)$ gate and one $\theta'(n-1)$
gate, shown as

After the first $\theta'(n)$ gate, we have

$$y' = y \oplus (((x_1x_2 \oplus s_1)x_3 \oplus s_2)x_4 \oplus s_3)x_5 \oplus \cdots)x_{n-1}$$
$$= y \oplus x_1x_2x_3\cdots x_{n-1} \oplus (((s_1x_3 \oplus s_2)x_4 \oplus s_3)x_5 \oplus \cdots)x_{n-1}$$
$$= y \oplus x_1x_2x_3\cdots x_{n-1} \oplus z.$$  \hspace{1cm} (7.3.4)

Note the redundant term $z$ can be rewritten as

$$z = (((x'_1x'_2 \oplus s'_1)x'_3 \oplus s'_2)x'_4 \oplus s'_3)x'_5 \oplus \cdots)x'_{n-2}.$$  \hspace{1cm} (7.3.5)

Therefore, one more $\theta'(n-1)$ gate can wipe it out,

$$y'' = y' \oplus z = y \oplus x_1x_2x_3\cdots x_{n-1}.$$  \hspace{1cm} (7.3.6)

We used totally $n - 3$ scratch bits and $(2n - 5) + (2(n - 1) + 5) = 4n - 12$ number of $\theta^{(3)}$ gates here.
Chapter 8

Quantum Circuit Model

8.1 Definition of quantum circuit

Def 8.1.1 (Quantum Circuit). Quantum Circuit, the circuit model of quantum computation, is a sequence of a finite number of Quantum gates acting on a finite number of qubits.

As we shall see that this definition is similar to the definition of Classical Circuit, but the following results can be very different. A comparison of Classical Circuit and Quantum Circuit is shown in Table 8.1.

<table>
<thead>
<tr>
<th>Computation</th>
<th>Note</th>
<th>Object</th>
<th>Operation</th>
<th>Reversibility</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Computation</td>
<td>n-bit</td>
<td>logic gate:</td>
<td>irreversible (most cases)</td>
<td>reversible</td>
<td>P vs. NP</td>
</tr>
<tr>
<td></td>
<td>({0,1}^n)</td>
<td>AND, OR, NOT,etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantum Computation</td>
<td>n-qubit</td>
<td>quantum gate:</td>
<td>reversible</td>
<td></td>
<td>BPP vs. BQP</td>
</tr>
<tr>
<td></td>
<td>(\mathbb{C}^2)</td>
<td>(U(2^n)) group</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8.1.1 One-qubit gates

The one qubit gates are \(U(2)\) or \(SU(2)\) operators. And we shall see that \(U(2)\) operator is equivalent to \(SU(2)\) operator up to a phase coefficient. There are some examples of the one-qubit gate:

- Hadamard gate \(H\),
  \[
  \begin{pmatrix}
  H \\
  \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix}
  1 & 1 \\
  1 & -1
  \end{pmatrix};
  \] (8.1.1)

- Phase gate \(S\),
  \[
  \begin{pmatrix}
  S \\
  \end{pmatrix} = \begin{pmatrix}
  1 & 0 \\
  0 & i
  \end{pmatrix};
  \] (8.1.2)

- \(\frac{\pi}{8}\) gate \(T\):
  \[
  \begin{pmatrix}
  T \\
  \end{pmatrix} = \begin{pmatrix}
  1 & 0 \\
  0 & e^{i\pi/4}
  \end{pmatrix} = e^{-i\pi/8} \begin{pmatrix}
  e^{i\pi/8} & 0 \\
  0 & e^{i\pi/8}
  \end{pmatrix};
  \] (8.1.3)
Pauli $X$ gate:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad (8.1.4)$$

Pauli $Z$ gate:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad (8.1.5)$$

Phase-shift gate $R_\theta$:

$$R_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}; \quad (8.1.6)$$

Thm 8.1.1.1. Any $U(2)$ group element can be expressed as

$$U = e^{i\alpha} D_z(\beta) D_y(\gamma) D_z(\delta), \quad (8.1.7)$$

where $D_z(\beta)$, $D_y(\gamma)$, and $D_z(\delta)$ are rotation gates defined as

$$D_z(\beta) := e^{-i\sigma_z \beta/2} = \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix}; \quad (8.1.8a)$$

$$D_y(\gamma) := e^{-i\sigma_y \gamma/2} = \begin{pmatrix} \cos \gamma/2 & -\sin \gamma/2 \\ \sin \gamma/2 & \cos \gamma/2 \end{pmatrix}, \quad (8.1.8b)$$

$$D_z(\delta) := e^{-i\sigma_z \delta/2} = \begin{pmatrix} e^{-i\delta/2} & 0 \\ 0 & e^{i\delta/2} \end{pmatrix}. \quad (8.1.8c)$$

Proof. Now, we are going to prove the theorem step by step.

**Step 1.** Any $U(2)$ group element $U$ should satisfy the unitary condition

$$UU^\dagger = U^\dagger U = I_2. \quad (8.1.9)$$

Let’s assume that

$$U := \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}, \quad \text{with} \quad a', b', c', d' \in \mathbb{C}. \quad (8.1.10)$$

From E.Q. (8.1.9) we can get

$$|\det U|^2 = \det \left( UU^\dagger \right) = 1,$$

namely

$$\det U = e^{2i\alpha}, \quad \text{with} \quad 0 \leq \alpha < 2\pi \quad \text{and} \quad \alpha \in \mathbb{R}. \quad (8.1.11)$$

**Step 2.** With the E.Q. (8.1.11) in hand, we can denote $U$ as

$$U = e^{i\alpha} U', \quad (8.1.12)$$

with

$$\det U' = 1, \quad \text{and} \quad U'U'^\dagger = U'^\dagger U' = I_2, \quad (8.1.13)$$

and

$$U' := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2), \quad \text{with} \quad a, b, c, d \in \mathbb{C}. \quad (8.1.14)$$
Thus

\[
\begin{align*}
(a & b) \left( a^* \right) = aa^* + bb^* = 1, \quad (8.1.15a) \\
(c & d) \left( c^* \right) = cc^* + dd^* = 1, \quad (8.1.15b) \\
(a & b) \left( c^* \right) = ac^* + bd^* = 0, \quad (8.1.15c) \\
\text{det} \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) = ad - bc = 1. \quad (8.1.15d)
\end{align*}
\]

Thus, we can conclude that

- E.Q. (8.1.15a) and E.Q. (8.1.15b) both can set a constraint on \( U' \), respectively. As we shall see that the imaginary parts of the both two equations are absolutely 0.
- E.Q. (8.1.15c) can spell two constraint on \( U' \).
- It seems that E.Q. (8.1.15d) can also put two constraint on \( U' \). But, we can find out that the constraint (8.1.15d) can be derived from the former three equations (8.1.15a) \( \sim \) (8.1.15c).

Therefore, the E.Q. (8.1.15a) \( \sim \) (8.1.15d) would set totally five constraints on the parameters of \( U' \). Consequently, there are only \( 8 - 5 = 3 \) degrees of freedom left for the \( SU(2) \) group element \( U' \).

**Step 3.** From the constraint E.Q. (8.1.15c) we can derive

\[
\begin{align*}
ac^* + bd^* &= 0 \\
ac^* + bd^* &= 0
\end{align*}
\]

namely

\[
|a|^2 |c|^2 = |b|^2 |d|^2. \quad (8.1.16)
\]

Combine E.Q. (8.1.15a) and E.Q. (8.1.15b) with E.Q. (8.1.16), and we obtain

\[
|a|^2 (1 - |d|^2) = (1 - |a|^2) |d|^2,
\]

i.e.,

\[
|a| = |d|.
\]

Therefore

\[
\begin{align*}
|a| &= |d|, \\
|b| &= |c|,
\end{align*}
\]

\[
1 = |a|^2 + |b|^2. \quad (8.1.17)
\]

Thus,

\[
\begin{align*}
|a| &= |d|, \quad (8.1.18a) \\
|b| &= |c|, \quad (8.1.18b) \\
1 &= |a|^2 + |b|^2, \quad (8.1.18c) \\
0 &= ac^* + bd^*, \quad (8.1.18d) \\
1 &= ad - bc. \quad (8.1.18e)
\end{align*}
\]
Hence, we can denote \( a, b, c, d \) in the form of

\[
\begin{align*}
\begin{aligned}
a &= f_a \cos \frac{\gamma}{2}, \\
b &= -f_b \sin \frac{\gamma}{2}, \\
c &= f_c \sin \frac{\gamma}{2}, \\
d &= f_d \cos \frac{\gamma}{2}.
\end{aligned}
\end{align*}
\] (8.1.19a-d)

Step 4. Then, we have

\[
\begin{align*}
\begin{aligned}
|f_a|^2 &= |f_b|^2 = |f_d|^2 = 1, \\
f_a^* f_c &= f_b^* f_d, \\
f_a f_d &= 1 = f_b f_c,
\end{aligned}
\end{align*}
\] (8.1.20a-c)

namely

\[
\begin{align*}
\begin{aligned}
|f_a| &= |f_b| = |f_c| = |f_d| = 1, \\
f_c &= f_b^*, \\
f_d &= f_a^*.
\end{aligned}
\end{align*}
\] (8.1.21a-c)

Therefore, we can assume that

\[
\begin{align*}
\begin{aligned}
f_a &= e^{-i(\beta+\delta)/2}, \\
f_b &= e^{-i(\beta-\delta)/2}, \\
f_c &= e^{i(\beta-\delta)/2}, \\
f_d &= e^{i(\beta+\delta)/2},
\end{aligned}
\end{align*}
\] (8.1.22a-d)

i.e.,

\[
U = e^{i\alpha} \begin{pmatrix}
  e^{-i(\beta+\delta)/2} \cos \frac{\gamma}{2} & -e^{-i(\beta-\delta)/2} \sin \frac{\gamma}{2} \\
  e^{i(\beta-\delta)/2} \sin \frac{\gamma}{2} & e^{i(\beta+\delta)/2} \cos \frac{\gamma}{2}
\end{pmatrix}
\]

\[
= e^{i\alpha} \begin{pmatrix}
  e^{-i\beta} & 0 \\
  0 & e^{i\delta/2}
\end{pmatrix} \begin{pmatrix}
  e^{-i\delta/2} \cos \frac{\gamma}{2} & -e^{i\delta/2} \sin \frac{\gamma}{2} \\
  e^{-i\delta/2} \sin \frac{\gamma}{2} & e^{i\delta/2} \cos \frac{\gamma}{2}
\end{pmatrix}
\]

\[
= e^{i\alpha} \begin{pmatrix}
  e^{-i\beta} & 0 \\
  0 & e^{i\delta/2}
\end{pmatrix} \begin{pmatrix}
  \cos \frac{\gamma}{2} & -\sin \frac{\gamma}{2} \\
  \sin \frac{\gamma}{2} & \cos \frac{\gamma}{2}
\end{pmatrix} \begin{pmatrix}
  e^{-i\delta/2} & 0 \\
  0 & e^{i\delta/2}
\end{pmatrix}
\]

\[
= e^{i\alpha} \mathcal{D}_z(\beta) \mathcal{D}_y(\gamma) \mathcal{D}_z(\delta). 
\] (8.1.23)

Notes:

- Relations between rotations:

\[
\begin{align*}
\begin{aligned}
\mathcal{D}_y \left( -\frac{\pi}{2} \right) \mathcal{D}_x(\alpha) \mathcal{D}_y \left( \frac{\pi}{2} \right) &= \mathcal{D}_z(\alpha) \\
\mathcal{D}_z \left( \frac{\pi}{2} \right) \mathcal{D}_x(\alpha) \mathcal{D}_z \left( -\frac{\pi}{2} \right) &= \mathcal{D}_y(\alpha)
\end{aligned}
\end{align*}
\] (8.1.24-25)
• Examples:

\[
\begin{align*}
H &= \mathcal{D}_x(\pi) \mathcal{D}_y\left(\frac{\pi}{2}\right) \text{Ph}\left(\frac{\pi}{2}\right), & \text{Hadamard gate,} && (8.1.26) \\
\text{NOT} &= \mathcal{D}_x(\pi) \text{Ph}\left(\frac{\pi}{2}\right), & \text{NOT gate,} && (8.1.27)
\end{align*}
\]

with

\[
\text{Ph}(\theta) := e^{i\theta} I_2. \quad (8.1.28)
\]

These can be verified easily.

– Hadamard gate:

\[
\begin{align*}
\mathcal{D}_x(\pi) &= \cos\left(\frac{\pi}{2}\right) - i \sin\left(\frac{\pi}{2}\right) \sigma_x = -i \sigma_x, \\
\mathcal{D}_y\left(\frac{\pi}{2}\right) &= \cos\left(\frac{\pi}{4}\right) - i \sin\left(\frac{\pi}{4}\right) \sigma_y = \frac{1}{\sqrt{2}} (I_2 - i \sigma_y),
\end{align*}
\]

from which we can get the left-hand-side of E.Q. (8.1.26)

\[
\begin{align*}
\mathcal{D}_x(\pi) \mathcal{D}_y\left(\frac{\pi}{2}\right) \text{Ph}\left(\frac{\pi}{2}\right) &= \frac{1}{\sqrt{2}} \sigma_x (I_2 - i \sigma_y) \\
&= \frac{1}{\sqrt{2}} (\sigma_x - i \sigma_x \sigma_y) \\
&= \frac{1}{\sqrt{2}} \sigma_z \\
&= H.
\end{align*}
\]

– NOT gate \(X\):

\[
\begin{align*}
\mathcal{D}_x(\pi) \text{Ph}\left(\frac{\pi}{2}\right) &= \left(\cos\frac{\pi}{2} - i \sin\frac{\pi}{2} \sigma_x\right) i \\
&= \sigma_x \\
&= X.
\end{align*}
\]

Notes: For an arbitrary single-qubit gate, we exploit the symbol \(\mathcal{D}\) to denote it, since such the notation is often used to denote the spinor representation of the \(SU(2)\) group. In the literature, we may use another the symbol \(\mathcal{R}\) instead of \(\mathcal{D}\), since this notation means the an arbitrary single-qubit gate may represent a rotation in three dimensional space.

8.1.2 Controlled two-qubit gates and controlled three-qubit gates

Def 8.1.2 (Controlled two-qubit gates). Controlled operation is defined as

\[
\begin{align*}
\begin{cases}
\text{if } A \text{ is true, then do } B; \\
\text{if } A \text{ is false, then do } C.
\end{cases}
\end{align*}
\]

Controlled operation is essential in computer science.

Def 8.1.3 (Controlled \(U\) gate). Let’s denote the Controlled \(U\) gate as \(CU\), then

\[
CU : \ket{c} \ket{t} \mapsto \ket{c} U^c \ket{t}, \quad (8.1.29)
\]
which means

\[
\begin{align*}
\text{if } c = 0, \quad |0\rangle |t\rangle & \rightarrow |0\rangle |t\rangle, \\
\text{if } c = 1, \quad |1\rangle |t\rangle & \rightarrow |1\rangle U |t\rangle,
\end{align*}
\]

where \( |c\rangle \) is the control qubit, and \( |t\rangle \) is the target qubit. This can also be represented in the diagrammatical formalism:

\[
CU = \begin{array}{c}
\bullet \\
U
\end{array}
\]

One example of the Controlled \( U \) gate is the CNOT gate, i.e.,

\[
\begin{array}{c}
\begin{array}{c}
\bullet \\
\times
\end{array}
\end{array} = \begin{array}{c}
\begin{array}{c}
\bullet \\
\times
\end{array}
\end{array}
\]

### 8.1.2.1 Quantum Toffoli gate and Fredkin gate

**Remarks:** Quantum Toffoli gate and Fredkin gate can be regarded as controlled-controlled-NOT gate and controlled-swap gate respectively.

**Def 8.1.4.** *Analogy to the classical reversible gate, the Quantum Toffoli gate is defined as*

\[
\theta^{(3)} (x,y,z) = |x,y,z \oplus xy\rangle,
\]

(8.1.30)

where \( x, y, z = 0, 1 \) and it has the diagram formalism,

\[
\begin{array}{c}
|x\rangle \\
|y\rangle \\
|z\rangle
\end{array} \oplus
\begin{array}{c}
|x\rangle \\
|y\rangle \\
|z\rangle
\end{array} =
\begin{array}{c}
|x\rangle \\
|y\rangle \\
|z\rangle \oplus xy
\end{array}.
\]

**Remark:** Quantum Toffoli gate with the Hadamard gate and phase gate is the universal quantum gate set, namely these three gates can perform universal quantum computation.

**Def 8.1.5.** *The Quantum Fredkin gate is defined as*

\[
\theta^{(3)} (x,y,z) = |x, xz \oplus xy, xy \oplus \bar{xz}\rangle,
\]

(8.1.31)

where \( x, y, z = 0, 1 \) and it has the diagram formalism,

\[
\begin{array}{c}
|x\rangle \\
|y\rangle \\
|z\rangle
\end{array} \oplus
\begin{array}{c}
|xz \oplus xy\rangle \\
|xy \oplus \bar{xz}\rangle
\end{array} =
\begin{array}{c}
|x\rangle \\
|y\rangle \\
|z\rangle \oplus xy
\end{array}.
\]

### 8.1.3 Quantum circuit model of GHZ states

Bell states \[\text{Most important and popular state} \]
\[\text{Two-qubit maximally entangled} \]
\[\text{Bell inequality} \]

GHZ states \[\text{Most important and popular state} \]
\[\text{Multi-qubit maximally entangled} \]
\[\text{GHZ theorem} \]

82
Thm 8.1.3.1.

\[ |\text{GHZ}\rangle_n := \frac{1}{\sqrt{2}} \left( |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right) \]

\[ = \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1x_2\cdots x_n\rangle. \quad (8.1.32) \]

where \( \text{CNOT}_{ij} \) is the \( \text{CNOT} \) gate with the \( i \)-th qubit as the control qubit and \( j \)-th qubit as the target qubit. And this can be represented in the diagram formulism, shown in Figure 8.1.

\[ \begin{array}{c}
|x_1\rangle \quad \text{H} \\
|x_2\rangle \\
|x_3\rangle \\
\vdots \\
|x_n\rangle \\
\end{array} \]

\[ |\text{GHZ}\rangle \\
\]

Figure 8.1: \( |\text{GHZ}\rangle \) state generated by the \( \text{CNOT} \) and \( H \) gates.

e.g.1. Three qubit GHZ state.

\[ |\text{GHZ}\rangle_3 = \frac{1}{\sqrt{2}} \left( |000\rangle + |111\rangle \right). \quad (8.1.33) \]

\[ \begin{array}{c}
|0\rangle \quad |0\rangle \\
|0\rangle \\
\vdots \\
\end{array} \]

\[ \begin{array}{c}
|\text{H}\rangle \\
|\text{GHZ}\rangle_3 \\
\vdots \\
\end{array} \]

\[ \begin{array}{c}
t_0 \\
t_1 \\
t_2 \\
t_3 \\
\end{array} \]

Figure 8.2: \( |\text{GHZ}\rangle_3 \) state generated by the \( \text{CNOT} \) and \( H \) gates.

- \( t_0 : |000\rangle; \)
- \( t_1 : \frac{1}{\sqrt{2}} \left( |000\rangle + |100\rangle \right); \)
- \( t_2 : \frac{1}{\sqrt{2}} \left( |000\rangle + |110\rangle \right); \)
- \( t_3 : \frac{1}{\sqrt{2}} \left( |000\rangle + |111\rangle \right). \)

e.g.2. Four qubit GHZ state.

\[ |\text{GHZ}\rangle_4 = \frac{1}{\sqrt{2}} \left( |0011\rangle - |1100\rangle \right). \quad (8.1.34) \]
Let’s now consider the complete set of observables defining the GHZ state.

- The phase-bit operator $X_1 \otimes X_2 \otimes \cdots \otimes X_n$
  
  \[ (X_1 \otimes X_2 \otimes \cdots \otimes X_n) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \]
  
  \[ = \text{CNOT}_{1n} (X_1 \otimes X_2 \otimes \cdots \otimes X_{n-1} \otimes I_2) \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \]
  
  \[ = \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \]

- $t_0 : |1011 \rangle$;
- $t_1 : \frac{1}{\sqrt{2}} (|0011 \rangle - |1011 \rangle)$;
- $t_2 : \frac{1}{\sqrt{2}} (|0011 \rangle + |1111 \rangle)$;
- $t_3 : \frac{1}{\sqrt{2}} (|0011 \rangle + |1101 \rangle)$;
- $t_4 : \frac{1}{\sqrt{2}} (|0011 \rangle - |1100 \rangle)$.

Let’s now consider the complete set of observables defining the GHZ state.

- The phase-bit operator $X_1 \otimes X_2 \otimes \cdots \otimes X_n$:

\[
(X_1 \otimes X_2 \otimes \cdots \otimes X_n) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \\
= \text{CNOT}_{1n} (X_1 \otimes X_2 \otimes \cdots \otimes X_{n-1} \otimes I_2) \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \\
= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \\
= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 (Z \otimes I_2 \otimes \cdots \otimes I_2) |x_1 x_2 \cdots x_n \rangle, \\
\]

i.e.,

\[
(X_1 \otimes X_2 \otimes \cdots \otimes X_n) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle \\
= (-1)^{x_1} \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} \text{H}_1 |x_1 x_2 \cdots x_n \rangle, \\
\]

where we have utilized the relation (4.4.22), namely

\[
(X_1 \otimes X_2) \text{CNOT}_{12} = \text{CNOT}_{12} (X_1 \otimes I_2), \\
\]

and the property

\[
\begin{align*}
H X H &= Z \\
H^2 &= I_2
\end{align*} \\
\Rightarrow X H = H Z.
\]
• The first parity-bit operator $Z_1 \otimes Z_2 \otimes I_3 \otimes \cdots \otimes I_2$:

$$\left(Z_1 \otimes Z_2 \otimes I_3 \otimes \cdots \otimes I_2 \right) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{13}$$

$$\left(Z_1 \otimes Z_2 \otimes I_3 \otimes \cdots \otimes I_2 \right) \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} (I_2 \otimes Z_2 \otimes I_3 \otimes \cdots \otimes I_2) H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= -1^{x_2} \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle,$$

(8.1.37)

where we have used the relations (4.4.22d) and (4.4.22f), namely

$$\left\{ \begin{array}{l}
(Z \otimes Z) \text{CNOT}_{12} = \text{CNOT}_{12} (I_2 \otimes Z),

(Z \otimes I_2) \text{CNOT}_{12} = \text{CNOT}_{12} (Z \otimes I_2).
\end{array} \right.$$  

• The second parity-bit $I_2 \otimes Z_2 \otimes Z_3 \otimes I_3 \otimes \cdots \otimes I_2$:

$$\left(I_2 \otimes Z_2 \otimes Z_3 \otimes I_3 \otimes \cdots \otimes I_2 \right) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= \text{CNOT}_{1n} \cdots \text{CNOT}_{14} (I_2 \otimes Z_2 \otimes Z_3 \otimes I_3 \otimes I_2) \text{CNOT}_{13} \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{13}$$

$$\left(Z_1 \otimes Z_2 \otimes Z_3 \otimes I_3 \otimes I_2 \right) \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} (I_2 \otimes Z_2 \otimes Z_3 \otimes I_3 \otimes I_2) H_1 |x_1 x_2 \cdots x_n\rangle$$

$$= -1^{x_2+x_3} \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n\rangle,$$

(8.1.37)

to derive this result we have to employ the relations (4.4.22d) and (4.4.22f), i.e.,

$$\left\{ \begin{array}{l}
(Z \otimes Z) \text{CNOT}_{12} = \text{CNOT}_{12} (I_2 \otimes Z),

(I_2 \otimes Z) \text{CNOT}_{12} = \text{CNOT}_{12} (Z \otimes Z).
\end{array} \right.$$
• The third parity-bit $I_2 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes I_2$:

$$
(I_2 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes \cdots \otimes I_2) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \cdots \text{CNOT}_{15}
$$

$$
(I_2 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes \cdots \otimes I_2) \text{CNOT}_{14} \text{CNOT}_{13} \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \cdots \text{CNOT}_{14}
$$

$$
(Z_1 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes \cdots \otimes I_2) \text{CNOT}_{13} \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \cdots \text{CNOT}_{13}(I_2 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes \cdots \otimes I_2) \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \cdots \text{CNOT}_{12} H_1 (I_2 \otimes I_2 \otimes Z_3 \otimes Z_4 \otimes I_2 \otimes \cdots \otimes I_2) \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= (-1)^{x_3 + x_4} \text{CNOT}_{1n} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n).
$$

(8.1.38)

• The $(n-1)$-th parity-bit $I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n$:

$$
(I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n) \text{CNOT}_{1n} \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n}(Z_1 \otimes I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n) \text{CNOT}_{1(n-1)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \text{CNOT}_{1(n-1)}
$$

$$
(I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n) \text{CNOT}_{1(n-2)} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n)
$$

$$
= \text{CNOT}_{1n} \cdots \text{CNOT}_{12} H_1 (I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n) |x_1 x_2 \cdots x_n)
$$

$$
= (-1)^{x_{n-1} + x_n} \text{CNOT}_{1n} \cdots \text{CNOT}_{12} H_1 |x_1 x_2 \cdots x_n).
$$

(8.1.39)

## 8.2 Universal quantum computation

### 8.2.1 Quantum universal gate set

The definition of the Universal Quantum Gate set is analogous to the Universal Classical gate set.

**Def 8.2.1 (Universal Quantum Gate set).** A set of quantum gates is universal if any unitary operator $U \in SU(2)$ (or $U \in U(2)$, up to a global phase), can be expressed as a product of the elementary gates from this very set.

One example of the Universal Quantum Gate set is

$$
\{\text{CNOT}, \text{one-qubit gate } \epsilon SU(2)\}.
$$

This is a Universal Quantum Gate set, but the number of the elementary gates in the set is infinity. Thus it is not practical to construct a Quantum Computer with this gate set.
Def 8.2.2 (Finite Approximately Universal Quantum Gate set). A finite gate set is approximately universal, if any unitary operator in $SU(2^n)$ can be approximately expressed as a product of elementary gates in this set to arbitrary precisions.

Examples:

1) $\{H, S, \theta^{(3)}\}$, $H$ being the Hadamard gate, $S$ representing the Phase gate, and $\theta^{(3)}$ for the three-bit quantum Toffoli gate, i.e.,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad \theta^{(3)}|x, y, z\rangle = |x, y, z \oplus xy\rangle.$$

For this case, there are only three elementary gates in this set, we can construct a very small Quantum Computer. And we can get arbitrary precisions, though this gate set is not a "real" Universal gate set.

2) $\{H, T, \text{CNOT}\}$ with $T$ as the $\frac{\pi}{8}$ gate and CNOT denoted the CNOT gate, namely,

$$T = e^{i\frac{\pi}{8}} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}, \quad \text{CNOT}|x, y\rangle = |x, x \oplus y\rangle.$$

Remark: $S = T^2$.

NOTE: Designing of a unitary operator as product of elementary gates may require exponential number of gates, therefore, this sort of design is not sufficient and Quantum Computer may run very slowly in certain conditions.

8.2.2 Universal quantum gate set of two-qubit gates

Def 8.2.3 (Generic two-qubit gates). The two-qubit gate with eigenvalues $e^{i\theta_1}$, $e^{i\theta_2}$, $e^{i\theta_3}$, and $e^{i\theta_4}$, where $\frac{\theta_1}{\pi}$ and $\frac{\theta_2}{\pi}$ are irrational numbers, is defined as Generic two-qubit gate.

To completely understand this, we have to learn the Number Theory in mathematics.

Thm 8.2.2.1. {Any Generic two-qubit gate} is an universal quantum gate set.

Thm 8.2.2.2 (Barenco’s gate). Barenco’s gate (1995) is defined as

$$\text{Controlled-Ph}\left( -\frac{\pi}{4} \right) \mathcal{D}_{x} \left( \frac{\theta}{2} \right), \quad \text{(8.2.1)}$$

where $\frac{\theta}{\pi}$ is an irrational number. It can be represented in the diagram formalism as

```
\[ U \]
```

with

$$U := \text{Ph}\left( -\frac{\pi}{4} \right) \mathcal{D}_{x} \left( \frac{\theta}{2} \right) = e^{-i\frac{\pi}{4}} \begin{pmatrix} \cos \frac{\theta}{4} & -i \sin \frac{\theta}{4} \\ -i \sin \frac{\theta}{4} & \cos \frac{\theta}{4} \end{pmatrix}. \quad \text{(8.2.2)}$$

Barenco’s gate is not a generic gate. One thing to keep in mind is that we want to avoid the irrational number. The set $\{\text{Barenco's gate, SWAP}\}$ is a universal quantum gate set.
Thm 8.2.3. \(\{\text{CNOT}, \text{arbitrary single-qubit gates}\}\) is a universal quantum gate set.

**Proof.** Since Barenco’s gate together with the \(\text{SWAP}\) gate can perform universal quantum computation, we describe the \(\text{SWAP}\) gate as

\[
\text{SWAP} = \text{CNOT}_{12}\text{CNOT}_{23}\text{CNOT}_{12}; \tag{8.2.3}
\]

and with the following Lemmas, reformulate Barenco’s gate as an expression in terms of the \(\text{CNOT}\) gate and single-qubit gates.

**Lemma 8.2.2.1.** Any single-qubit gate \(U\), can be described in terms of Euler angles, namely

\[
U = e^{i\alpha} \mathcal{D}_z(\beta) \mathcal{D}_y(\gamma) \mathcal{D}_z(\delta), \quad \text{with } \alpha, \beta, \gamma, \delta \in \mathbb{R}. \tag{8.2.4}
\]

**Lemma 8.2.2.2.** Any single-qubit gate \(U\) can be formulated as

\[
U = \text{Ph}(\alpha) \text{AXBXC}, \quad \text{with } ABC = I_2, \ A, B, C \in SU(2) \text{ and } \alpha \in \mathbb{R}. \tag{8.2.5}
\]

**Lemma 8.2.2.3.** Any controlled two-qubit gate can be decomposed in the following way

\[
U = R_{\alpha} C \oplus B \oplus A \tag{8.2.6}
\]

with

\[
R_{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}.
\]

**Proof.** We are now going to prove the three lemmas 8.2.2.1, 8.2.2.2 and 8.2.2.3.

**Lemma 1:** The lemma 8.2.2.1 is proved in subsection 8.1.1, section 8, chapter 7.

**Lemma 2:** The lemma 8.2.2.2 can be proved through lemma 8.2.2.1. Because \(A, B, C\) are all elements of the \(SU(2)\) group, therefore, from lemma 8.2.2.1, we can construct \(A, B, C\) in the form of:

\[
\begin{align*}
A &= \mathcal{D}_z(\phi) \mathcal{D}_y(\varphi) \mathcal{D}_z(\psi), \tag{8.2.7a} \\
B &= \mathcal{D}_z(\phi') \mathcal{D}_y(\varphi') \mathcal{D}_z(\psi'), \tag{8.2.7b} \\
C &= \mathcal{D}_z(\phi'') \mathcal{D}_y(\varphi'') \mathcal{D}_z(\psi''), \tag{8.2.7c}
\end{align*}
\]

where we have introduced 9 parameters. The expressions (8.2.7a)~(8.2.7c) ensures that \(A, B\) and \(C\) all belong to \(SU(2)\), therefore \(ABC \in SU(2)\). Since we can parameterize an arbitrary \(SU(2)\) operator with only two independent real variables, the expressions (8.2.7a)~(8.2.7c) have two constraints. And the relation

\[
ABC = I_2 \tag{8.2.8}
\]

gives rise to another constraint on the parameters that we have introduced in the expressions (8.2.7a)~(8.2.7c). Therefore, we actually have 6 independent variables in the expressions (8.2.7a)~(8.2.7c). We do not need that much of independent
variables, since \( U \in U(2) \) can be determined with three independent real variables. We may set

\[
\begin{align*}
A &= \mathcal{D}_z(\beta) \mathcal{D}_y \left( \frac{\gamma}{2} \right), \\
B &= \mathcal{D}_y \left( \frac{-\gamma}{2} \right) \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right), \\
C &= \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right).
\end{align*}
\tag{8.2.9a-b-c}
\]

As we shall see that

\[
ABC = \mathcal{D}_z(\beta) \mathcal{D}_y \left( \frac{\gamma}{2} \right) \mathcal{D}_y \left( \frac{-\gamma}{2} \right) \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right) \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right)
\]

\( = I_2, \)

and

\[
AXBXC = \mathcal{D}_z(\beta) \mathcal{D}_y \left( \frac{\gamma}{2} \right) \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right) \mathcal{D}_z \left( -\frac{\beta + \delta}{2} \right)
\]

\( = I_2, \)

from which we can see that theorem 8.2.2.2 can be derived from theorem 8.2.2.1.

**Lemma 3:** The lemma 8.2.2.3 can be proved by utilizing the lemma 8.2.2.2. The controlled-\( U \) gate can be expressed as

\[
CU = |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes U = \begin{pmatrix} I_2 & U \end{pmatrix}
\tag{8.2.11}
\]

By inserting E.Q.\( (8.2.5) \) into E.Q.\( (8.2.11) \), and we can get

\[
CU = \left( \begin{array}{cc} ABC & e^{i\alpha}AXBXC \\ e^{i\alpha}I_2 \end{array} \right)
\]

\( = \begin{pmatrix} I_2 & e^{i\alpha}I_2 \\ e^{i\alpha}I_2 & I_2 \end{pmatrix} \begin{pmatrix} A & B \\ A & X \end{pmatrix} \begin{pmatrix} I_2 & X \\ B & C \end{pmatrix} \begin{pmatrix} C & C \end{pmatrix}
\]

\( = C - \text{Ph}(\alpha) \begin{pmatrix} A & B \\ A & C \end{pmatrix} \text{CNOT} \begin{pmatrix} B & C \\ B & C \end{pmatrix}, \)

\( \tag{8.2.12} \)

where \( C - \text{Ph}(\alpha) \) is the controlled-phase gate,

\[
C - \text{Ph}(\alpha) = |0\rangle \langle 0| \otimes I_2 + |1\rangle \langle 1| \otimes e^{i\alpha}I_2
\]

\( = \begin{pmatrix} I_2 & e^{i\alpha}I_2 \\ e^{i\alpha}I_2 & I_2 \end{pmatrix}
\]

\( = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} \otimes I_2
\]

\( = R_\alpha \otimes I_2, \)

\( \tag{8.2.13} \)
namely
\[ \Phi(\alpha) = R_\alpha. \] (8.2.14)

On the other hand, we can also obtain
\[ \begin{align*}
(A)_{I_2 \otimes A} &= A; \\
(B)_{I_2 \otimes B} &= B; \\
(C)_{I_2 \otimes C} &= C.
\end{align*} \] (8.2.15a)

On the other hand we know that
\[ C U = U. \] (8.2.16)

By substituting E.Q. (8.2.13), (8.2.15a)~(8.2.15c) and (8.2.16) in E.Q. (8.2.12), we can acquire E.Q. (8.2.6).

8.2.3 Deutsch’s gate is a universal quantum gate

Thm 8.2.3.1 (Deutsch’s gate). Deutsch gate (1989) is a universal quantum gate,

Deutsch’s gate: \( |x\rangle |y\rangle |z\rangle \rightarrow |x\rangle |y\rangle R^x |z\rangle, \) (8.2.17)

where
\[ R := \Phi\left( -\frac{\pi}{2} \right) D_x(\theta) = -ie^{-i\sigma_2^x}, \] (8.2.18)

with \( \frac{\theta}{\pi} \) being an irrational real number. And it has the diagrammatical formalism,

Deutsch’s gate
\[ \begin{align*}
\text{Deutsch's gate} &= R.
\end{align*} \]

This is the first universal quantum gate, and is a three-qubit gate. Note that
\[ R_{\text{Deutsch}} = \Phi\left( -\frac{\pi}{2} \right) D_x(\theta) = \left[ \Phi\left( -\frac{\pi}{4} \right) D_x\left( \frac{\theta}{2} \right) \right]^2 = \left( U_{\text{Barenco}} \right)^2, \]

and we show that Deutsch’s gate can be expressed in terms of the two-qubit gates, i.e., CNOT and Barenco’s gate,

\[ \begin{align*}
\text{CNOT} &= U_0 \\
\text{Barenco} &= U_1
\end{align*} \] (8.2.19)
If we let the quantum circuit act on $|x\rangle \otimes |y\rangle \otimes |z\rangle$ state, then we can get the tripartite states in different steps as labeled in the right-hand-side of E.Q. (8.2.19):

1. $\psi(t_0) = |x\rangle |y\rangle |z\rangle$;
2. $\psi(t_1) = |x\rangle |y\rangle U^y |z\rangle$;
3. $\psi(t_2) = |x\rangle |x \oplus y\rangle U^y |z\rangle$;
4. $\psi(t_3) = |x\rangle |x \oplus y\rangle (U^x)^{x \otimes y} U^y |z\rangle$
   $= |x\rangle |x \oplus y\rangle U^{-x \otimes y} U^y |z\rangle$;
5. $\psi(t_4) = |x\rangle |x \oplus y\oplus x\rangle U^{-x \otimes y} U^y |z\rangle$
   $= |x\rangle |y\rangle U^{-x \otimes y} U^y |z\rangle$;
6. $\psi(t_5) = |x\rangle |y\rangle U^x U^{-x \otimes y} U^y |z\rangle$
   $= |x\rangle |y\rangle U^{x+y-x \otimes y} |z\rangle$
   $= |x\rangle |y\rangle U^{x+y-(x+y-2xy)} |z\rangle$
   $= |x\rangle |y\rangle U^{2xy} |z\rangle$
   $= |x\rangle |y\rangle R^{xy} |z\rangle$.

Note: the calculation here is binary, namely

$$x \oplus y \equiv (x + y) \mod 2$$
$$= x + y - 2xy, \quad (8.2.20)$$

where $x, y \in \{0, 1\}$.  


Chapter 9

Physical Realization of Quantum Computers
Chapter 10

Quantum Algorithms

Quantum mechanics: Real Black Magic Calculus.

—Albert Einstein

Computer programming is an art of form, like the creation of poetry or music.

—Donald Knuth

Deutsch’s algorithm combines quantum parallelism with a property of quantum mechanics known as interference.

—Nielsen & Chuang

The quantum search algorithm is essentially optimal, is both exciting and disappointing. It is exciting because it tells us that for this problem, at least we have fully plumbed the depths of quantum mechanics, no further improvement is possible. The disappointment arises because we might have hoped to do much better than the square root speedup offered by the quantum search algorithm.

—Nielsen & Chuang

The essence of the design of many quantum algorithms is that a clever choice of function and final transformation allows efficient determination of useful global information about the function—information which can not be attained quickly on a classical computer.

—Nielsen & Chuang

Reference:

• [Preskill] Chapter 6: Quantum computation;
• [Nielsen & Chuang] Chapter 6: Quantum search algorithms

10.1 Classical and quantum algorithm

Def 10.1.1 (Algorithm). An algorithm is

1. a well-defined procedure;

2. with finite descriptions;
3. designed for realising an information processing task;
4. designed to solve a computational problem.

This definition is proper for both classical algorithm and quantum algorithm. The only thing we have to notice is that, the classical algorithm is designed to process classical information, while the quantum algorithm is designed to perform information processing task using the fundamental principle’s of Quantum Mechanics.

But, what advantage can we get from the Quantum Algorithm? One of the benefit is that Quantum Algorithm is more efficient than the Classical Algorithm in some specific problems. And there are some examples, as shown in Table 10.1:

Table 10.1: Examples of Quantum Algorithm

<table>
<thead>
<tr>
<th>Speed up</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-exponential speed up</td>
<td>Grover’s search algorithm (Oracle model)</td>
</tr>
<tr>
<td>Relativized exponential speed up</td>
<td>Simon’s algorithm (Oracle model)</td>
</tr>
<tr>
<td>Exponential speed up</td>
<td>Shor’s factoring algorithm</td>
</tr>
</tbody>
</table>

10.2 Oracle model

Oracle model is the black-box model.

**Def 10.2.1 (Oracle (black box)).** *A subroutine evaluating a function, but we have no idea about how this subroutine is performed. We are only concerned about how this subroutine is performed, or we only make a query, on the Oracle (Black-box).*

**Remark:** For the oracle (black box), usually we don’t care about how the black box works, and we are only concerned about how to use it.

- We can make a query on $U_f$.
- We are only concerned about the input and output.
- No idea about how $U_f$ is performed.

**Examples:**

(1) Classical black box:

$$x \xrightarrow{f} f(x),$$

which may be not a reversible operation. And we can define the classical black box which is a reversible gate to compute the function $f(x)$ shown as

$$x \xrightarrow{U_f} x$$

$$y \xrightarrow{y \oplus f(x)}$$
(2) Quantum black box is a quantum gate $U_f$ computing $f(x)$:

$$
|x\rangle \xrightarrow{U_f} |x\rangle \\
|y\rangle \xrightarrow{U_f} |y\oplus f(x)\rangle,
$$

where $|x\rangle$ is called the register qubit, and $|y\rangle$ is called the ancilla qubit. Note that quantum gate $U_f$ satisfies the unitary (reversible) condition given by

$$
U_f^\dagger U_f = U_f U_f^\dagger = I_d.
$$

### The Query Complexity:

- In the oracle model, we only count the minimal number of the black box to characterize the circuit complexity. Because the time of performing $U_f$ gate takes far greater than the time of performing other gates.
- Complexity is very abstract, but very important in computer science.

**Note:** Time of performing $U_f$ $\gg$ Time of perform other gates. For example

\[
\text{Time (} U_f \text{)} = 1 \text{ year;} \\
\text{Time (other gates)} = 1 \text{ sec.}
\]

### 10.3 Deutsch’s algorithm

The Deutsch’s algorithm is the first Quantum Algorithm, presented in 1989.

#### 10.3.1 Definitions

**Def 10.3.1** (Constant function and Balanced function). *Let’s define a function $f$:

$$
f: \forall x \in \{0,1\} \rightarrow y = f(x) \in \{0,1\}.
$$

Then, the function $f$ is

\[
\begin{cases}
\text{Constant function, if } f(0)=f(1); \\
\text{Balanced function, if } f(0)\neq f(1).
\end{cases}
\]

- In the case that $f$ is constant, i.e., $f(0) = f(1)$, we can acquire

$$
f(0) = f(1) = 0, \quad \text{or} \quad f(0) = f(1) = 1.
$$

Therefore,

$$
f(0) \oplus f(1) = 0
$$

with the binary addition.

- In the case that $f$ is balanced, i.e., $f(0) \neq f(1)$, we can know

$$
f(0) = 0, f(1) = 1, \quad \text{or} \quad f(0) = 1, f(1) = 0.
$$

Similarly,

$$
f(0) \oplus f(1) = 1
$$

which suggests that the number of $x$ which satisfies $f(x) = 0$ and $x$ that $f(x) = 1$ is same, namely

\[
1 = \#\{x|f(x) = 0, \forall x \in \{0,1\}\} = \#\{x|f(x) = 1, \forall x \in \{0,1\}\}.
\]
**Question:** Assume that the $U_f$ is a black box of computing $f(x)$, what is the minimum number of performing $U_f$ to judge whether $f(x)$ is a constant or balanced function?

### 10.3.2 Classical algorithm

The minimum number of performing $U_f$ to know about whether $f(x)$ is constant or balanced is two. For

$$
\begin{array}{c}
x - U_f - x \\
y - y \oplus f(x),
\end{array}
$$

we can construct the following classical circuit to verify whether $f(x)$ is constant or balanced,

$$
\begin{array}{c}
0 - U_f - 0 \\
0 - f(0) \\
1 - U_f - 1 \\
0 - f(0) \oplus f(1),
\end{array}
$$

where obviously the $f(0)$ and $f(1)$ have to be calculated independently.

### 10.3.3 Deutsch’s problem

Deutsch’s problem is firstly presented in 1989, and it is of conceptual importance. This problem can be stated in the following way.

- **Input:** A black-box for computing an unknown function $f(x)$, i.e.,

$$
\begin{array}{c}
|x\rangle - U_f - |x\rangle \\
|y\rangle - |y \oplus f(x)\rangle.
\end{array}
$$

- **Promise:** $f(x)$ is constant or balanced.
- **Question:** Determine $f(x)$ constant or balanced.
- **Answer:** 1 time of using $U_f$.

**Note:** In quantum mechanics, $f(0)$ and $f(1)$ can be computed simultaneously, because of quantum superposition principle.

### 10.3.4 Phase kick-back

**Lemma 10.3.4.1.** With the quantum black box $U_f$ defined as

$$
U_f: |x\rangle |y\rangle \mapsto |x\rangle |y \oplus f(x)\rangle,
$$

with the diagrammatical representation

$$
\begin{array}{c}
|x\rangle - U_f - |x\rangle \\
|y\rangle - |y \oplus f(x)\rangle,
\end{array}
$$

we can prove the following algebraic formula,

$$
U_f |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}},
$$

96
i.e., the quantum circuit model given by

\[
|x\rangle \quad \text{U}_f \quad |x\rangle \quad \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}} \quad (-1)^{f(x)} \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}}.
\]

or equivalently

\[
|x\rangle \quad \text{U}_f \quad (-1)^{f(x)} |x\rangle \quad \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}}.
\]

**Note:** Why is the above quantum circuit model called the phase kick-back? The global phase \((-1)^{f(x)}\) can be kicked back from the target qubit to the control qubit, so that the target qubit \(\frac{0\langle 0| - 1\langle 1|}{\sqrt{2}}\) seems unchanged from the input to the output.

This lemma can be proved in the following manner.

**Proof.**

\[
\text{U}_f |x\rangle \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}} = \frac{\text{U}_f |x\rangle |0\rangle - \text{U}_f |x\rangle |1\rangle}{\sqrt{2}} = \frac{|x\rangle |f(x)\rangle - |x\rangle |\bar{f}(x)\rangle}{\sqrt{2}} = \frac{|x\rangle |f(x)\rangle - |\bar{f}(x)\rangle}{\sqrt{2}} = \begin{cases} 
\text{if } f(x) = 0, & |x\rangle \frac{1}{\sqrt{2}}(0\langle 0| - 1\langle 1|) \\
\text{if } f(x) = 1, & |x\rangle \frac{1}{\sqrt{2}}(1\langle 1| - 0\langle 0|) = (-1)|x\rangle \frac{1}{\sqrt{2}}(0\langle 0| - 1\langle 1|) 
\end{cases} = (-1)^{f(x)} |x\rangle \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}}.
\]

There we get E.Q. (10.3.13) verified. \(\square\)

**Example:** If \(f(x) = x\), then \(\text{U}_x = \text{CNOT}\), since

\[
\text{U}_x |x\rangle |y\rangle = |x\rangle |y\oplus f(x)\rangle = |x\rangle |y\oplus x\rangle = \text{CNOT} |x\rangle |y\rangle.
\]

And we can see that

\[
\text{CNOT} |x\rangle \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}} = \frac{|x\rangle |x\rangle - |x\rangle |\bar{x}\rangle}{\sqrt{2}} = (-1)^x |x\rangle \frac{0\langle 0| - 1\langle 1|}{\sqrt{2}},
\]

so that in such the circumstance we have \(\text{CNOT} |x\rangle = (-1)^x |x\rangle\).

**Note:** The phase kick-back is the key technique in the performance of such quantum algorithms as Deutsch’s algorithm, Deutsch-Jozsa’a algorithm, Grover’s search algorithm, and so on.
10.3.5 Deutsch’s algorithm

This is the first quantum algorithm in quantum information science. Although this algorithm may be nonsense in practice, but is of conceptual importance. The associated quantum circuit is of the form,

\[
\begin{array}{c}
|0\rangle - |1\rangle \\
\sqrt{2}
\end{array}
\quad \begin{array}{c}
H \\
\downarrow
\end{array}
\quad \begin{array}{c}
U_f \\
\downarrow
\end{array}
\quad \begin{array}{c}
H \\
\downarrow
\end{array}
\quad \begin{array}{c}
|0\rangle - |1\rangle \\
\sqrt{2}
\end{array}
\quad \begin{array}{c}
\boxed{\text{input}}
\end{array}
\]

(10.3.19)

where we study it in the following five steps:

1. Input: state preparation.
   
   \[|\psi(t_1)\rangle = |0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}},\]

2. |\psi(t_2)⟩ = (H⊗I₂)|ψ(t₁)⟩
   \[= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}},\]
   
   We have utilized the relation that
   \[\begin{cases}
   H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
   H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}
   \end{cases}\iff H|i\rangle = \frac{1}{\sqrt{2}} \sum_{j=0}^{1} (-1)^{i\cdot j} |j\rangle \quad (10.3.20)
   
   where \(i \cdot j\) is AND operation.

3. |ψ(t₃)⟩ = U_f|ψ(t₂)⟩
   \[= \frac{1}{\sqrt{2}} \sum_{i=0}^{1} U_f|i\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\]
   \[= \frac{1}{\sqrt{2}} \sum_{i=0}^{1} (-1)^{f(i)}|i\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}},\]

4. |ψ(t₄)⟩ = (H⊗I₂)|ψ(t₃)⟩
   \[= \frac{1}{\sqrt{2}} \sum_{i=0}^{1} (-1)^{f(i)}(H⊗I₂)|i\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\]
   \[= \frac{1}{2} \sum_{i,j=0}^{1} (-1)^{f(i)+i\cdot j}|j\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}.
   
5. Output: quantum measurement.

And the icon \(\text{input} \rightarrow \text{output}\) stands for the measurement with respect to the standard basis \{⟨0|, |1⟩\}. If we perform the measurement \[|0\rangle \langle 0| \otimes I,\]
   there we can get the state after measurement for the bipartite system as,

\[j = 0: \quad |\tilde{\psi}(t₄)\rangle = \frac{1}{2} \sum_{i=0}^{1} (-1)^{f(i)}|0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (10.3.21)\]

And we can see that,
Then function $f$ is a constant function, i.e., $f(0) = f(1)$, the probability that the first qubit is in the state $|0\rangle$ is 1, since

$$(0) \langle 0| \otimes I_2 ) |\psi(t_4)\rangle = (-1)^{f(0)} |0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \quad (10.3.22)$$

$$(0) \langle 1| \otimes I_2 ) |\psi(t_4)\rangle = \frac{1}{2} \sum_{i=0}^{1} (-1)^{f(i+1)} |1\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$= (-1)^{f(0)} \frac{1}{2} \sum_{i=0}^{1} (-1)^i |1\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$= 0; \quad (10.3.23)$$

(2) if $f$ is a balanced function, i.e., $f(0) \neq f(1)$, the probability that the first qubit is in the state $|0\rangle$ is 0, due to

$$(0) \langle 0| \otimes I_2 ) |\psi(t_4)\rangle = \frac{1}{2} \sum_{i=0}^{1} (-1)^{f(i)} |0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$= 0. \quad (10.3.24)$$

**Note:** What is underlying Deutsch’s algorithm? The $f(0)$ and $f(1)$ are calculated simultaneously so that $f(0) \otimes f(1)$ is calculated using the black box $U_f$ once. This is called quantum parallelism due to quantum interference or quantum superposition principle.

**10.4 Deutsch-Jozsa’s algorithm**

As we have seen that the Deutsch’s algorithm uses the one-qubit register, we are going to consider the Deutsch-Jozsa’s Algorithm which uses an $n$-qubit register.

**10.4.1 Constant and balanced function in $n$-qubit**

**Def 10.4.1** (Constant and Balanced function $(n$-qubit)). *Let’s define the function $f$ to be*

$$f : \quad \forall x \in \{0,1\}^n \Rightarrow f(x) \in \{0,1\}. \quad (10.4.1)$$

Then function $f$ is

- **Constant function**, if $f(x) = c, \; \forall x \in \{0,1\}^n$, where $c$ is a constant in $\{0,1\}$;

- **Balanced function**, if the number of elements in the set $\{x | f(x) = 1, \forall x \in \{0,1\}^n\}$ is the same as that in $\{x | f(x) = 0, \forall x \in \{0,1\}^n\}$, i.e., the numbers both equal to $2^{n-1}$.

$$\#\{x | f(x) = 0, \forall x \in \{0,1\}^n\} = \#\{x | f(x) = 1, \forall x \in \{0,1\}^n\} = 2^{n-1}. \quad (10.4.2)$$

**10.4.2 Constant or balanced function?**

**Question:** If we know that $f$ is either a Constant function or a Balanced function, how can we determine it? And, how many queries would have to be made?

There are some optional ways to solve this problem:
• If we choose to use the classical deterministic algorithm, there we should make at least $2^{n-1} + 1$ queries in the worst case.

• If we use the classical random algorithm, there are at least order(1) queries.

• If we choose Deutsh-Jozsa’s algorithm, one query will suffice, which is due to the application of the quantum superposition principle.

### 10.4.3 Notation and lemma

- **State:**
  \[
  \begin{align*}
  |x⟩ &:= |x_1 x_2 \cdots x_n⟩, \\
  |y⟩ &:= |y_1 y_2 \cdots y_n⟩.
  \end{align*}
  \]
  (10.4.3a)

- **“Product”:**
  \[
  x \cdot y := x_y y_1 \oplus x_2 y_2 \oplus \cdots \oplus x_n y_n,
  \]
  (10.4.4)

  with
  \[
  x_i y_i := x_i \text{ AND } y_i.
  \]
  (10.4.5)

- **Summation:**
  \[
  \sum_{y \in \{0,1\}^n} := \sum_{y_1=0}^{1} \sum_{y_2=0}^{1} \cdots \sum_{y_n=0}^{1}.
  \]
  (10.4.6)

**Lemma 10.4.3.1.**

\[
H^{(n)}|x⟩ := \left( H \otimes H \otimes \cdots \otimes H \right) |x_1 x_2 \cdots x_n⟩
\]

\[
= \frac{1}{2^{n/2}} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} |y⟩.
\]

(10.4.7)

**Proof.**

\[
H^{(n)}|x⟩ = \left( H \otimes H \otimes \cdots \otimes H \right) |x_1 x_2 \cdots x_n⟩
\]

\[
= H|x_1⟩ \otimes H|x_2⟩ \otimes \cdots \otimes H|x_n⟩
\]

\[
= \left( \frac{1}{\sqrt{2}} \right)^n \sum_{y_1=0}^{1} \sum_{y_2=0}^{1} \cdots \sum_{y_n=0}^{1} (-1)^{x_1 y_1} |y_1⟩ \sum_{y_2=0}^{1} (-1)^{x_2 y_2} |y_2⟩ \cdots \sum_{y_n=0}^{1} (-1)^{x_n y_n} |y_n⟩
\]

\[
= \left( \frac{1}{\sqrt{2}} \right)^n \sum_{y_1=0}^{1} \sum_{y_2=0}^{1} \cdots \sum_{y_n=0}^{1} (-1)^{x_1 y_1 + x_2 y_2 + \cdots + x_n y_n} |y_1 y_2 \cdots y_n⟩
\]

\[
= \left( \frac{1}{\sqrt{2}} \right)^n \sum_{y_1=0}^{1} \sum_{y_2=0}^{1} \cdots \sum_{y_n=0}^{1} (-1)^{x_1 y_1 \oplus x_2 y_2 \oplus \cdots \oplus x_n y_n} |y_1 y_2 \cdots y_n⟩
\]

\[
= \frac{1}{2^{n/2}} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} |y⟩,
\]

(10.4.8)

where we have utilized E.Q. (10.3.20). Therefore, we have verified E.Q. (10.4.7). □

**Note:**

\[
H^{(n)}|0⟩ = \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x⟩.
\]

(10.4.9)
With the action of the \( n \)-fold Hadamard gates, we have the superposition of all the computational basis vectors, namely, all product basis vectors. In other words, all computational basis vectors can be encoded in the one quantum state due to the quantum superposition principle.

**Lemma 10.4.3.2** (Phase kick-back for \( n \)-qubit). Let’s define \( U_f \) as

\[
U_f : |a\rangle|y\rangle \mapsto |a\rangle|y \oplus f(a)\rangle
\]  

(10.4.10)

with

\[
f : \forall a \in \{0,1\}^n \mapsto f(a) \in \{0,1\},
\]

and

\[
y \in \{0,1\}.
\]

And we call \( |a\rangle \) as the first register, which is an \( n \)-qubit state, and \( |y\rangle \) as the second register, which is a single qubit.

Therefore,

\[
U_f |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}},
\]

(10.4.11)

with \( x \in \{0,1\}^n \). This can also be presented in the diagram formalism:

\[
\begin{array}{c}
|a_1\rangle \\
|a_2\rangle \\
\vdots \\
|a_n\rangle \\
\frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{array}
\begin{array}{c}
U_f \\
|a_1\rangle \\
|a_2\rangle \\
\vdots \\
|a_n\rangle \\
\frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{array}
\begin{array}{c}
(-1)^{f(x)} |a_1\rangle \\
(-1)^{f(x)} |a_2\rangle \\
\vdots \\
(-1)^{f(x)} |a_n\rangle \\
\frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{array}
\]

(10.4.12)

Proof.

\[
U_f |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}
\]

\[
= |x\rangle \frac{f(x) - |1 + f(x)\rangle}{\sqrt{2}}
\]

\[
= |x\rangle \frac{f(x) - \bar{f}(x)}{\sqrt{2}}
\]

\[
= (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}.
\]

(10.4.13)

We have made use of the same technique exploited to prove lemma 10.3.4.1 here. \( \square \)

### 10.4.4 Deutsch-Jozsa’s algorithm

Deutsch-Jozsa’s algorithm can be expressed in quantum circuit model:

\[
\begin{array}{c}
|0\rangle_1 \\
|0\rangle_2 \\
\vdots \\
|0\rangle_n \\
\frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{array}
\begin{array}{c}
H \\
H \\
\vdots \\
H
\end{array}
\begin{array}{c}
U_f \\
H \\
\vdots \\
H
\end{array}
\begin{array}{c}
H \\
H \\
\vdots \\
H
\end{array}
\begin{array}{c}
(-1)^{f(x)} |0\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}.
\end{array}
\]

(10.4.14)

where
(1) Input: state preparation.
\[ |\psi(t_1)\rangle = |0\rangle^{\otimes n} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \]

(2) \[ |\psi(t_2)\rangle = \left[ (H^{\otimes n}) \otimes I_2 \right] |\psi(t_1)\rangle \]
\[ = \frac{1}{2^n/2} \sum_{x \in \{0,1\}^n} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \]

(3) The phase kick-back technique.
\[ |\psi(t_3)\rangle = U_\omega |\psi(t_2)\rangle \]
\[ = \frac{1}{2^n/2} U_\omega \sum_{x \in \{0,1\}^n} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = \frac{1}{2^n/2} \sum_{x \in \{0,1\}^n} (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \]

(4) \[ |\psi(t_4)\rangle = \left[ (H^{\otimes n}) \otimes I_2 \right] |\psi(t_3)\rangle \]
\[ = \frac{1}{2^n/2} \left[ (H^{\otimes n}) \otimes I_2 \right] \sum_{x \in \{0,1\}^n} (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = \frac{1}{2^n} \sum_{x,y \in \{0,1\}^n} (-1)^{f(x) + x \cdot y} |y\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \]

(5) Output: quantum measurement.

If the post-measurement state of the first \( n \)-qubit is \( |0\rangle^{\otimes n} \), then the corresponding state of the \( n+1 \)-qubit system should be
\[ |\tilde{\psi}(t_4)\rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{f(x)} |0\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \] \[ (10.4.15) \]

Then, we can infer that

- if \( f \) is a constant function, i.e.,
\[ f(x) = f(x'), \forall x, x' \in \{0,1\}^n, \]
then the probability to find out that the first \( n \)-qubit is in the state \( |0\rangle^{\otimes n} \) is 1, since
\[ \langle 0|^{\otimes n} \langle 0 | \otimes I_2 |\psi(t_4)\rangle \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{f(0)} |0\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = (-1)^{f(0)} |0\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \] \[ (10.4.16) \]

and
\[ \langle 1|^{\otimes n} \langle 1 | \otimes I_2 |\psi(t_4)\rangle \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{f(x)+x} |1\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{f(0)+x} |1\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = (-1)^{f(0)} \frac{1}{2^n} \left[ \sum_{x \in \{0,1\}^n} (-1)^x \right] |1\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \]
\[ = 0, \] \[ (10.4.17) \]

with
\[ (-1)^x := (-1)^{x_1 + \cdots + x_n}. \]
• if $f$ is a balanced function, i.e.,

$$\# \{ x | f(x) = 0, \forall x \in \{0,1\}^n \} = \# \{ x | f(x) = 1, \forall x \in \{0,1\}^n \},$$

then the probability to find out that the first $n$-qubit is in the state $|n\rangle^n$ is 0, since

$$\langle 0 \rangle^n \langle 0 | \otimes I_2 | \psi(t_4) \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^f(x) |0\rangle^n |0\rangle - |1\rangle \sqrt{2}$$

$$= \frac{1}{2^n} \left[ \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \right] |0\rangle^n |0\rangle - |1\rangle \sqrt{2}$$

$$= 0.$$  \hspace{1cm} (10.4.18)

**Remark:** Quantum superposition principle makes it possible that computing $f(x)$ for all $x \in \{0,1\}^n$ simultaneously, which is impossible in classical computation.

### 10.5 Bernstein-Vazirani’s algorithm

Firstly, let’s introduce a lemma.

**Lemma 10.5.0.1.**

$$\sum_{x \in \{0,1\}^n} (-1)^{a+y} x = 2^n \delta_{a,y}, \hspace{0.5cm} \forall a, y \in \{0,1\}^n.$$  \hspace{1cm} (10.5.1)

This lemma can be proved in the following way.

**Proof.**

$$\sum_{x \in \{0,1\}^n} (-1)^{a+y} x = \sum_{x_1=0} \sum_{x_2=0} \cdots \sum_{x_n=0} (-1)^{a_1+y_1} x_1 (-1)^{a_2+y_2} x_2 \cdots (-1)^{a_n+y_n} x_n$$

$$= 2^n \delta_{a_1,y_1} \delta_{a_2,y_2} \cdots \delta_{a_n,y_n}$$

$$= 2^n \delta_{a,y},$$  \hspace{1cm} (10.5.2)

which is equivalent to E.Q.  \hspace{1cm} (10.5.1).

Let’s assume that there is a function $f_a$ defined as

$$f_a : \forall x \in \{0,1\}^n \rightarrow f_a(x) \in \{0,1\}.$$  \hspace{1cm} (10.5.3)

And we also know some facts about the function (black box) that

$$f_a(x) = a \cdot x$$

$$= a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_n x_n.$$  \hspace{1cm} (10.5.4)

**Question:** How can we determine $a$, which is an $n$-bit string?

**Solution 1** With classical algorithm, we need totally $n$ queries, i.e., $n$ equations, to get $a.$
**Solution 2** With Quantum Algorithm, only one query is sufficient to obtain \(a\). For the Quantum Algorithm, the corresponding Quantum Circuit model should be

\[
\begin{align*}
|0\rangle_1 & \quad H & |0\rangle_1 \\
|0\rangle_2 & \quad H & |0\rangle_2 \\
\vdots & \quad H & \vdots \\
|0\rangle_n & \quad H & |0\rangle_n \\
\frac{|0\rangle-|1\rangle}{\sqrt{2}} & \quad U_{f_a} & \frac{|0\rangle-|1\rangle}{\sqrt{2}},
\end{align*}
\]

(10.5.5)

where

1. \(|\psi(t_1)\rangle = |0\rangle^n \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}}\),
2. \(|\psi(t_2)\rangle = [(H^{\otimes n}) \otimes \mathbf{I}_2] |\psi(t_1)\rangle\)
   \[= \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}},\]
3. \(|\psi(t_3)\rangle = U_{f_a} |\psi(t_2)\rangle\)
   \[= \frac{1}{2^{n/2}} U_{f_a} \sum_{x \in \{0,1\}^n} |x\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}},\]
4. \(|\psi(t_4)\rangle = [(H^{\otimes n}) \otimes \mathbf{I}_2] |\psi(t_3)\rangle\)
   \[= \frac{1}{2^{n/2}} [(H^{\otimes n}) \otimes \mathbf{I}_2] \sum_{x \in \{0,1\}^n} (-1)^{a \cdot x} |x\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}}\]
   \[= \frac{1}{2^n} \sum_{x,y \in \{0,1\}^n} (-1)^{x \cdot (a+y)} |y\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}}\]
   \[= \sum_{y \in \{0,1\}^n} \delta_{a,y} |y\rangle \frac{|0\rangle-|1\rangle}{\sqrt{2}}\]
   \[= |a\rangle \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}}.\]

With the quantum measurement on the \(n\) register qubits, we can get \(|a\rangle\), and equivalently we obtain \(a = (a_1, a_2, \ldots, a_n)\).

### 10.6 Simon’s algorithm

Simon’s Algorithm

- has no phase kickback;
- is the 1-st algorithm that shows Quantum Algorithm can have exponential speed up for apparently hard problems.

### 10.7 Grover’s algorithm

Let’s now consider a search problem, the unstructured database search, with \(N \gg 1\) items. The unstructured database implies no special constrains or symmetries on the database. The task is to locate one particular term amid an unstructured set, with \(N\) items. Metaphorically, we say the task is to find a needle in a haystack, or to catch a particular fish in the sea.
10.7.1 Overview of the problem

Let’s assume that there is a black-box for computing an unknown function $f_\omega$ defined as

$$f_\omega : \forall x \in \{0,1\}^n \mapsto f_\omega(x) \in \{0,1\}.$$  \hfill (10.7.1)

\textbf{Question}: How can we find a unique marked term $x = \omega \in \{0,1\}^n$ such that

$$f_\omega(x) = \begin{cases} 1, & \text{if } x = \omega; \\ 0, & \text{if } x \neq \omega. \end{cases} \hfill (10.7.2a)$$

\textbf{Solution 1} With classical deterministic algorithm, we need $N = 2^n$ steps, in the order of $N$, to find $\omega$, namely checking the entire database item by item.

\textbf{Solution 2} With classical random algorithm, we need $N/2 = 2^{n-1}$ steps, in the order of $N$, to find $\omega$.

\textbf{Solution 3} With Grover’s algorithm which is a quantum algorithm, we only need the order of $(\sqrt{N})$ steps to find $\omega$, which is a big speed up if considering a huge database, like $N = 10^{10}$ items.

10.7.2 Grover’s algorithm

Let’s denote that the size of the database $\mathbb{Z}_2^n$ as $N = \#\mathbb{Z}_2^n = 2^n$. In Quantum Mechanics, it is possible to encode the entire database into one state, denoted as state $|s\rangle$ given by

$$|s\rangle := \frac{1}{\sqrt{N}} \sum_{x \in \{0,1\}^n} |x\rangle$$

$$= \frac{1}{\sqrt{N}} |\omega\rangle + \frac{1}{\sqrt{N}} \sum_{x \neq \omega} |x\rangle$$

$$= \frac{1}{\sqrt{N}} |\omega\rangle + \sqrt{\frac{N-1}{N}} |\omega^\perp\rangle, \hfill (10.7.3)$$

where

$$|\omega^\perp\rangle := \frac{1}{\sqrt{N-1}} \sum_{x \neq \omega} |x\rangle. \hfill (10.7.5)$$

Rewrite the state $|s\rangle$ into special type of two-dimensional Hilbert space $\mathcal{H}_2$ spanned by $\{ |\omega\rangle, |\omega^\perp\rangle \}$, namely, $\mathcal{H}_2 = \text{Span}\{ |\omega\rangle, |\omega^\perp\rangle \}$, where

$$\langle \omega | \omega^\perp \rangle = 0, \quad \langle \omega^\perp | \omega^\perp \rangle = 1. \hfill (10.7.6)$$

As we can show in Figure 10.1 where

$$\theta = \arcsin \frac{1}{\sqrt{N}} \hfill (10.7.7)$$

Therefore, we can also rewrite state $|s\rangle$ in the form of

$$|s\rangle = \sin \theta |\omega\rangle + \cos \theta |\omega^\perp\rangle. \hfill (10.7.8)$$
Figure 10.1: State $|s\rangle$ as a state vector in the two-dimensional Hilbert space $\mathcal{H}_2 = \text{Span}\{ |\omega\rangle, |\omega^\perp\rangle \}$.

**Question:** How to get $|\omega\rangle$ from $|s\rangle$? The quantum search problem becomes the problem of how to rotate $|s\rangle$ to $|\omega\rangle$ in the two-dimensional Hilbert space $\mathcal{H}_2 = \text{Span}\{ |\omega\rangle, |\omega^\perp\rangle \}$. The hint is that a geometric rotation can be decomposed as a product of two reflection operations.

Now, we will show the way to find $|\omega\rangle$ from $|s\rangle$ in the following steps.

1° The first reflection operator $U_\omega$ is defined as

$$U_\omega := I_2 - 2 |\omega\rangle \langle \omega|.$$  \hfill (10.7.9)

And we can get that

$$U_\omega |\omega\rangle = (I_2 - 2 |\omega\rangle \langle \omega|) |\omega\rangle = |\omega\rangle - 2 |\omega\rangle = - |\omega\rangle,$$

$$U_\omega |\omega^\perp\rangle = (I_2 - 2 |\omega\rangle \langle \omega|) |\omega^\perp\rangle = |\omega^\perp\rangle.$$  \hfill (10.7.10)

And we can show that

$$U_\omega^2 = (I_2 - 2 |\omega\rangle \langle \omega|)^2 = I_2 - 4 |\omega\rangle \langle \omega| + (2 |\omega\rangle \langle \omega|)^2 = I_2.$$  \hfill (10.7.11)

Thus, with $U_\omega$ acting on the state $|s\rangle$, we can get

$$|s'\rangle := U_\omega |s\rangle = U_\omega (\sin \theta |\omega\rangle + \cos \theta |\omega^\perp\rangle) = - \sin \theta |\omega\rangle + \cos \theta |\omega^\perp\rangle,$$  \hfill (10.7.12)

which can be shown in Figure 10.2 and the geometric implication of operator $U_\omega$ is reflecting the state $|s\rangle$ around the $|\omega^\perp\rangle$ axis.

2° The second reflection operator $U_s$ is defined as

$$U_s := 2 |s\rangle \langle s| - I_2.$$  \hfill (10.7.13)
As we can see that

\[
U_s |s\rangle = (2 |s\rangle \langle s| - I_2) |s\rangle = 2|s\rangle - |s\rangle = |s\rangle, \quad (10.7.15)
\]

\[
U_s |s^\perp\rangle = (2 |s\rangle \langle s| - I_2) |s^\perp\rangle = -|s^\perp\rangle, \quad (10.7.16)
\]

where

\[
\langle s|s^\perp\rangle = 0 \quad \text{and} \quad \langle s^\perp|s^\perp\rangle = 1. \quad (10.7.17)
\]

With the second reflection operator acting on the first reflected state, we can get

\[
|s'_1\rangle := U_s |s'\rangle = U_s U_\omega |s\rangle = (2 |s\rangle \langle s| - I_2) (|s\rangle - \sin \theta |\omega\rangle + \cos \theta |\omega^\perp\rangle)
\]

\[
= -2 \sin \theta |s\rangle + 2 \cos \theta |s\rangle \langle s|\omega\rangle + \sin \theta |\omega\rangle - \cos \theta |\omega^\perp\rangle
\]

\[
= -2 \sin^2 \theta |s\rangle + 2 \cos^2 \theta |s\rangle + \sin \theta |\omega\rangle - \cos \theta |\omega^\perp\rangle
\]

\[
= 2 \cos 2 \theta |s\rangle + \sin \theta |\omega\rangle - \cos \theta |\omega^\perp\rangle
\]

\[
= (2 \cos 2 \theta + 1) \sin \theta |\omega\rangle + (2 \cos 2 \theta - 1) \cos \theta |\omega^\perp\rangle
\]

\[
= (\cos 2 \theta + 2 \cos^2 \theta \sin \theta |\omega\rangle) + (\cos 2 \theta - 2 \sin^2 \theta) \cos \theta |\omega^\perp\rangle
\]

\[
= (\cos 2 \theta \sin \theta + \cos \theta \sin 2 \theta) |\omega\rangle + (\cos 2 \theta \cos \theta - \sin 2 \theta \sin \theta) |\omega^\perp\rangle
\]

\[
= \sin 3 \theta |\omega\rangle + \cos 3 \theta |\omega^\perp\rangle. \quad (10.7.18)
\]

And this can be shown in Figure 10.3 as a kind of geometrical interpretation. The second reflection operator is defined as reflecting the state around \(|s\rangle\) axis.

3\(^\circ\) Grover’s rotation is defined as

\[
R_{\text{grov}} := U_s \circ U_\omega. \quad (10.7.19)
\]

Therefore, we can attain

\[
|s'_1\rangle = R_{\text{grov}} |s\rangle. \quad (10.7.20)
\]
As we can see that the $R_{\text{grov}}$ rotates the original state $|s\rangle$ by the angle $2\theta$. We can define further that

$$R^n_{\text{grov}} = R_{\text{grov}} \circ R_{\text{grov}} \circ \cdots \circ R_{\text{grov}}.$$  

Therefore, we can infer that

$$R^n_{\text{grov}} |s\rangle = \sin(2n + 1)\theta |\omega\rangle + \cos(2n + 1)\theta |\omega^\perp\rangle.$$  

This is the so-called Grover iteration.

As we shall see that, in the Grover problem, the angle $\theta$ is very small, since

$$\sin \theta = \frac{1}{\sqrt{N}} = \frac{1}{2n-1}, \text{ with } n \gg 1.$$  

Suppose that with $T$-iteration we can get the marked $|\omega\rangle$, therefore

$$\begin{align*}
(2T + 1)\theta &= \frac{\pi}{2} \\
\sin \theta &= \frac{1}{\sqrt{N}}
\end{align*}$$

$$\begin{cases}
T = \frac{\pi}{4\theta} - \frac{1}{2} \\
\theta \approx \frac{1}{\sqrt{N}},
\end{cases}$$

namely

$$T = \frac{\sqrt{N} \pi}{4} \left[ 1 + O(N^{-1/2}) \right] \propto \sqrt{N}.$$  

10.7.3 Example: $N = 4$

$$f_\omega: \forall x \in \{0,1\}^2 \mapsto f_\omega(x) \in \{0,1\},$$  

more explicitly, it can be rewritten as

$$\begin{align*}
(0,0) &\mapsto f_\omega(0,0), \\
(0,1) &\mapsto f_\omega(0,1), \\
(1,0) &\mapsto f_\omega(1,0), \\
(1,1) &\mapsto f_\omega(1,1).
\end{align*}$$
Suppose that $|10\rangle$ is the marked item we are looking for in the database $\mathbb{Z}_2^2$, i.e.,

$$f_\omega(1,0) = 1, \quad f_\omega(0,0) = f(0,1) = f(1,1) = 0,$$

and the initial state encoded the database can be set as

$$|s\rangle = \sin \theta |\omega\rangle + \cos \theta |\omega^\perp\rangle,$$

where

$$\theta = \frac{\pi}{6}, \quad \text{and} \quad |\omega\rangle = |10\rangle.$$

With one Grover's rotation $R_{\text{grov}}$, we can get the initial state rotated counterclockwise through the angle $2\theta = \frac{\pi}{3}$, which will lead to the angle between second reflected state and the state $|\omega^\perp\rangle$ as

$$2\theta + \theta = 3\theta = \frac{\pi}{2},$$

which implies that the marked state $|10\rangle$ is found through one $R_{\text{grov}}$ operation. Check the result in algebraic approach

$$|s\rangle = \frac{1}{2}|10\rangle + \frac{\sqrt{3}}{2}|10^\perp\rangle;$$

$$|s'\rangle = U_{\omega}|s\rangle$$

$$= -\frac{1}{2}|10\rangle + \frac{\sqrt{3}}{2}|10^\perp\rangle;$$

$$|s'_1\rangle = U_sU_{\omega}|s\rangle$$

$$= \sin \left(\frac{\pi}{6} \times 2 + \frac{\pi}{6}\right)|10\rangle + \cos \left(\frac{\pi}{6} \times 2 + \frac{\pi}{6}\right)|10^\perp\rangle;$$

$$= |10\rangle$$

This can be shown in the Figure 10.4.

**Figure 10.4: One Grover’s Rotation with $N = 4$**

**Note:** The other diagrammatical approach to quantum search algorithm is shown below.

$$|s\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

$$= \frac{1}{2} \left| \begin{array}{cccc}
00 & 01 & 10 & 11
\end{array} \right\rangle$$

109
\[ |s'\rangle = U_\omega |s\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle + |11\rangle) \]

\[ = \frac{1}{2} \left| \begin{array}{cccc}
00 & 01 & 10 & 11 \\
\end{array} \right\rangle \]  \hspace{1cm} (10.7.33)

\[ |s'_1\rangle = U_1 U_\omega |s\rangle = |10\rangle \]

\[ = \left| \begin{array}{cccc}
00 & 01 & 10 & 11 \\
\end{array} \right\rangle \]  \hspace{1cm} (10.7.34)

The “needle” is represented for the state, and the direction of the needle is represented for the relative sign of the state.

### 10.7.4 Quantum circuit model of Grover’s algorithm

1° Initial state is prepared in the way as

\[ |s\rangle = H^\otimes n |0\rangle^\otimes n, \]

in quantum circuit model, which can be shown as

\[ |0\rangle \longrightarrow H \\
|s\rangle = |0\rangle \longrightarrow H \\
\vdots \\
|0\rangle \longrightarrow H. \]

(10.7.36)

2° As we know that

\[ U_\omega = I_2 - 2|\omega\rangle \langle \omega|, \]

we can define the unitary gate \( U_{f_\omega} \) as

\[ U_{f_\omega} : |x\rangle |y\rangle \mapsto |x\rangle |y \oplus f_\omega(x)\rangle, \]

with \(|x\rangle\) being a \(n\)-qubit state, and \(|y\rangle\) being a single qubit state, i.e.,

\[ \begin{array}{l}
\text{n qubits} \\
|y\rangle \\
\text{single qubit} \\
\end{array} \longrightarrow \begin{array}{c}
\text{H} \\
|y \oplus f_\omega(x)\rangle \.
\end{array} \]

(10.7.39)

Therefore, we can get the phase kick-back

\[ U_{f_\omega} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f_\omega(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \]

(10.7.40)

**Lemma 10.7.4.1.**

\[ U_{f_\omega} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (U_\omega \otimes I_2) |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \]

(10.7.41)

with the function \( f_\omega \) defined as

\[ f_\omega(x) = \begin{cases} 
1, & \text{if } x = \omega; \\
0, & \text{if } x \neq \omega.
\end{cases} \]

(10.7.42a)
This lemma can be represented in the form of Quantum Circuit diagram

\[ U_{f,\omega} \left( \omega \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) = (\omega \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}) \]

As we shall see that with \(|\omega\rangle\) as an unknown state, \(U_{f,\omega}\) is actually an oracle.

**Proof.**

- In the case \(|x\rangle = |\omega\rangle\).

\[
U_{f,\omega} \left( |\omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) = (-1)^{f(\omega)} |\omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
= -|\omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
= U_{\omega} |\omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} ; \\
\]

- In the case \(|x\rangle \neq |\omega\rangle\).

\[
U_{f,\omega} \left( |x \neq \omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) = (-1)^{f(x \neq \omega)} |x \neq \omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
= |x \neq \omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
= U_{\omega} |x \neq \omega\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} . \\
\]

**Note:** How to perform \(U_{\omega}\) has to use of the black box \(U_{f}\), which suggests that performing \(U_{\omega}\) takes the main part of the total performing time.

3° The second reflection operator

\[
U_s = 2 |s\rangle \langle s| - I_2, \\
\]

is at hand, because its realization does not involve the black-box and only includes elementary quantum gates. And we can derive that

\[
U_s = H^{\otimes n} \left( 2 |0\rangle^{\otimes n} \langle 0| - I_2^{\otimes n} \right) H^{\otimes n} \\
= H^{\otimes n} X^{\otimes n} \left( 2 |1\rangle^{\otimes n} \langle 1| - I_2^{\otimes n} \right) X^{\otimes n} H^{\otimes n} \\
= -H^{\otimes n} X^{\otimes n} \left( I_2^{\otimes (n-1)} \otimes H \right) H^{(n)} \left( I_2^{\otimes (n-1)} \otimes H \right) X^{\otimes n} H^{\otimes n}, \\
\]

(10.7.47)
where $\theta^{(n)}$ is the $n$-qubit Toffoli gate, which can be written as

$$\theta^{(n)} = (I_2^\otimes(n-1) - |1\rangle^\otimes(n-1) \langle 1|) \otimes I_2 + \left( |1\rangle^\otimes(n-1) \langle 1| \right) \otimes X. \quad (10.7.48)$$

Therefore, we can verify the E.Q. (10.7.47) in the following manner,

$$\left( I_2^\otimes(n-1) \otimes H \right) \theta^{(n)} \left( I_2^\otimes(n-1) \otimes H \right)$$

$$= \left( I_2^\otimes(n-1) \otimes H \right) \left[ \left( I_2^\otimes(n-1) - |1\rangle^\otimes(n-1) \langle 1| \right) \otimes I_2 + \left( |1\rangle^\otimes(n-1) \langle 1| \right) \otimes X \right] \left( I_2^\otimes(n-1) \otimes H \right)$$

$$= \left( I_2^\otimes(n-1) - |1\rangle^\otimes(n-1) \langle 1| \right) \otimes I_2 + \left( |1\rangle^\otimes(n-1) \langle 1| \right) \otimes Z$$

$$= I_2^\otimes(1) + |1\rangle^\otimes(n-1) \langle 1| \otimes (-2|1\rangle \langle 1|)$$

$$= I_2^\otimes(n) - 2|1\rangle^\otimes(n) \langle 1|. \quad (10.7.49)$$

Now, we can make use of the above components to construct the quantum circuit model for Grover’s algorithm.

**Step 1** The quantum circuit should be

\[ \begin{array}{c}
|0\rangle_1 \\
|0\rangle_2 \\
\vdots \\
|0\rangle_n \\
|0\rangle_{\langle -1|} \sqrt{2} \\
H^\otimes n \\
\vdots \\
\tilde{R}_{\text{grov}} \\
\vdots \\
\tilde{R}_{\text{grov}} \\
\vdots \\
\tilde{R}_{\text{grov}} \\
\vdots \\
\vdots \\
|0\rangle_{\langle -1|} \sqrt{2}
\end{array} \]

Note that $\tilde{R}_{\text{grov}} := (U_s \otimes I_2) U_{\omega}$.

**Step 2** Suppose that the result we get in **Step 1** is state $|\omega\rangle'$, then we can use the oracle $U_{\omega}$ to check if it is the state $|\omega\rangle$, i.e.,

\[ |\omega\rangle' \rightarrow U_{\omega} \]

\[ |0\rangle_{\langle -1|} \sqrt{2} \rightarrow |0\rangle_{\langle -1|} \sqrt{2}. \quad (10.7.51) \]

With the measurement, we can do the check. If we find out the result is wrong, then we should start from **Step 1** again. If it is the rightful state, namely, the measurement results are 1, then the work is done.
Chapter 11

Quantum Circuit Complexity

11.1 Circuit complexity

11.1.1 Definitions

Def 11.1.1 (Classical circuit model). A classical circuit model is finite sequence of elementary gates applied to a specified string of input bits.

Def 11.1.2 (Circuit Complexity). Circuit complexity is a quantity which quantifies how many resources are exploited in computation, such as the number of elementary gates.

Thm 11.1.1.1 (Characterization of Circuit Complexity). Circuit complexity is characterized by the size, depth, and width of the smallest circuit to compute a given problem.

Def 11.1.3 (Size, Depth, Width). The size, depth, and width of the circuit can be defined as below:

1. Size: the number of elementary gates used to solve a problem;
2. Depth: the number of time steps to solve a problem;
3. Width: the maximum number of gates that act in any one time step.

Example: Construction of the GHZ state is shown in Figure 11.1. There we can get the complexity of this circuit:

1° Size: there are one Hadamard gate $H$, and $(n - 1)$ CNOT gates;
2° Depth: there are totally $n$ time steps, which is a polynomial of $n$ and that is a good thing;
3° Width: since there is only one gate is exploited for every time step, then the depth should be 1.
11.1.2 Complexity class

We can classify the circuit into different categories according to the size the circuit:

\[
\begin{align*}
\text{size} & \propto \text{poly}(n) \quad \text{excellent,} \\
\text{size} & \propto \text{Exp}(n) \quad \text{the worst case.}
\end{align*}
\]

**Def 11.1.4** (Decision problem). A decision problem is a “yes” or “no” problem.

**Def 11.1.5** (P-problem).

\[P\text{-problem} = \{\text{decision problems which can be solved by polynomial size circuit family}\}\]

**Def 11.1.6** (NP-problem).

\[NP\text{-problem} = \{\text{decision problems which can be verified by polynomial size circuit family}\}\]

**Note:** verifying a problem is very different from solving a problem. For example, it would be very difficult to work out a password, but with a password in hand, it would be very easy to verify if it works.

Question: What’s the relation between N-problem and NP-problem? It is widely believed, but not solved, that these two categories of problems are not equivalent to each other. But, how can we prove or falsify that kind of consensus?

**Thm 11.1.2.1** (Cook’s theorem(1971)). Every problem in NP is polynomially reducible to NP-Completeness (NPC) problem.

Some relations between the different types of problems expressed in diagram formalism shown in Figure 11.2.

![Decision problem, NP, NPC, NPI and P](image)

Figure 11.2: Decision problem, NP, NPC, NPI and P

To be clarified in the future (2015).

11.1.3 Quantum complexity

There are two kinds of classical circuit, namely the deterministic circuits and the probabilistic circuit with probabilistic gates:

Classical Circuit

\[
\begin{align*}
\text{deterministic circuits,} \\
\text{probabilistic circuits with probabilistic gates}
\end{align*}
\]

Probabilistic gates

\[
\begin{align*}
\text{AND} & \quad \text{with probability } \varepsilon, \\
\text{OR} & \quad \text{with probability } 1 - \varepsilon,
\end{align*}
\]

e.g., simulation Monte Carlo Annealing. Classical circuit model is described by Bounded-error Probabilistic Polynomial size (BPP) problem. Quantum circuit, on the other hand, is described by Bounded-error Quantum Probabilistic Polynomial size (BQP) problem.
11.1.4 Accuracy

For the exact Quantum Computer, we need infinite number of quantum gates. But, that is impossible to realize in practice. For practical consideration, we would turn to Finite Quantum Approximately Universal Gate set, with which we can get the result in arbitrary accuracy. So the precision is an important topics.
Chapter 12

Quantum Simulation
Part III

Density Matrix and Quantum Entanglement
Chapter 13

Quantum Mechanics (II): Density Matrix

References:

- [Preskill] Chapter 2: Foundations I: states and ensembles;

13.1 Density matrix as state of quantum open system

Density matrix (operator) describes the state of a quantum open system. It can be introduced in both physical and mathematical approaches. Usually, the quantum computer is an open subsystem, and with the environment by which suggests lots of noises, and they form a closed system.

In the mathematical sense, we can view the projective measurement theory in terms of the density operator,

\[
\rho = \sum_{a_k} \langle \psi | P_n | a_k \rangle \langle a_k | \psi \rangle
\]

(13.1.1)

with the density operator defined as

\[
\rho := |\psi\rangle \langle \psi|
\]

(13.1.2)

and all \(|a_k\rangle \in \{|a_j\rangle | A |a_j\rangle = a_j |a_j\rangle, |a_j\rangle \in \mathcal{H}\} \) being normalized. Therefore,

\[
\langle A \rangle = \sum_n a_n \langle \psi | P_n | \psi \rangle = \sum_n a_n \text{tr}(\rho P_n) = \text{tr}(\rho A).
\]

(13.1.3)

We can claim that the probability to obtain the outcome \(a_n\) when measuring \(A\) is actually

\[
\text{Prob}(a_n) = \text{tr}(\rho P_n).
\]

(13.1.4)
In the physical sense, for the quantum closed system with the initial state $|\psi\rangle$,

$$|\psi\rangle = \sum_n c_n |n\rangle,$$

where $\{ |n\rangle \}$ is an orthonormal basis for the Hilbert space $\mathcal{H}$. The probability to get the post-measurement states $|n\rangle$ is

$$\text{Prob}(|n\rangle) = \frac{|\langle n|\psi\rangle|^2}{|\langle \psi|\psi\rangle|^2}. \quad (13.1.5)$$

**Question:** What is the post-measurement state of an open system?

$$\rho := \sum_n p_n |n\rangle \langle n|, \quad \text{with} \quad \sum_n p_n = 1 \quad \text{and} \quad p_n \geq 0. \quad (13.1.6)$$

The $\rho$ is the so-called density matrix, which will be defined later. And the density operator $\rho$ is equivalent to state ensemble $\{ |n\rangle, p_n \}$, in the Quantum Mechanical sense.

**Remark.** In the description of the density matrix, the ambiguity of the global phase factor with non-physical meaning is removed,

$$|\psi\rangle \equiv e^{i\alpha} |\psi\rangle, \quad (13.1.7)$$

$$\rho = |\psi\rangle \langle \psi|, \quad \rho_\alpha = e^{i\alpha} |\psi\rangle \langle \psi|e^{-i\alpha} = \rho. \quad (13.1.8)$$

### 13.1.1 State ensemble formalism of density matrix

**Def 13.1.1.** The state ensemble is a set of state $|\psi_i\rangle$ each attached with a probability $p_i$, namely

$$\{ |\psi_i\rangle, p_i \mid i = 1, \ldots, n \}, \quad \text{with} \quad p_i \in [0, 1] \quad \text{and} \quad \sum_{i=1}^nP_i = 1. \quad (13.1.9)$$

It means that the physical system has the probability of $p_i$ to be prepared in the state $|\psi_i\rangle$.

For any observable $A$, the expectation value of $A$ is

$$\langle A \rangle = \sum_{i=1}^n p_i \langle \psi_i | A | \psi_i \rangle$$

$$= \sum_{k=1}^{\dim \mathcal{H}} \sum_{i=1}^n p_i \langle \psi_i | k \rangle \langle k | A | \psi_i \rangle$$

$$= \sum_{k=1}^{\dim \mathcal{H}} \sum_{i=1}^n \langle k | A | \psi_i \rangle p_i \langle \psi_i | k \rangle$$

$$= \sum_{k=1}^{\dim \mathcal{H}} \langle k | A \sum_{i=1}^n (|\psi_i\rangle p_i \langle \psi_i|) |k \rangle$$

$$= \sum_{k=1}^{\dim \mathcal{H}} \langle k | A \rho | k \rangle$$

$$= \text{tr}(A \rho),$$

with $\rho = \sum_{i=1}^n p_i |\psi_i\rangle \langle \psi_i|$, 

$$\langle A \rangle = \sum_{i=1}^n p_i \langle \psi_i | A | \psi_i \rangle = \text{tr}(A \rho), \quad (13.1.10)$$

where $\{ |k\rangle \mid k = 1, \ldots, \dim \mathcal{H} \}$ is an orthonormal basis for the Hilbert space $\mathcal{H}$.
Def 13.1.2. The density matrix for the system described by the state ensemble, as shown in Eq. (13.1.9), is defined as

\[ \rho = \sum_{i=1}^{n} p_i |\psi_i\rangle \langle \psi_i| , \]  

(13.1.11)

- if \( n = 1 \), this represents a pure state, and the system is completely specified;
- if \( n \geq 2 \), this represents a mixed state, and the system is partially specified.

Thm 13.1.1.1. A sufficient and necessary condition to tell whether the density operator \( \rho \) represents a pure state or a mixed state is

\[
\begin{align*}
\text{pure state} & \iff \rho^2 = \rho, \quad \text{tr}\rho^2 = 1; \\
\text{mixed state} & \iff \rho^2 \neq \rho, \quad \text{tr}\rho^2 < 1.
\end{align*}
\]  

(13.1.12)

Proof. As we shall show later that the density operators are non-negative and self-adjoint with unit trace. It is always possible to use a set of orthonormal eigenvectors of the density operator \( \rho \) as the basis to expand the whole Hilbert space \( \mathcal{H} \). In that way, the density matrix \( \rho \) is diagonalized, with its eigenvalues \( \lambda_i \) lying along the diagonal position of the matrix:

\[ \rho = \sum_{i=1}^{\dim\mathcal{H}} \lambda_i |\lambda_i\rangle \langle \lambda_i| . \]  

(13.1.13)

There we can also get \( \rho^2 \):

\[
\rho^2 = \sum_{i=1}^{\dim\mathcal{H}} \lambda_i \lambda_j |\lambda_i\rangle \langle \lambda_i| \langle \lambda_j| \lambda_j\rangle \\
= \sum_{i,j=1}^{\dim\mathcal{H}} \lambda_i \lambda_j |\lambda_i\rangle \langle \lambda_j| \delta_{ij} \\
= \sum_{i=1}^{\dim\mathcal{H}} \lambda_i^2 |\lambda_i\rangle \langle \lambda_i| .
\]  

(13.1.14)

\[
\text{tr}\rho^2 = \sum_{i=1}^{\dim\mathcal{H}} \lambda_i^2 \leq \left( \sum_{i=1}^{\dim\mathcal{H}} \lambda_i \right)^2 = 1.
\]  

(13.1.15)

Therefore, the sufficient and necessary condition that \( \rho \) is equal to \( \rho^2 \) should be

\[ \lambda_i = \lambda_i^2, \quad \text{with} \quad i = 1, \ldots, \dim\mathcal{H}. \]  

(13.1.16)

With \( \rho \) being self-adjoint and non-negative with unit trace, namely

\[ \sum_{i=1}^{\dim\mathcal{H}} \lambda_i = 1, \quad \text{with} \lambda_i \in \mathbb{R} \text{ and } \lambda_i \geq 0, \quad i = 1, \ldots, \dim\mathcal{H}, \]  

(13.1.17)

it then can be inferred that Eq. (13.1.16) is equivalent to that there should be only one \( \lambda_j \) being 1 while all other \( \lambda_i \) with \( i \neq j \) are vanishing. And it also says that the same thing as Eq. (13.1.12), since with only one eigenvalue for the density operator, it is sufficient and necessary to get the pure state. \( \square \)

Thm 13.1.1.2. The density operator \( \rho \) should be subject to the following time-evolution equation

\[ i\hbar \frac{\partial}{\partial t} \rho(t) = [H, \rho(t)]. \]  

(13.1.18)
Notice: We can see that E.Q. (13.1.18) is very similar to the Heisenberg equation, but it’s not a Heisenberg equation at all. This can be shown in the following way.

Proof. To get the time evolution of the physical system, we may start from the state ensemble as given in E.Q. (13.1.9). Every state vector $|\psi_i(t)\rangle$ within the state ensemble should follow the Shrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi_i(t)\rangle = H |\psi_i(t)\rangle,$$

or we can get the state ensemble at time $t$

$$\{ |\psi_i(t)\rangle , p_i | \; i = 1, \ldots, n \}, \quad (13.1.19)$$

with

$$|\psi_i(t)\rangle := \mathcal{U}(t) |\psi_i\rangle , \quad (13.1.20)$$

where $\mathcal{U}(t)$ is defined in E.Q. (3.1.17). Therefore, the density at time $t$ should be

$$\rho(t) = \sum_{i=1}^{n} p_i |\psi_i(t)\rangle \langle \psi_i(t)| , \quad (13.1.21)$$

Now, we can derive the time-evolution equation for $\rho$:

$$i\hbar \frac{\partial}{\partial t} \rho(t) = \sum_{i=1}^{n} p_i \left( i\hbar \frac{\partial}{\partial t} |\psi_i(t)\rangle \langle \psi_i(t)| + |\psi_i(t)\rangle \langle \psi_i(t)| i\hbar \frac{\partial}{\partial t} \langle \psi_i(t)| \right)$$

$$= \sum_{i=1}^{n} p_i \left( H |\psi_i(t)\rangle \langle \psi_i(t)| + |\psi_i(t)\rangle \langle \psi_i(t)| (-H) \right)$$

$$= \left[ \rho, H \right]$$

which is equivalent to E.Q. (13.1.18).

13.1.2 Operator formalism of density matrix

Now, we come to discuss the properties of the density matrix.

Def 13.1.3. The density matrix $\rho$ is a non-negative self-adjoint operator with unit trace, i.e.,

$$\begin{cases} 
\rho^i = \rho \text{ self-adjoint operator}, \\
\rho \geq 0 \text{ non-negative}, \\
\text{tr} \rho = 1 \text{ unit trace}.
\end{cases} \quad (13.1.22)$$

Notice: $\rho$ being non-negative means that all its eigenvalues are non-negative, which is also equivalent to that given any state its expectation value should be non-negative.

Thm 13.1.2.1. The operator formalism of the density matrix $\rho$ is equivalent to its state ensemble description.

Proof. To derive the three properties of the density operator $\rho$, we may start from the definition (13.1.11), for it not only represents the mixed state case, but also the pure state case.
• Self-adjoint

\[ \rho^\dagger = \left( \sum_{i=1}^{n} p_i |\psi_i\rangle \langle \psi_i| \right)^\dagger \]  
\[ = \sum_{i=1}^{n} p_i^* |\psi_i\rangle \langle \psi_i| \]  
\[ = \sum_{i=1}^{n} p_i |\psi_i\rangle \langle \psi_i| \]  
\[ = \rho. \]  

(13.1.23)

since \( p_i \in \mathbb{R} \) for \( i = 1, \ldots, n \).

• Non-negative

\[ \langle \phi \mid \rho \mid \phi \rangle = \sum_{i=1}^{n} p_i \langle \phi \mid \psi_i \rangle \langle \psi_i \mid \phi \rangle \]  
\[ = \sum_{i=1}^{n} p_i |\langle \psi_i \mid \phi \rangle|^2, \]  

(13.1.24)

where \( |\phi\rangle \) is an arbitrary vector in Hilbert space \( \mathcal{H} \). Given \( p_i \geq 0 \) for all \( i = 1, \ldots, n \), it’s obvious that \( \langle \phi \mid \rho \mid \phi \rangle \) is non-negative, namely \( \rho \) is non-negative.

• Unit trace

\[ \text{tr}\rho = \text{tr}\left( \sum_{i=1}^{n} p_i |\psi_i\rangle \langle \psi_i| \right) \]  
\[ = \sum_{i=1}^{n} p_i \text{tr}(|\psi_i\rangle \langle \psi_i|) \]  
\[ = \sum_{i=1}^{n} p_i \]  
\[ = 1. \]  

(13.1.25)

because \( p_i \) with \( i = 1, \ldots, n \) would add up to 1 and all \( |\psi_i\rangle \) are assumed to be normalized.

13.1.3 Reduced density matrix (State for subsystem)

Let’s assume that the state of the composite physical system \( C \) consisted of subsystem \( A \) and subsystem \( B \), is in state \( |\psi\rangle_C \). Then, we can get the density matrix for the composite physical system \( C \),

\[ \rho_{AB} = |\psi\rangle_{AB} \langle \psi|. \]  

(13.1.32)

That represents a pure state, as we consider the composite physical system as a whole. But, now we examine only the subsystem \( A \) (Alice) of the composite system. We can get the observable \( M_A \) for the subsystem of Alice, expressed in Hilbert space \( \mathcal{H}_{AB} \) of the composite system \( C \) as

\[ M_A \otimes I_B. \]

Therefore, the expectation value of the observable \( M_A \) is then

\[ \langle M_A \rangle = AB \langle \psi \mid M_A \otimes I_B |\psi\rangle_{AB} \]  
\[ = \text{tr}\left( (M_A \otimes I_B)\rho_{AB} \right) \]  
\[ = \text{tr}_A(M_A\rho_A), \]  

(13.1.33)
where
\[ \rho_A = \text{tr}_B \rho_{AB}. \]  
(13.1.34)

And we have used the facts that
\[
\begin{align*}
\text{tr} \mathcal{O}_{AB} &= \text{tr}_A (\text{tr}_B \mathcal{O}_{AB}), \\
\text{tr} (M_A \mathcal{O}_{AB}) &= \text{tr}_A (M_A \text{tr}_B \mathcal{O}_{AB}),
\end{align*}
\]
which can be proved simply by expressing all the operators in the form of kets and bras.

The reduced density matrix \( \rho_A \) can describe the subsystem A, given the density operator of the compound physical system is \( \rho_{AB} \), in the following sense. If \( \mathcal{O}_A \) is an operator corresponding to an arbitrary observable defined in the subsystem A, namely \( \mathcal{O}_A \in \mathcal{B}(\mathcal{H}_A) \), then we can express it in the total Hilbert space \( \mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B \) as
\[ \mathcal{O}_A \otimes I_B. \]  
(13.1.36)

Therefore, the expectation value of the observable \( \mathcal{O}_A \) should be
\[
\begin{align*}
\langle \mathcal{O}_A \rangle &= \text{tr}_{AB} (\mathcal{O}_A \rho_{AB}) \\
&= \text{tr}_A \left[ \text{tr}_B (\mathcal{O}_A \rho_{AB}) \right] \\
&= \text{tr}_A (\mathcal{O}_A \text{tr}_B \rho_{AB}) \\
&= \text{tr}_A (\mathcal{O}_A \rho_A).
\end{align*}
\]  
(13.1.37)

On the other hand, it really makes a density operator because

- **self-adjoint:**
\[
\begin{align*}
\rho_A^\dagger &= (\text{tr}_B \rho_{AB})^\dagger \\
&= \text{tr}_B \rho_{AB}^\dagger \\
&= \text{tr}_B \rho_{AB} \\
&= \rho_A;
\end{align*}
\]  
(13.1.38)

- **unit trace:**
\[
\begin{align*}
\text{tr}_A \rho_A &= \text{tr}_A (\text{tr}_B \rho_{AB}) \\
&= \text{tr}_{AB} \rho_{AB} \\
&= 1;
\end{align*}
\]  
(13.1.39)

- **nonnegative:**
\[
\begin{align*}
A \langle \psi | \rho_A | \psi \rangle_A &= \text{tr}_A (\rho_A | \psi \rangle_A \langle \psi |) \\
&= \text{tr}_{AB} (\rho_{AB} | \psi \rangle_A \langle \psi | \otimes I_B)]
\end{align*}
\]

namely
\[
A \langle \psi | \rho_A | \psi \rangle_A \geq 0, \quad \forall | \psi \rangle_A \in \mathcal{H}_A
\]  
(13.1.40)

since the probability for the compound system \( \rho_{AB} \) found in state \( | \psi \rangle_A \otimes I_B \) should be nonnegative.

*E.g.* Reduced density matrix \( \rho_A \) of the general bipartite system
\[
| \psi \rangle_{AB} = \sum_i \sum_\mu a_{i\mu} | i \rangle_A \otimes | \mu \rangle_B,
\]  
(13.1.41)

\[
\rho_A = \text{tr}_B \rho_{AB} = \sum_\mu \langle \mu | \rho_{AB} | \mu \rangle_B = \sum_\mu \sum_i \sum_j a_{i\mu} a_{ij}^* | i \rangle_A \langle j |,
\]  
(13.1.42)

where the partial trace
\[
\text{tr}_B (| \mu \rangle_B \langle \nu |) = \langle \nu | \mu \rangle_B = \delta_{\mu\nu}.
\]  
(13.1.43)

123
13.2 Mixed state formalism of a qubit

**Thm 13.2.0.1.** Density matrix in the two-dimension Hilbert space $\mathcal{H}_2$, can be represented as

$$\rho(\vec{p}) = \frac{1}{2}(I_2 + \vec{p}\vec{\sigma}), \quad (13.2.1)$$

with $\vec{p} \in \mathbb{R}^3$ satisfying $||\vec{p}|| \leq 1$. And we give the vector $\vec{p}$ the name of the “polarization vector”.

**Proof.** Firstly, we would prove that any density matrix in the two-dimensional Hilbert space $\mathcal{H}_2$ can be expressed in the form as E.Q. (13.2.1). Secondly, we would show that with the constraint $||\vec{p}|| \leq 1$, E.Q. (13.2.1) is a density matrix.

a) Any density matrix in the 2-dimensional Hilbert space $\mathcal{H}_2$ can be represented in the form of E.Q. (13.2.1).

The most general 2-dimensional matrix would have the form

$$\rho = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (13.2.2)$$

which has $2 \times 4 = 8$ degrees of freedom. But to make a density matrix for a physical system, the matrix $\rho$ must satisfy three constraints.

(i) **Self-adjoint**: A density matrix must be self-adjoint, i.e., $\rho^\dagger = \rho$, which means

$$\begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \iff \begin{cases} a^* = a \\ d^* = d \\ b^* = c \\ c^* = b \end{cases}$$

which actually makes four constraints for $\rho$, that would reduce the degrees of freedom from eight to four.

(ii) **Unit trace**: The trace of a density matrix must be equal to one, i.e.,

$$\text{tr}\rho = 1 \iff a + d = 1, \quad (13.2.4)$$

which makes a fifth constraint to $\rho$. The degrees of freedom has come down to three now.

(iii) **Positive**: The density matrix must have all its eigenvalues to be non-negative.

This would appear as inequalities, namely “holonomic constraints” in the language of Classical Mechanics, which would never reduce a freedom but could absolutely set some kind of range for the parameters.

On the other hand, we know that, the density matrices for the Hilbert space $\mathcal{H}_2$ can actually be viewed as a vector space $\mathcal{B}(\mathcal{H}_2)$. The three Pauli matrices plus the identity matrix can be the vector basis for $\mathcal{B}(\mathcal{H}_2)$, because they are linearly independent to each other. Therefore, it would be clear that any density matrix can be expressed in the form of E.Q. (13.2.1).
Now, we set to prove that iff $|\bar{p}| \leq 1$, E.Q. (13.2.1) would be a density matrix. From E.Q. (13.2.1) we get

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + p_3 & p_1 - ip_2 \\ p_1 + ip_2 & 1 - p_3 \end{pmatrix} \quad \text{with} \quad p_1, p_2, p_3 \in \mathbb{R}. \quad (13.2.5)$$

This definition implies apparently that $\rho$ is self-adjoint. And we can immediately get the trace and the determinant of $\rho$:

$$\text{tr}\rho = \frac{1}{2} ((1 + p_3) + (1 - p_3)) = 1, \quad (13.2.6)$$

$$\det\rho = \frac{1}{4}[(1 + p_3)(1 - p_3) - (p_1 - ip_2)(p_1 + ip_2)] = \frac{1}{4}(1 - |\bar{p}|^2). \quad (13.2.7)$$

Let’s denote the two eigenvalues of $\rho$ with $\lambda_1$ and $\lambda_2$, then

$$\left\{ \begin{array}{lcl} \text{tr}\rho &=& \lambda_1 + \lambda_2 = 1, \\
\det\rho &=& \lambda_1\lambda_2 = \frac{1}{4}(1 - |\bar{p}|^2), \end{array} \right.$$ 

from which we can infer that $\rho$ being nonnegative is equivalent to

$$\left\{ \begin{array}{lcl} \lambda_1 &\geq& 0, \\
\lambda_2 &\geq& 0 \end{array} \right. \iff |\bar{p}| \leq 1. \quad (13.2.8)$$

Therefore, $\rho$ defined in E.Q. (13.2.1) is a density matrix iff $|\bar{p}| \leq 1$.

We can now claim that, any density operator for two-level system attached with Hilbert space $\mathcal{H}_2$, can be written as E.Q. (13.2.1).

### 13.2.1 Why polarization vector?

We firstly calculate the expectation value $\langle \hat{\sigma} \cdot \bar{n} \rangle$, given the physical system is in the state described by the density operator $\rho$ defined in E.Q. (13.2.1),

$$\langle \hat{\sigma} \cdot \bar{n} \rangle = \text{tr}[ (\hat{\sigma} \cdot \bar{n}) \rho ] = \text{tr} \left[ \frac{1}{2} (\hat{\sigma} \cdot \bar{n}) (1 + \hat{\bar{p}} \hat{\sigma}) \right] = \frac{1}{2} \left[ \text{tr}(\hat{\sigma} \cdot \bar{n}) + \text{tr}( (\hat{\sigma} \cdot \bar{n})(\hat{\bar{p}} \hat{\sigma}) ) \right]. \quad (13.2.9)$$

Amid, we can get

$$\text{tr}(\hat{\sigma} \cdot \bar{n}) = \text{tr}(n_1 \sigma_1 + n_2 \sigma_2 + n_3 \sigma_3) = n_1 \text{tr}\sigma_1 + n_2 \text{tr}\sigma_2 + n_3 \text{tr}\sigma_3 = 0, \quad (13.2.10)$$

where we have used the fact that

$$\text{tr}\sigma_i = 0, \quad \text{with} \quad i = 1, 2, 3. \quad (13.2.11)$$
And, we can also get

\[
\text{tr}\left[(\hat{\sigma} \cdot \hat{n})(\hat{\rho} \cdot \hat{\sigma})\right] = \text{tr}\left(\sum_{i,j=1}^{3} n_i p_j \sigma_i \sigma_j\right)
\]

\[
= \sum_{i,j=1}^{3} n_i p_j \text{tr}(\sigma_i \sigma_j)
\]

\[
= 2 \sum_{i,j=1}^{3} n_i p_j \delta_{ij}
\]

\[
= 2 \hat{n} \cdot \hat{p}.
\]  

(13.2.12)

In the above derivation we have used the relation

\[
\text{tr}(\sigma_i \sigma_j) = 2 \delta_{ij},
\]  

(13.2.13)

which can be derived in the following manner

\[
i \neq j: \quad \text{tr}(\sigma_i \sigma_j) = \text{tr}(i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k) = i \sum_{k=1}^{3} \epsilon_{ijk} \text{tr}\sigma_k = 0 \Rightarrow \text{tr}(\sigma_i \sigma_j) = 2 \delta_{ij}, \quad i, j = 1, 2, 3.
\]

Now, by substituting E.Q. (13.2.10) and E.Q. (13.2.12) into E.Q. (13.2.9), we can conclude that

\[
\langle \hat{\sigma} \cdot \hat{n} \rangle = \hat{n} \cdot \hat{p},
\]

which is equivalent to

**Thm 13.2.1.1.**

\[
\text{tr}\left((\hat{\sigma} \cdot \hat{n}) \rho(\hat{p})\right) = \hat{n} \cdot \hat{p},
\]

(13.2.14)

with \( \hat{n}, \hat{p} \in \mathbb{R}^3 \) and \( \|\hat{n}\| = 1 \).

With this theorem, we can understand why we call \( \hat{p} \) the “polarization vector”:

(i) \( \langle \hat{\sigma} \cdot \hat{n} \rangle = \text{tr}\left((\hat{\sigma} \cdot \hat{n}) \rho(\hat{p})\right) = 0 \), for \( \hat{p} \perp \hat{n} \);

(ii) \( \langle \hat{\sigma} \cdot \hat{n} \rangle = \text{tr}\left((\hat{\sigma} \cdot \hat{n}) \rho(\hat{p})\right) = 1 \), for \( \hat{p} = \hat{n} \);

(iii) \( \langle \hat{\sigma} \cdot \hat{n} \rangle = \text{tr}\left((\hat{\sigma} \cdot \hat{n}) \rho(\hat{p})\right) = -1 \), for \( \hat{p} = -\hat{n} \);

(iv) \( \langle \hat{\sigma} \cdot \hat{n} \rangle = \text{tr}\left((\hat{\sigma} \cdot \hat{n}) \rho(\hat{p})\right) = 0 \), for \( \hat{p} = 0 \), \( \forall \hat{n} \in \mathbb{R}^3 \).

We can see that \( \hat{p} = 0 \) implies no polarization orientation, which shows “no information”.

**13.2.2 Pure state and mixed state in two-dimensional Hilbert space \( \mathcal{H}_2 \)**

**Thm 13.2.2.1.** If the state in the two-dimensional Hilbert space \( \mathcal{H}_2 \), which means that the state has the density matrix defined in E.Q. (13.2.1). Then the state of the physical system is

\[
\begin{align*}
\text{a pure state}, & \quad \text{if} \ \hat{p} \ \text{is on the Bloch sphere,} \ |\hat{p}| = 1, \\
\text{a mixed state}, & \quad \text{if} \ \hat{p} \ \text{is inside the Bloch sphere,} \ |\hat{p}| < 1.
\end{align*}
\]
We can show that

$$\rho^2(\vec{p}) = \frac{1}{4} \left( 1 + 2 \hat{\vec{p}} \vec{\sigma} + (\hat{\vec{p}} \vec{\sigma})^2 \right) = \frac{1}{4} \left( 1 + 2 \hat{\vec{p}} \vec{\sigma} + \rho^2 \right), \tag{13.2.15}$$

because

$$\begin{align*}
(\hat{\vec{p}} \vec{\sigma})^2 &= \sum_{i=1}^{\tilde{N}} p_i^2 \sigma_i^2 + \sum_{i,j \neq i} p_i p_j \{\sigma_i, \sigma_j\} \\
&= \sum_{i=1}^{\tilde{N}} p_i^2 + 0 \\
&= |\hat{\vec{p}}|^2.
\end{align*}$$

Therefore,

$$\begin{align*}
\rho^2(\hat{\vec{p}}) &= \rho(\hat{\vec{p}}), \text{ if } \hat{\vec{p}} \text{ is on the Bloch sphere, } |\hat{\vec{p}}| = 1; \\
\rho^2(\hat{\vec{p}}) &\neq \rho(\hat{\vec{p}}), \text{ if } \hat{\vec{p}} \text{ is inside the Bloch sphere, } |\hat{\vec{p}}| < 1.
\end{align*} \tag{13.2.16}$$

Thus $\rho^2(\hat{\vec{p}}) \neq \rho(\hat{\vec{p}})$ or $|\hat{\vec{p}}| < 1$ is sufficient to claim that $\rho(\hat{\vec{p}})$ is a mixed state. On the other hand, we can get some examples of case $|\hat{\vec{p}}| = 1$:

- $\hat{\vec{p}} = \hat{e}_z, \quad \rho(\hat{e}_z) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |0\rangle \langle 0| = |\uparrow_z\rangle \langle \uparrow_z|;

- $\hat{\vec{p}} = \hat{e}_x, \quad \rho(\hat{e}_x) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} (|0\rangle + |1\rangle) (|0\rangle + |1\rangle) = |+\rangle \langle +| = |\uparrow_x\rangle \langle \uparrow_x|;

- $\hat{\vec{p}} = \hat{e}_y, \quad \rho(\hat{e}_y) = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} = \frac{1}{2} (|0\rangle + i|1\rangle) (|0\rangle - i|1\rangle) = |\uparrow_y\rangle \langle \uparrow_y|.

In general, if $|\hat{\vec{p}}| = 1$, we can express $\hat{\vec{p}}$ as $\hat{\vec{p}} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$,

$$\rho(\hat{\vec{p}}) = \frac{1}{2} \begin{pmatrix} 1 + \cos \theta & \sin \theta \cos \varphi - i \sin \theta \sin \varphi \\ \sin \theta \cos \varphi + i \sin \theta \sin \varphi & 1 - \cos \theta \end{pmatrix}.$$ 

Therefore, $|\hat{\vec{p}}| = 1$ is a sufficient condition for that $\rho(\hat{\vec{p}})$ is a pure state.

Given that any state for a two-level physical system, such as spin-1/2 systems, which can be described by $\rho(\hat{\vec{p}})$ as defined in E.Q. \[13.2.1\], we can conclude that now:

- $|\hat{\vec{p}}| = 1$ is a sufficient and necessary condition for that $\rho(\hat{\vec{p}})$ is a pure state;

- $|\hat{\vec{p}}| < 1$ is a sufficient and necessary condition for that $\rho(\hat{\vec{p}})$ is a mixed state.
Pure state vs. Mixed state in 2-dimensional Hilbert space

<table>
<thead>
<tr>
<th>pure state</th>
<th>mixed state</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\psi\rangle = \frac{1}{\sqrt{2}}(</td>
</tr>
<tr>
<td>$\sigma_x</td>
<td>\uparrow\rangle =</td>
</tr>
<tr>
<td>$\rho =</td>
<td>\psi\rangle \langle \psi</td>
</tr>
<tr>
<td>$\rho = \frac{1}{2}(</td>
<td>0\rangle \langle 0</td>
</tr>
</tbody>
</table>

Measurement $M_x = |\uparrow\rangle \langle \uparrow|$, $N_x = |\downarrow\rangle \langle \downarrow|$  

$\langle M_x \rangle = 1$, $\langle N_x \rangle = 0$  
$\langle M_x \rangle = \frac{1}{2}$, $\langle N_x \rangle = \frac{1}{2}$

13.3 Convexity of density matrix

Thm 13.3.0.2 (Convexity of density matrix). The density matrices form a convex set, and the pure states are the external points of this convex set.

This is transparent for the two-dimensional Hilbert space $\mathcal{H}_2$. As we can show that by using the concept of the Bloch ball, in $\mathcal{H}_2$, the most general density matrix can be expressed as

$$\rho(\vec{p}) := \frac{1}{2} \left( I_2 + \hat{p} \hat{\sigma} \right), \quad \text{with} \ |\vec{p}| \leq 1. \quad (13.3.1)$$

- If $|\vec{p}| = 1$, then $\rho(\vec{p})$ represents a pure state. $|\vec{p}| = 1$ also means that the state can be represented as one point on the Bloch sphere. So, the pure states is the external points of the Bloch ball, which is a convex set.

- If $|\vec{p}| < 1$, then $\rho(\vec{p})$ represents a mixed state. $|\vec{p}| < 1$ also means that the state can be represented as one point inside the Bloch ball.

- For any vector $\vec{p}$ with the constraint $|\vec{p}| \leq 1$, we can express it in the form of

$$\vec{p} = \vec{n}_2 + \lambda(\vec{n}_1 - \vec{n}_2), \quad (13.3.2)$$

with $0 \leq \lambda \leq 1$ and $|\vec{n}_1| = |\vec{n}_2| = 1$, $\vec{n}_1, \vec{n}_2 \in \mathbb{R}^3$. Therefore, we can get the density matrix $\rho(\vec{p})$ expressed as

$$\rho(\vec{p}) = \frac{1}{2} \left( I_2 + [\vec{n}_2 + \lambda(\vec{n}_1 - \vec{n}_2)] \hat{\sigma} \right)$$

$$= \frac{1}{2} \left[ \lambda(I_2 + \vec{n}_1 \hat{\sigma}) + (1 - \lambda)(I_2 + \vec{n}_2 \hat{\sigma}) \right]$$

$$= \frac{\lambda}{2}(I_2 + \vec{n}_1 \hat{\sigma}) + \frac{1 - \lambda}{2}(I_2 + \vec{n}_2 \hat{\sigma}),$$

i.e.,

$$\rho(\vec{p}) = \lambda \rho(\vec{n}_1) + (1 - \lambda) \rho(\vec{n}_2), \quad (13.3.3)$$

with

$$\rho(\vec{n}_1) = \frac{1}{2} (I_2 + \vec{n}_1 \hat{\sigma}),$$

$$\rho(\vec{n}_2) = \frac{1}{2} (I_2 + \vec{n}_2 \hat{\sigma}).$$

128
Therefore, we can also get
\[
\rho(\bar{\rho}) \iff \{\lambda, |\psi_+ (\bar{n}_1)\rangle; (1 - \lambda), |\psi_+ (\bar{n}_2)\rangle\}.
\]

(13.3.4)

Remark: The same mixed state density matrix may have different interpretations of state ensembles. We can make an easy example to demonstrate this, for instance
\[
\rho := \frac{1}{2} I.
\]

(13.3.5)

The density matrix defined in this form can be interpreted as state ensemble consisted of an arbitrary orthonormal basis vector set for the Hilbert space $\mathcal{H}$, with each state vector attached with the probability $1/\dim \mathcal{H}$.

Question: Why can the same mixed state density matrices have different interpretations of state ensembles?

Information encoded in the density matrix of the mixed state (which may denote the subsystem of the entangled pure state), can only be partially specified. And the entanglement information (as partial information of the composite system) can not be reveled in the subsystem individually, which leads to the ambiguity of the interpretation of state ensembles. As we could see, the state is more entangling, the information is more hidden.

13.4 Two-qubit system and its subsystem

Two-qubit system and its subsystem

<table>
<thead>
<tr>
<th>entire system</th>
<th>subsystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>two-qubit $\in \mathcal{H}_2 \otimes \mathcal{H}_2$</td>
<td>one-qubit $\in \mathcal{H}_2$</td>
</tr>
<tr>
<td>density matrix</td>
<td>reduced density matrix</td>
</tr>
<tr>
<td>pure state</td>
<td>(maybe) mixed state</td>
</tr>
<tr>
<td>(exactly known)</td>
<td>(partially known or completely unknown)</td>
</tr>
</tbody>
</table>

13.4.1 Example: EPR pair (Bell state)

We consider a state of the composite physical system, which is made up of two subsystems A (Alice) and B (Bob),
\[
|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}\left( |00\rangle + |11\rangle \right)_{AB} = \frac{1}{\sqrt{2}}\left( |0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\right).
\]

(13.4.1)

And the corresponding density matrix should be
\[
\rho_{AB} := |\psi\rangle_{AB} \langle \psi| = \frac{1}{2} \left( |00\rangle_{AB} \langle 00| + |01\rangle_{AB} \langle 01| + |10\rangle_{AB} \langle 10| + |11\rangle_{AB} \langle 11| \right).
\]

(13.4.2)

We can check that $\rho_{AB}^2 = \rho_{AB}$,
\[
\rho_{AB}^2 = |\psi\rangle_{AB} \langle \psi| \rho_{AB} \langle \psi| = |\psi\rangle_{AB} \langle \psi| = \rho_{AB},
\]

(13.4.3)
amid we have use the fact that

$$\rho_{AB} = \frac{1}{2} \left( |0 \rangle_B \langle 0|_A \otimes |0 \rangle_B + |1 \rangle_A \otimes |1 \rangle_B \right)$$

Therefore, the entire physical system is in pure state.

Now, we consider the reduced density matrix for the subsystem $A$ (Alice), which is a one-qubit system,

$$\rho_A := \text{tr}_B \rho_{AB}. \quad \text{(13.4.4)}$$

From above definition, we can get

$$\rho_A = \sum_{i=0}^{1} b_i |i \rangle_B \langle i|_B$$

$$= \frac{1}{2} \left( |0 \rangle_A \langle 0| + |1 \rangle_A \langle 1| \right)$$

$$= \frac{1}{2} I_2. \quad \text{(13.4.5)}$$

For an arbitrary orientation, $\hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$,

$$\langle \hat{n} \vec{\sigma}_A \rangle = \text{tr}_A \left[ (\hat{n} \vec{\sigma}_A) \rho_A \right]$$

$$= \frac{1}{2} \text{tr}_A (\hat{n} \vec{\sigma}_A)$$

$$= 0. \quad \text{(13.4.6)}$$

Therefore, in this case the polarization vector for the subsystem $A$ should be a null vector in $\mathbb{R}^3$, and we can extract “no information” from this subsystem.

**Remark**: Generally, the composite system with pure state, is exactly known. While, for the subsystem, which is usually in mixed state, is often partially known or nothing known.

**Def 13.4.1 (Maximally entanglement)**. If a bipartite $\rho_{AB}$ is a pure state with $\rho_A = \frac{1}{2} I_A$, then $\rho_{AB}$ is called maximally entangled.

### 13.4.2 Maximally entangled two-qubit pure states

**Thm 13.4.2.1**. All Bell states are maximally entangled.

**Proof**. This is transparent from E.Q. [4.3.11],

$$|\psi(i,j)\rangle_{AB} = \frac{1}{\sqrt{2}} \left[ |i\rangle_{AB} \pm (-1)^j |\bar{i}\rangle_{AB} \right]$$

130
But, Alice and Bob can avoid it by Bell measurements $X$ing each other their measurements results. A third party, Eve. Eve wants to destroy the state shared by Alice and Bob, i.e., destroy Alice and Bob share a maximal entangled state (e.g. the Bell state $\psi$).

\[
\rho_A = \text{tr}_B |\psi(i,j)\rangle_{AB} \langle \psi(i,j)| \\
= B \langle i | \psi(i,j) \rangle_{AB} \langle \psi(i,j) | i \rangle_B + B \langle i | \psi(i,j) \rangle_{AB} \langle \psi(i,j) | i \rangle_B \\
= \frac{1}{2} |0\rangle_A \langle 0 | + \frac{1}{2} |1\rangle_A \langle 1 | \\
= \frac{1}{2} I_A, 
\]

and

\[
\rho_B = \text{tr}_A |\psi(i,j)\rangle_{AB} \langle \psi(i,j)| \\
= B \langle 0 | \psi(i,j) \rangle_{AB} \langle \psi(i,j) | 0 \rangle_B + B \langle 1 | \psi(i,j) \rangle_{AB} \langle \psi(i,j) | 1 \rangle_B \\
= \frac{1}{2} |i\rangle_A \langle i | + \frac{1}{2} |i\rangle_A \langle i | \\
= \frac{1}{2} I_B.
\]

Therefore, we can conclude now that every Bell state $|\psi(i,j)\rangle$ is maximally entangled. □

**Remark:**

1. Local unitary transformations, e.g. $I_2 \otimes X_i Z_j$, preserve the entangling property;
2. With $\rho_A = \rho_B = \frac{1}{2} I_2$, we cannot acquire any information from local measurement on any one of the two subsystems.
3. For product pure state, there is no entanglement. For example, $|\phi\rangle_{AB} = |\bar{n}\rangle_A \otimes |\bar{n}\rangle_B$, has the corresponding reduced density matrices

\[
\begin{align*}
\rho_A &= |\bar{n}\rangle_A \langle \bar{n} | \\
\rho_B &= |\bar{n}\rangle_B \langle \bar{n} |
\end{align*}
\]

with $|\bar{n}| = |\bar{n}| = 1$. There is no information can be hidden in the subsystem in this case. We may say that the more entangled a state is, the more hidden of the quantum information is.

### 13.4.3 Monogamy of maximal entanglement

Alice and Bob share a maximal entangled state (e.g. the Bell state $|\phi^+\rangle$). And here comes a third party, Eve. Eve wants to destroy the state shared by Alice and Bob, i.e., destroy $|\phi^+\rangle$.

\[
\begin{align*}
|\phi^+\rangle_{AB} \otimes |0\rangle_E & \quad \text{separable} \\
\downarrow & \\
|\psi\rangle_{ABE} & \quad \text{entangled} \\
& = |00\rangle_{AB} |e_{00}\rangle_E + |01\rangle_{AB} |e_{01}\rangle_E \\
& \quad + |10\rangle_{AB} |e_{10}\rangle_E + |11\rangle_{AB} |e_{11}\rangle_E.
\end{align*}
\]

But, Alice and Bob can avoid it by Bell measurements $X_A \otimes X_B$ and $Z_A \otimes Z_B$ and informing each other their measurements results.
• After measuring $Z_A \otimes Z_B$, they get the result of 1, then

$$\frac{\ket{\psi}_{ABE}}{\text{entangled}} \downarrow \frac{\ket{\psi'}_{ABE}}{\text{entangled}} = \ket{00}_{AB} |e_{00}\rangle_E + \ket{11}_{AB} |e_{11}\rangle_E \quad (13.4.10)$$

• After measurement of $X_A \otimes X_B$, the result is 1, then the final state has to be

$$\frac{\ket{\psi'}_{ABE}}{\text{entangled}} \downarrow \frac{\ket{\psi''}_{ABE}}{\text{separable}} = \frac{1}{\sqrt{2}} (\ket{00}_{AB} + \ket{11}_{AB}) \otimes |e\rangle_E \cdot \quad (13.4.11)$$

in which the state of Eve $|e\rangle_E$ has been decoupled with $|\phi^+\rangle_{AB}$.
Chapter 14

Schmidt Decomposition, Purification and GHJW Theorem

References:
- [Preskill] Chapter 2: Foundations I: states and ensembles;

14.1 Introduction

Let’s consider a bipartite composite system with two subsystems A (Alice) and B (Bob), interacting with each other. The state of the composite system can be described by the pure state $\psi_{AB}$ with the corresponding density matrix $\rho_{AB} = |\psi\rangle_{AB}\langle\psi|$. And the density matrix of the subsystem A should be $\rho_A = \text{tr}_B \rho_{AB} = \text{tr}_A (|\psi\rangle_{AB}\langle\psi|)$. Before further talking about the measurement and quantum operation on the subsystem of the composite system, in this chapter, we would introduce some relations between subsystem and composite system. And Figure 14.1 illustrates the plan for this chapter.

![Figure 14.1: Plan for chapter 14](image)

14.2 Schmidt decomposition and quantum entanglement

14.2.1 Schmidt decomposition

**Thm 14.2.1.1** (Schmidt decomposition). In a bipartite system $\mathbb{H}_A \otimes \mathbb{H}_B$, a bipartite pure state $|\psi\rangle_{AB}$ can be expressed in the form of

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i} |i\rangle_A \otimes |i\rangle_B,$$

with $p_i \geq 0$, $\sum_i p_i = 1$, (14.2.1)
where \{i\}_A and \{i\}'_B are orthonormal basis for \mathcal{H}_A and \mathcal{H}_B respectively. This is the so-called Schmidt decomposition.

Example: Bell state (EPR state)

\[ |\psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right) \]

\[ = \frac{1}{\sqrt{2}} \sum_{i=0}^{1} |ii\rangle_{AB} \]

\[ = \frac{1}{\sqrt{2}} \sum_{i=0}^{1} |i\rangle_A \otimes |i\rangle_B , \]

(14.2.2)

with \sqrt{p_0} = \sqrt{p_1} = \frac{1}{\sqrt{2}}. We can also get

\[ \begin{aligned}
\rho_A &= \frac{1}{2} I_A , \\
\rho_B &= \frac{1}{2} I_B .
\end{aligned} \]

(14.2.3)

Remark: For \[ |\psi\rangle_{AB} = \sum_i \sqrt{p_i} |i\rangle_A \otimes |i\rangle_B , \]
we may have \dim \mathcal{H}_A \neq \dim \mathcal{H}_B. But \rho_A and \rho_B would have the same diagonalized formalism, i.e.,

\[ \rho_A = \sum_i p_i |i\rangle_A \langle i| , \quad \rho_B = \sum_i p_i |i\rangle_B \langle i| . \]

(14.2.4)

Def 14.2.1. Schmidt coefficient and Schmidt number:

1. The non-vanishing \(p_i\) in E.Q. (14.2.1) are called Schmidt coefficients.

2. The number of Schmidt coefficients is called Schmidt number.

14.2.2 Quantum entanglement

Def 14.2.2. Separable and Entangled bipartite pure states:

• if the Schmidt number is 1, \[ |\psi\rangle_{AB} \] is called separable, which means that the state vector of the composite physical system can be expressed as

\[ |\psi\rangle_{AB} = |\phi\rangle_A \otimes |\chi\rangle_B , \]

(14.2.5)

which is the so-called product state;

• if the Schmidt number is greater than 1, then the state \[ |\psi\rangle_{AB} \] is called entangled state;

• if the reduced density matrices for the subsystems are all identity matrices, namely

\[ \begin{aligned}
\rho_A &= \frac{1}{2} I_A , \\
\rho_B &= \frac{1}{2} I_B ,
\end{aligned} \]

(14.2.6)

then \[ |\psi\rangle_{AB} \] is called maximally entangled. (We have actually mentioned this at the end of the last chapter (Ref. page 130).)

Note: We would discuss about the entanglement definition for bipartite mixed state and multipartite mixed state later.
14.2.3 Proof for the theorem of the Schmidt decomposition

**Proof.** The key point to get the Schmidt decomposition of an arbitrary state $|\psi\rangle_{AB}$ defined as

$$|\psi\rangle_{AB} := \sum_{i,\mu} b_{i\mu} \langle i | A \otimes | \mu \rangle_B ,$$

(14.2.7)

with $\{|i\rangle_A\}$ and $\{|\mu\rangle_B\}$ being orthonormal basis for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, is to get both the reduced density matrices $\rho_A$ and $\rho_B$ diagonalized. And therefore the corresponding density matrix of the composite physical system should be

$$\rho_{AB} := |\psi\rangle_{AB} \langle \psi| = \sum_{i,j,\mu,\nu} b_{i\mu} b^{*}_{j\nu} \langle i | A \otimes | j \rangle_B \otimes | \mu \rangle_B \langle \nu | .$$

(14.2.8)

Now, we can go on step by step.

**Step 1:**

Compute $\rho_A$ by the taking the partial trace of $\rho_{AB}$ over $\mathcal{H}_B$:

$$\rho_A = \operatorname{tr}_B \rho_{AB} = \sum_{\mu'} \langle \mu' | B \rho_{AB} | \mu' \rangle_B = \sum_{i,\mu,j,\nu,\nu'} b_{i\mu} b^{*}_{j\nu} \langle i | A \otimes | j \rangle_B \otimes (| \mu \rangle_B \langle \mu' | B \langle \nu | \rangle_B) = \sum_{i,\mu,j} b_{i\mu} b^{*}_{j\mu} \langle i | A \otimes | j \rangle_B .$$

(14.2.9)

**Step 2:**

Diagonalize $\rho_A$ by choosing another orthonormal basis $\{|i\rangle_A\}$, all vectors in which are all eigenvectors of $\rho_A$, such that

$$\rho_A = \sum_i p_i |i\rangle_A \langle i| , \quad \text{with} \quad \sum_i p_i = 1 , \quad \text{and} \quad p_i \geq 0 ,$$

(14.2.10)

since $\rho_A$ is the density matrix for the subsystem $A$, which should be nonnegative and self-adjoint operator with unit trace. And $\{|i\rangle_A\}$ are associated with a unitary transformation.

**Step 3:**

With the basis $\{|i\rangle_A\}$ for $\mathcal{H}_A$ and the basis $\{|\mu\rangle_B\}$ for $\mathcal{H}_B$, we can reformulate the state vector for the composite system,

$$|\psi\rangle_{AB} = \sum_{i,\mu} a_{i\mu} |i\rangle_A \otimes | \mu \rangle_B .$$

(14.2.11)

Now, we can choose another basis for $\mathcal{H}_B$ to simplify the expression of $|\psi\rangle_{AB}$:

$$|\tilde{i}\rangle_B := A \langle i | \psi \rangle_{AB} = \sum_{j,\mu} a_{j\mu} A \langle i | j \rangle_A \otimes | \mu \rangle_B = \sum_{\mu} a_{i\mu} | \mu \rangle_B ,$$

(14.2.12)

with which we can rewrite the state vector $|\psi\rangle_{AB}$ again,

$$|\psi\rangle_{AB} = \sum_i |i\rangle_A \otimes |\tilde{i}\rangle_B .$$

(14.2.13)
Correspondingly, the density matrix $\rho_{AB}$ should take the form of

$$\rho_{AB} = \sum_{i,j} |i\rangle_A \langle j| \otimes |\tilde{i}\rangle_B \langle \tilde{j}|. \quad (14.2.14)$$

**Step 4:**

We can prove that the newly chosen basis $\{|\tilde{i}\rangle_B\}$ is orthogonal and we also could normalize it. To accomplish those, again we can compute $\rho_A$ through partial trace calculation of $\rho_{AB}$ over $\mathcal{H}_B$:

$$\rho_A = \sum_\mu p_\mu |\tilde{i}\rangle_A \langle j| \otimes (|\tilde{j}\rangle_B \langle \mu|_B \otimes |\mu\rangle_B) \quad (14.2.15)$$

On the other hand, we know from E.Q. (14.2.10) that

$$\rho_A = \sum_{i,j} p_i \delta_{ij} |i\rangle_A \langle j|. \quad (14.2.16)$$

By comparing E.Q. (14.2.15) with E.Q. (14.2.16), we can get that

$$B \langle \tilde{j}\rangle_B = p_i \delta_{ij}, \quad (14.2.17)$$

which means that the basis $\{|\tilde{i}\rangle_B\}$ is actually orthogonal.

Now we can normalize $\{|\tilde{i}\rangle_B\}$ by defining

$$|i'\rangle_B := \frac{1}{\sqrt{p_i}} |\tilde{i}\rangle_B, \quad (14.2.18)$$

where $p_i > 0$, which means that we are discussing in the subspaces of $\mathcal{H}_A$ and $\mathcal{H}_B$ which ensures $p_i > 0$.

Now, with the basis $\{|i\rangle\}$ for $\mathcal{H}_A$ and $\{|i'\rangle\}$ for $\mathcal{H}_B$ we can rewrite the state vector for the third time

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i} |i\rangle_A \otimes |i'\rangle_B, \quad (14.2.1)$$

which is exactly the same as E.Q. (14.2.1).

### 14.3 Example for the Schmidt decomposition

As an example of the Schmidt decomposition, here we present the solution of the exercise 2.2, Chapter 2 of John Preskill’s online lecture notes.

**Problem:** For the two-qubit state

$$|\Phi\rangle = \frac{1}{\sqrt{2}} |\uparrow\rangle_A \left(\frac{1}{2} |\uparrow\rangle_B + \frac{\sqrt{3}}{2} |\downarrow\rangle_B\right) + \frac{1}{\sqrt{2}} |\downarrow\rangle_A \left(\frac{\sqrt{3}}{2} |\uparrow\rangle_B + \frac{1}{2} |\downarrow\rangle_B\right) \quad (14.3.1)$$

(a) Compute $\rho_A = \text{tr}_B(|\Phi\rangle\langle\Phi|)$ and $\rho_B = \text{tr}_A(|\Phi\rangle\langle\Phi|)$.
(b) Find the Schmidt decomposition of $|\Phi\rangle$.

**Notation:**

(1) $|\uparrow\rangle = |0\rangle$, $|\downarrow\rangle = |1\rangle$.

(2) 2×2 matrix:

$$N = \sum_{i,j=0}^{1} |i\rangle N_{ij} |j\rangle,$$

where $i$ is the row index and $j$ is the column index.

(3) 4×4 matrix:

$$M = \sum_{i,j,k,l=0}^{1} |i,j\rangle M_{ij,kl} |k,l\rangle,$$

where $i, j$ are the row indexes and $k, l$ are the column indexes.

(4) Product basis in Hilbert space $\mathcal{H}_2 \otimes \mathcal{H}_2$:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} ; |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} ; |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Rewrite the state $|\Psi\rangle_{AB}$ as

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{2}} |0\rangle_A \left( \frac{1}{2} |0\rangle_B + \frac{\sqrt{3}}{2} |1\rangle_B \right) + \frac{1}{\sqrt{2}} |1\rangle_A \left( \frac{\sqrt{3}}{2} |0\rangle_B + \frac{1}{2} |1\rangle_B \right)$$

$$= \frac{1}{\sqrt{8}} \left( |00\rangle_{AB} + \sqrt{3} |01\rangle_{AB} + \sqrt{3} |10\rangle_{AB} + |11\rangle_{AB} \right)$$

$$= \frac{1}{\sqrt{8}} \sum_{i,j=0}^{1} a_{ij} |i,j\rangle_{AB}.$$

And in the matrix expression, we have

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{8}} \begin{pmatrix} 1 \\ \sqrt{3} \\ \sqrt{3} \\ 1 \end{pmatrix}.$$

**Note:** $|\Psi\rangle_{AB}$ is symmetric under the exchange of system $A$ and system $B$, which implies $\rho_A = \rho_B$. 

137
\[ \rho_{AB} = |\Psi\rangle_{AB}\langle\Psi| \]
\[ = \frac{1}{8} \sum_{i,j,k,l=0} [ij]_{AB}(kl)a_{ij}^*a_{kl} \]
\[ = \frac{1}{8} \sum_{i,j,k,l=0} [ij]_{AB}(kl)\tilde{\rho}_{ij,kl} \]
\[ = \frac{1}{8} |00\rangle_{AB}(00) + \frac{\sqrt{3}}{8} |00\rangle_{AB}(01) + \frac{\sqrt{3}}{8} |00\rangle_{AB}(10) + \frac{1}{8} |00\rangle_{AB}(11) \]  \hspace{1cm} (14.3.7)
\[ + \frac{\sqrt{3}}{8} |01\rangle_{AB}(00) + \frac{3}{8} |01\rangle_{AB}(01) + \frac{3}{8} |01\rangle_{AB}(10) + \frac{\sqrt{3}}{8} |01\rangle_{AB}(11) \]  \hspace{1cm} (14.3.8)
\[ + \frac{\sqrt{3}}{8} |10\rangle_{AB}(00) + \frac{3}{8} |10\rangle_{AB}(01) + \frac{3}{8} |10\rangle_{AB}(10) + \frac{\sqrt{3}}{8} |10\rangle_{AB}(11) \]  \hspace{1cm} (14.3.9)
\[ + \frac{1}{8} |11\rangle_{AB}(00) + \frac{\sqrt{3}}{8} |11\rangle_{AB}(01) + \frac{\sqrt{3}}{8} |11\rangle_{AB}(10) + \frac{1}{8} |11\rangle_{AB}(11). \]  \hspace{1cm} (14.3.10)

And in the matrix expression, we have

\[ \rho_{AB} = \frac{1}{8} \left( \begin{array}{ccc} 1 & \sqrt{3} & 0 \\ \sqrt{3} & 3 & 0 \\ 0 & 0 & 1 \end{array} \right) \]
\[ = \frac{1}{8} \left( \begin{array}{ccc} 1 & \sqrt{3} & 0 \\ \sqrt{3} & 3 & 0 \\ 0 & 0 & 1 \end{array} \right) \]  \hspace{1cm} (14.3.12)

(a)

\[ \rho_A = \text{tr}_B \rho_{AB} \]
\[ = \sum_{m=0}^{1} B\langle m|\rho_{AB}|m \rangle_{B} \]
\[ = \sum_{m=0}^{1} \sum_{i,j,k,l=0} B\langle m|ij\rangle_{AB}\tilde{\rho}_{ij,kl}(kl|m \rangle_{B} \]
\[ = \sum_{m=0}^{1} \sum_{k=0}^{1} [i]_{A}(k)\tilde{\rho}_{im,km} \]
\[ = \left( \begin{array}{ccc} \tilde{\rho}_{00,00} + \tilde{\rho}_{01,01} & \tilde{\rho}_{00,10} + \tilde{\rho}_{01,11} \\ \tilde{\rho}_{10,00} + \tilde{\rho}_{11,01} & \tilde{\rho}_{10,10} + \tilde{\rho}_{11,11} \end{array} \right) \]
\[ = \frac{1}{8} \left( \begin{array}{ccc} 1 + 3 & \sqrt{3} + \sqrt{3} \\ \sqrt{3} + \sqrt{3} & 3 + 1 \end{array} \right) \]
\[ = \frac{1}{2} \left( \begin{array}{c} 1 \sqrt{3} \\ \sqrt{3} 2 \end{array} \right) \]
\[ = \frac{1}{2} \mathbb{I}_2 + \frac{\sqrt{3}}{4} X, \]  \hspace{1cm} (14.3.13)
where $X$ is the Pauli matrix,

$$X = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$  \hfill (14.3.14)

\[
\begin{align*}
\rho_B &= \text{tr}_A \rho_{AB} \\
&= \sum_{m=0}^1 A(m|\rho_{AB}|m)_A \\
&= \sum_{m=0, i, j, k, l=0}^1 A(m|ij\rangle\langle jk\rangle|k\rangle\langle m\rangle)_A \\
&= \sum_{m=0, i, k=0}^1 |j\rangle_B \langle l|\tilde{\rho}_{mj,kl} \\
&= \begin{pmatrix}
\tilde{\rho}_{00,00} + \tilde{\rho}_{10,10} & \tilde{\rho}_{00,01} + \tilde{\rho}_{10,11} \\
\tilde{\rho}_{01,00} + \tilde{\rho}_{11,10} & \tilde{\rho}_{01,01} + \tilde{\rho}_{11,11}
\end{pmatrix} \\
&= \frac{1}{8} \begin{pmatrix}
1 + 3 & \sqrt{3} + \sqrt{3} \\
\sqrt{3} + \sqrt{3} & 3 + 1
\end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix}
1 & \sqrt{3} \\
\sqrt{3} & 1
\end{pmatrix} \\
&= \frac{1}{4} B^2 + \sqrt{3} X \\
&= \rho_A. \quad \hfill (14.3.15)
\end{align*}
\]

(b) Eigenstate of Pauli gate $Z$.

$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle. \quad \hfill (14.3.16)$

Eigenstate of Pauli gate $X$.

\[
X|\pm\rangle = \pm|\pm\rangle. \quad \hfill (14.3.17)
\]

\[
\begin{align*}
|\pm\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle); \\
\{+\rangle &= \{-\rangle = 1, \quad \{+\rangle \langle -\} = \langle +\rangle = \langle +\rangle = 0; \\
I_2 &= |+\rangle(+\rangle + |-\rangle)(-\rangle).
\end{align*} \hfill (14.3.18)
\]

Find the eigenstate of reduced density matrix $\rho_A$ and $\rho_B$.

\[
\rho_{A, B}|\pm\rangle = \lambda_{\pm}|\pm\rangle, \quad \lambda_{\pm} = \frac{1}{2} \pm \frac{\sqrt{3}}{4}, \quad \hfill (14.3.19)
\]

\[
\rho_{A, B} = \lambda_{+} |+\rangle A, B |+\rangle + \lambda_{-} |-\rangle A, B |-\rangle. \quad \hfill (14.3.20)
\]

Therefore

\[
|\Psi\rangle_{AB} = |+\rangle A (+|\Psi\rangle_{AB} + |-\rangle A (-|\Psi\rangle_{AB}) \\
= |+\rangle A |\tilde{\varphi}_1\rangle B + |-\rangle A |\tilde{\varphi}_2\rangle B. \quad \hfill (14.3.21)
\]

where

\[
|\tilde{\varphi}_1\rangle_B = A (+|\Psi\rangle_{AB} = \frac{\sqrt{6} + \sqrt{2}}{4} |+\rangle_B = \sqrt{\lambda_+} |+\rangle_B, \quad \hfill (14.3.22)
\]

\[
|\tilde{\varphi}_2\rangle_B = A (-|\Psi\rangle_{AB} = -\frac{\sqrt{6} - \sqrt{2}}{4} |-\rangle_B = -\sqrt{\lambda_-} |-\rangle_B.
\]

Thus, the Schmidt decomposition of state $|\Psi\rangle_{AB}$ shows as

\[
|\Psi\rangle_{AB} = \sqrt{\lambda_+} |+\rangle A |+\rangle_B - \sqrt{\lambda_-} |-\rangle A |-\rangle_B. \quad \hfill (14.3.23)
\]

139
14.4 The purification theorem and GHJW theorem

14.4.1 Purification

Thm 14.4.1.1 (Purification). For any given mixed state $\rho_A$ for system $A$, there is a bipartite pure state $|\Phi\rangle_{AB}$ such that

$$\rho_A = \text{tr}_B(|\Phi\rangle_{AB}\langle \Phi|).$$

(14.4.1)

Then $|\Phi\rangle_{AB}$ is called the purification of $\rho_A$.

For example, the density matrix $\Phi$ as defined in E.Q. (14.4.2). As we can see that

$$\rho_A = I_A/2,$$

as we have shown in E.Q. (13.4.5), has the purification

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}).$$

Proof. The key point is the Schmidt decomposition. Suppose that

$$\rho_A := \sum_i p_i |\varphi_i\rangle_A \langle \varphi_i|,$$

with $\sum_i p_i = 1$, and $p_i > 0$, (14.4.2)

where $\{|\varphi_i\rangle\}$ is an orthonormal basis of $H_A$.

Now, we introduce another system $B$ and construct the following state for the composite physical system $A \otimes B$, $\Phi_{AB}$:

$$|\Phi\rangle_{AB} = \sum_i \sqrt{p_i} |\varphi_i\rangle_A \otimes |\alpha_i\rangle_B,$$

with $\sum_i p_i = 1$, (14.4.3)

amid, $\{|\alpha_i\rangle\}$ is the orthonormal basis of $H_B$. And we can show that $|\Phi\rangle_{AB}$ is a pure state, which is self-convincing, and it is also normalized since

$$\text{tr}_A|\Phi\rangle_{AB} \langle \Phi|_{AB} = \sum_{i,j} \sqrt{p_j p_i} \langle \varphi_j|_A \langle \delta_{j,i}|_B \langle \varphi_i| \otimes \langle \alpha_j|_A \langle \alpha_i|_B$$

$$= \sum_{i,j} \sqrt{p_j p_i} \langle \varphi_j|_A \langle \varphi_i|_B \langle \alpha_j|_A \langle \alpha_i|_B$$

$$= \sum_{i,j} \sqrt{p_j p_i} \delta_{ij} \delta_{ji}$$

$$= \sum_i p_i$$

$$= 1.$$ 

We can also verify that $|\Phi\rangle_{AB}$ defined in E.Q. (14.4.3) is truly purification of the state $\rho_A$ as defined in E.Q. (14.4.2). As we can see that

$$\text{tr}_B\rho_{AB} = \sum_i |\alpha_i\rangle_B \langle \alpha_i|_B \langle \rho_{AB}|_B$$

$$= \sum_{i,j,k} \sqrt{p_j p_k} \langle \varphi_j|_A \langle \varphi_k|_B \langle \alpha_j|_A \langle \alpha_k|_B$$

$$= \sum_{i,j,k} \sqrt{p_j p_k} \langle \varphi_j|_A \langle \varphi_k|_B \langle \alpha_k|_B \langle \alpha_j|_A$$

$$= \sum_{i,j,k} \sqrt{p_j p_k} \delta_{ij} \delta_{ki}$$

$$= \sum_i p_i \langle \varphi_i|_A \langle \varphi_i|_B$$

$$= \rho_A.$$
i.e., \[ \rho_A = \text{tr}_B \rho_{AB}. \] (14.4.4)

Therefore, \( |\Phi\rangle_{AB} \) is a purification of \( \rho_A \).

There are some things that we should remark:

- \( \rho_A \) can be prepared as ensembles of pure states in many different ways. This can be inferred directly from the theorem of purification, since based on any state ensemble interpretation of the \( \rho_A \), we can always construct the corresponding purification, which can be different from each other.

- All these purification are experimentally indistinguishable, if one only observes the system \( A \). Because for any measurement \( M_A \) on the subsystem \( A \), we have

\[ \langle M_A \rangle = \text{tr}_A \left( M_A \rho_A \right). \]

- Different purifications are associated with a local unitary transformation \( U_B \) on the subsystem \( B \), which we will prove in the next subsection and introduce the GHJW theorem.

### 14.4.2 The GHJW theorem

**Thm 14.4.2.1 (GHJW).** Consider many ensembles that realising \( \rho_A \). There is a Hilbert space \( \mathcal{H}_B \) and a bipartite pure state \( |\Phi\rangle_{AB} \), by which any one of these ensembles can be realized measuring a suitable observable of \( B \).

**Proof.** Firstly, we would show that two arbitrary purification of \( \rho_A \), e.g. \( |\Phi_1\rangle \) and \( |\Phi_2\rangle \), are connected with a local unitary transformation on the subsystem \( B \). Let’s denote that

\[
\begin{align*}
|\Phi_1\rangle_A & := \sum_i \sqrt{p_i} |\varphi_i\rangle_A \otimes |\alpha_i\rangle_B, \\
|\Phi_2\rangle_A & := \sum_i \sqrt{q_i} |\psi_i\rangle_A \otimes |\beta_i\rangle_B,
\end{align*}
\]

(14.4.5)

where \( \{|p_i, |\varphi_i\rangle_A\} \) and \( \{|q_i, |\psi_i\rangle_A\} \) are two state ensemble interpretations of \( \rho_A \) and \( \{|\alpha_i\rangle_B\} \) and \( \{|\beta_i\rangle_B\} \) are two orthonormal basis for \( \mathcal{H}_B \).

Assume that \( \rho_A \) has the following eigenvectors and the associated eigenvalues

\[ \rho_A |k\rangle_A = \lambda_k |k\rangle_A, \quad \text{with} \quad \sum_k \lambda_k = 1, \quad \text{and} \quad \lambda_k \geq 0. \]

(14.4.6)

Since \( |\Phi_1\rangle_{AB} \) and \( |\Phi_2\rangle_{AB} \) are two different purification of \( \rho_A \), it’s apparent that

\[ \text{tr}_A |\Phi_1\rangle_{AB} = \text{tr}_A |\Phi_2\rangle_{AB} = \rho_A. \]

(14.4.7)

With E.Q. (14.4.7) in mind, we can see that the respective Schmidt decompositions for \( |\Phi_1\rangle_{AB} \) and \( |\Phi_2\rangle_{AB} \) are

\[
\begin{align*}
|\Phi_1\rangle_{AB} & = \sum_k \sqrt{\lambda_k} |k\rangle \otimes |\varphi_k\rangle_B, \\
|\Phi_2\rangle_{AB} & = \sum_k \sqrt{\lambda_k} |k\rangle \otimes |\psi_k\rangle_B.
\end{align*}
\]

(14.4.8)

Therefore, \( |\Phi_1\rangle_{AB} \) and \( |\Phi_2\rangle_{AB} \) are connected with the local unitary translation

\[ |\Phi_1\rangle_{AB} = (I_A \otimes U_B) |\Phi_2\rangle_{AB}, \]

(14.4.9)
with
\[ U_B = \sum_k |k'_1\rangle_B \langle k'_2| . \] (14.4.10)

Substitute E.Q. \([14.4.5]\) into E.Q. \((14.4.9)\), and we can get that
\[ |\Phi_1\rangle_{AB} = \sum_i \sqrt{q_i} |\psi_i\rangle_A \otimes U_B |\beta_i\rangle_B , \]

namely
\[ |\Phi_1\rangle_{AB} = \sum_i \sqrt{q_i} |\psi_i\rangle_A \otimes |\gamma_i\rangle_B , \] (14.4.11)

where
\[ |\gamma_i\rangle_B := U_B |\beta_i\rangle_B . \] (14.4.12)

Surely, \(\{|\gamma_i\rangle_B\}\) makes an orthonormal basis for \(\mathcal{H}_B\), too.

Secondly, let’s consider two observables \(\mathcal{M}_B\) and \(\mathcal{N}_B\) defined in the subsystem B. And we consider that case the two vector basis \(\{|\alpha_i\rangle_B\}\) and \(\{|\gamma_i\rangle_B\}\) respectively are eigenvectors of \(\mathcal{M}_B\) and \(\mathcal{N}_B\), namely
\[
\begin{align*}
\mathcal{M}_B |\alpha_i\rangle_B &= M_i |\alpha_i\rangle_B , \\
\mathcal{N}_B |\gamma_j\rangle_B &= N_j |\gamma_j\rangle_B , \quad \text{with } i,j = 1,\ldots, \dim \mathcal{H}_B .
\end{align*}
\]

Therefore, we can conclude that:

- After the measurement of \(\mathcal{M}_B\), there would be a probability of \(p_i\) for the subsystem B in the state \(|\alpha_i\rangle_B\), namely the subsystem A in the state \(|\varphi_i\rangle_A\). Hence, measuring \(\mathcal{M}_B\) in the subsystem B would prepare the subsystem A in the mixed state \(\rho_A\) with the state ensemble \(\{(p_i, |\varphi_i\rangle_A)\}\).

- After the measurement of \(\mathcal{N}_B\), there would be a probability of \(q_j\) for the subsystem B in the state \(|\gamma_j\rangle_B\), namely the subsystem A in the state \(|\psi_i\rangle_A\). Thus, the measurement of \(\mathcal{N}_B\) in the subsystem B would prepare the subsystem A in the mixed state \(\rho_A\) with the state ensemble \(\{(p_i, |\psi_i\rangle_A)\}\).

Therefore, we get the GHJW theorem.

\[ \square \]

### 14.5 Information is physics

Ambiguities in the concept of density matrix

(a) A density matrix has many different interpretations of state ensemble.

(b) A density matrix can have many different interpretations of purification.

Here we discuss an example of bipartite system which is made up of two subsystems Alice and Bob. Alice and Bob share a pure state of the form
\[ |\psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left( |00\rangle_{AB} + |11\rangle_{AB} \right) . \] (14.5.1)
We can easily get the reduced density matrices for Alice and Bob respectively,

\[ \rho_A = \text{tr}_B \rho_{AB} \]
\[ = \frac{1}{2} \sum_{i,j,j'} B \langle i | \rho_{AB} | j \rangle_B (A \langle j' | \otimes_B \langle j' | \rangle_B) \]
\[ = \frac{1}{2} \sum_{i,j,j'} B \langle i | \otimes_B \langle j | \rangle_B (A \langle j' | \otimes_B \langle j' | \rangle_B) \]
\[ = \frac{1}{2} \sum_{i,j,j'} \delta_{ij} \delta_{jj'} \]
\[ = \frac{1}{2} \sum_{i=0}^1 \langle i | \rangle_A \delta_{ii}, \]

i.e.,

\[ \rho_A = \frac{1}{2} I_2. \] (14.5.2)

Reasoning in the same manner, we can get

\[ \rho_B = \frac{1}{2} I_2. \] (14.5.3)

Now, let’s consider the following processes:

- **The first process:**
  
  - Bob performs a measurement along the \( z \)-direction on his system but doesn’t tell Alice via classical communication (e.g. phone call). Therefore, no information transfer between Alice and Bob. In this case, we can get \( \rho_A = \frac{1}{2} I_2 \), which is a mixed state.
  
  - Bob performs a measurement along the \( z \)-direction on his system and phone calls Alice about his measurement result (e.g. \( |0\rangle_B \)). There is information transferred between Alice and Bob. And, in this case \( \rho_A = |0\rangle_A \langle 0| \), which is a pure state.

From this process, we can conclude that information (via classical communication) changes the physics of the system of Alice.

- **The second process:**
  
  - Bob measures his qubit along the \( z \)-axis, and phone calls Alice but only tells her that he has measured his qubit along the \( z \)-direction. Therefore, Alice’s system is prepared in state ensemble

\[ \left\{ \left( \frac{1}{2} |0\rangle + \frac{1}{2} |1\rangle \right) \right\}. \] (14.5.4)

  - Bob measures his qubit along the \( x \)-axis, and phone calls Alice but only tells her that he has measured his qubit along the \( x \)-direction. Therefore, Alice’s system is prepared in state ensemble

\[ \left\{ \left( \frac{1}{2} |+\rangle - \frac{1}{2} |-\rangle \right) \right\}. \] (14.5.5)
The two state ensembles described by (14.5.4) and (14.5.5) respectively, are not distinguished by any conceivable measurement on the subsystem A only. The reason is that both these two state ensemble are corresponding to the same density operator $\rho_A = \frac{1}{2}I_2$. 
Chapter 15

Mixed State Entanglement and Multi-partite Entanglement

... entanglement is a fundamentally new resource in the world that goes essentially beyond classical resources; iron to the classical world’s bronze age.

—Nielsen & Chuang

A major task of quantum computation and quantum information is to exploit this new resource (quantum entanglement) to do information processing tasks impossibility or much more difficult with classical information.

—Nielsen & Chuang

A lot of mileage in quantum computation, and especially quantum information, has come from asking the simple question: “what would some entanglement buy me in this problem?”.

—Nielsen & Chuang

References:

- [Preskill] Chapter 4: Quantum entanglement;
- [Nielsen & Chuang] Chapter 12: Quantum information theory.

15.1 Bipartite mixed state entanglement

15.1.1 Separability

15.1.1.1 Bipartite pure states

A bipartite pure state is separable if it is a tensor product of two pure states, i.e.,

\[ |\psi\rangle_{AB} = |\alpha\rangle_A \otimes |\beta\rangle_B, \]

(15.1.1)

which is equivalent to say that the bipartite pure state has the Schmidt number as 1 is separable.
15.1.1.2 Bipartite mixed state

A bipartite mixed state $\rho_{AB}$ is separable if it admits the ensemble description given by

$$\rho_{AB} = \sum_{i,j} p_{ij} \rho_{A,i} \otimes \rho_{B,j},$$  \hspace{1cm} (15.1.2)

where

$$\sum_{i,j} p_{ij} = 1, \quad p_{ij} \geq 0$$

and $\{\rho_{A,i} \mid \forall \ i\}$ are density operators for subsystem A, while $\{\rho_{B,j} \mid \forall \ j\}$ for subsystem B.

**Example:** This is a separable mixed state

$$\rho_{AB} = \sum_{i} p_{i} |\alpha_{i}\rangle \langle \alpha_{i}| \otimes \sum_{j} q_{j} |\beta_{j}\rangle \langle \beta_{j}|$$  \hspace{1cm} (15.1.3)

where $p_{ij} = p_{i} q_{j}$

$$\sum_{i} p_{i} = 1, \quad \text{with} \quad p_{i} \geq 0, \quad \forall i,$$

and

$$\sum_{j} q_{j} = 1, \quad \text{with} \quad q_{j} \geq 0, \quad \forall j.$$

15.1.2 Quantum bipartite entanglement

A bipartite state is entangled if it is not a separable state. Here is one separable state

$$\rho_{1} = \frac{1}{4} I_{2} \otimes I_{2}.$$  \hspace{1cm} (15.1.4)

And another separable state shows as

$$\rho_{2} = \frac{1}{4} \sum_{i=0}^{1} |i\rangle_{A} \langle i| \otimes \sum_{j=0}^{1} |j\rangle_{B} \langle j|$$

$$= \frac{1}{4} \sum_{i,j=0}^{1} |ij\rangle_{AB} \langle ij|$$

$$= \frac{1}{4} \left( |00\rangle_{AB} \langle 00| + |01\rangle_{AB} \langle 01| + |10\rangle_{AB} \langle 10| + |11\rangle_{AB} \langle 11| \right).$$  \hspace{1cm} (15.1.5)

And for the state $\rho_{3}$, is expressed as

$$\rho_{3} = \frac{1}{4} \left( |\phi^{+}\rangle \langle \phi^{+}| + |\phi^{-}\rangle \langle \phi^{-}| + |\psi^{+}\rangle \langle \psi^{+}| + |\psi^{-}\rangle \langle \psi^{-}| \right),$$  \hspace{1cm} (15.1.6)

which is hard to determine whether it is separable or entangled.

However, as we shall see that $\rho_{1}$, $\rho_{2}$ and $\rho_{3}$ are actually same, since both four product states $\{|ij\rangle \mid i,j = 0,1\}$ and four Bell states $\{|\psi(ij)\rangle \mid i,j = 0,1\}$ can span the Hilbert space $\mathcal{H}_{2} \otimes \mathcal{H}_{2}$, namely

$$\rho_{1} = \rho_{2} = \rho_{3}.$$  \hspace{1cm} (15.1.7)

**Remarks:** Quantum entanglement is a very important but difficult issue in quantum physics. How to quantify multi-particle entanglement is an open problem up to now.
15.1.3 Positive-partial transpose (PPT) criterion for quantum bipartite separability

**Thm 15.1.3.1.** If $\rho_{AB}$ is separable, namely be of the form (15.1.2), then the partial transpose of $\rho_{AB}$, defined as

$$
(\rho_{AB})^{PT} = (I_2 \otimes T)\rho_{AB} = \sum_{i,j} p_{ij} \rho_{A,i} \otimes (\rho_{B,j})^T,
$$

(15.1.8)
is still non-negative, where $T$ denotes the transpose of matrix, namely

$$
T(|i\rangle \langle j|) = |j\rangle \langle i|.
$$

(15.1.9)

*Proof.* As we shall know that $\{\rho_{B,j} \mid \forall j\}$ are density operators of the subsystem B. Therefore, for an arbitrary $\rho_{B,j}$, the eigenvectors can span the local Hilbert space $\mathcal{H}_B$. If we can diagonalize $\rho_{B,j}$ with the basis being the complete set of independent orthonormal eigenvectors of $\rho_{B,j}$. After $\rho_{B,j}$ being diagonalized, we can get that its transpose is diagonalized, too. It would obvious that, that the transpose would preserve the eigenvalues of $\rho_{B,j}$. In another worlds, $(\rho_{B,j})^T$ should also be nonnegative, namely

$$
(\rho_{B,j})^T \geq 0, \; \forall j.
$$

(15.1.10)

On the other hand, the fact that $\{\rho_{A,i} \mid \forall i\}$ being the density operators of the subsystem A, ensures that they are all nonnegative. Therefore, $(\rho_{AB})^{PT}$ is nonnegative. \hfill \Box

**Corollary 15.1.3.1.** If $(\rho_{AB})^{PT}$ is not positive, then $\rho_{AB}$ is an entangled state.

This is a direct corollary of the theorem [15.1.3.1]

15.1.4 Example: the Werner state and the PPT criterion

Werner State, firstly presented in 1989 by Reinhard F. Werner, is defined as

$$
\rho(\lambda) := \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+|
$$

(15.1.11)

with $\lambda$ being a real number. We can show that if $\rho(\lambda)$ with definition [15.1.11] is a density matrix, iff

- Unit trace.

$$
\text{tr} [\rho(\lambda)] = \text{tr} \left[ \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+| \right]
$$

$$
= \frac{1}{4} (1 - \lambda) \text{tr} (I_4) + \lambda \text{tr} (|\phi^+\rangle \langle \phi^+|)
$$

$$
= (1 - \lambda) + \lambda
$$

$$
= 1.
$$
• Self-adjoint.

\[
[\rho(\lambda)]^\dagger = \left[ \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+| \right]^\dagger \\
= \left[ \frac{1}{4} (1 - \lambda) I_4 \right]^\dagger + (\lambda |\phi^+\rangle \langle \phi^+|)^\dagger \\
= \frac{1}{4} (1 - \lambda)^* I_4^\dagger + \lambda^* (|\phi^+\rangle \langle \phi^+|)^\dagger \\
= \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+| \\
= \rho(\lambda),
\]

since \(\lambda \in \mathbb{R}\).

• Nonnegative.

\[
\rho(\lambda) \geq 0 \iff \text{the eigenvalues of } \rho(\lambda) \geq 0.
\]

And it would be easy to verify that the Bell’s states \(|\psi(i, j)\rangle \text{ for } i, j = 1, 2\) are four eigenstates of \(\rho(\lambda)\).

\[
\rho(\lambda) |\psi(0, 0)\rangle = \rho(\lambda) |\phi^+\rangle \\
= \left[ \frac{1}{4} (1 - \lambda) I_4 + \lambda |\psi^-\rangle \langle \psi^-| \right]|\phi^+\rangle \\
= \frac{1}{4} (4 + 3\lambda) |\phi^+\rangle.
\]

(15.1.12)

and with \(i \neq 0\) and \(j \neq 0\), we can get

\[
\rho(\lambda) |\psi(i, j)\rangle = \rho(\lambda) |\psi(i, j)\rangle \\
= \left( \frac{1}{4} (1 - \lambda) I_4 + \lambda |\psi(0, 0)\rangle \langle \psi(0, 0)| \right)|\psi(i, j)\rangle \\
= \frac{1}{4} (1 - \lambda) |\psi(i, j)\rangle,
\]

(15.1.13)

as we know that \(|\psi(i, j)\rangle\), with \((i, j = 0, 1)\), are mutually independent. Therefore, we’ve gotten the complete set of mutually independent eigenvectors of \(\rho(\lambda)\).

Therefore, for the constrain of nonnegative:

\[
\rho(\lambda) \geq 0 \iff \begin{cases} 
\frac{1}{4} (1 + 3\lambda) \geq 0 \\
\frac{1}{4} (1 - \lambda) \geq 0
\end{cases} \iff -\frac{1}{3} \leq \lambda \leq 1.
\]

(15.1.14)

When \(\lambda = -\frac{1}{3}\)

\[
\rho\left(-\frac{1}{3}\right) = \frac{1}{3} (I_4 - |\phi^+\rangle \langle \phi^+|) = \frac{1}{3} (|\phi^-\rangle \langle \phi^-| + |\psi^+\rangle \langle \psi^+| + |\psi^-\rangle \langle \psi^-|),
\]

therefore

\[
\rho\left(-\frac{1}{3}\right) |\phi^+\rangle = \frac{1}{3} (I_4 - |\phi^+\rangle \langle \phi^+|)|\phi^+\rangle = 0|\phi^+\rangle,
\]

(15.1.15)
and

\[ \rho \left( \frac{1}{3} \right) |\psi(i, j)\rangle = \frac{1}{3} (I_4 - |\psi(0, 0)\rangle \langle \psi(0, 0)|) |\psi(i, j)\rangle = \rho \left( \frac{1}{3} \right) |\psi(i, j)\rangle \]  \hspace{1cm} (15.1.16)

with \( i \neq 0 \) and \( j \neq 0 \).

When \( \lambda = 1 \)

\[ \rho(1) = |\phi^+\rangle \langle \phi^+|, \]

which is a pure state with the relations

\[ \rho(1) |\phi^+\rangle = (|\phi^+\rangle \langle \phi^+|)|\phi^+\rangle = |\phi^+\rangle. \]  \hspace{1cm} (15.1.17)

and

\[ \rho(1) |\psi(i, j)\rangle = (|\phi^+\rangle \langle \phi^+|)|\psi(i, j)\rangle = 0 |\psi(i, j)\rangle \]  \hspace{1cm} (15.1.18)

with \( i \neq 0 \) and \( j \neq 0 \).

Calculate the partial transpose of Werner state \( \rho(\lambda) \).

\[ [\rho(\lambda)]^{PT} = (I_2 \otimes T) \rho(\lambda) \]

\[ = (I_2 \otimes T) \left[ \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+| \right] \]

\[ = \frac{1}{4} (1 - \lambda) (I_2 \otimes T) I_4 + \frac{1}{2} \lambda (I_2 \otimes T) |\phi^+\rangle \langle \phi^+| \]

\[ = \frac{1}{4} (1 - \lambda) I_4 + \frac{1}{2} \lambda (I_2 \otimes T) (|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|) \]

\[ = \frac{1}{4} (1 - \lambda) I_4 + \frac{1}{2} \lambda (|00\rangle \langle 00| + |01\rangle \langle 10| + |10\rangle \langle 01| + |11\rangle \langle 11|), \]  \hspace{1cm} (15.1.19)

from which we can see that

\[ [\rho(\lambda)]^{PT} |\phi^+\rangle = [\rho(\lambda)]^{PT} \frac{|00\rangle + |11\rangle}{\sqrt{2}} \]

\[ = \left[ \frac{1}{4} (1 - \lambda) + \frac{1}{2} \lambda \right] |\phi^+\rangle \]

\[ = \frac{1}{4} (1 + \lambda) |\phi^+\rangle, \]  \hspace{1cm} (15.1.20)

\[ [\rho(\lambda)]^{PT} |\phi^−\rangle = [\rho(\lambda)]^{PT} \frac{|00\rangle - |11\rangle}{\sqrt{2}} \]

\[ = \left[ \frac{1}{4} (1 - \lambda) + \frac{1}{2} \lambda \right] |\phi−\rangle \]

\[ = \frac{1}{4} (1 + \lambda) |\phi−\rangle, \]  \hspace{1cm} (15.1.21)

\[ [\rho(\lambda)]^{PT} |\psi^+\rangle = [\rho(\lambda)]^{PT} \frac{|01\rangle + |10\rangle}{\sqrt{2}} \]

\[ = \left[ \frac{1}{4} (1 - \lambda) + \frac{1}{2} \lambda \right] |\psi^+\rangle \]

\[ = \frac{1}{4} (1 + \lambda) |\psi^+\rangle, \]  \hspace{1cm} (15.1.22)

\[ [\rho(\lambda)]^{PT} |\psi^−\rangle = [\rho(\lambda)]^{PT} \frac{|01\rangle - |10\rangle}{\sqrt{2}} \]

\[ = \left[ \frac{1}{4} (1 - \lambda) - \frac{1}{2} \lambda \right] |\psi^−\rangle \]

\[ = \frac{1}{4} (1 - 3\lambda) |\psi^−\rangle. \]  \hspace{1cm} (15.1.23)
In another viewpoint, we can see

\[
(I_2 \otimes T)|\phi^+\rangle\langle \phi^+| = (I_2 \otimes T)\frac{1}{2} \sum_{i,j=0}^1 |ij\rangle\langle jj|
\]

\[
= \frac{1}{2} \sum_{i,j=0}^1 (I_2 \otimes T)(|ij\rangle \otimes |ij\rangle)
\]

\[
= \frac{1}{2} \sum_{i,j=0}^1 |ij\rangle \otimes T(|ij\rangle)
\]

\[
= \frac{1}{2} \sum_{i,j=0}^1 |ij\rangle \otimes |j\rangle
\]

\[
= \frac{1}{2} \sum_{i,j=0}^1 |ij\rangle
\]

\[
= \frac{1}{2} \text{SWAP}
\]  

(15.1.24)

where the Swap gate is defined in (7.2.4) and has the matrix expression

\[
\text{SWAP} = \sum_{i,j=0}^1 |ij\rangle\langle jj| = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]  

(15.1.25)

Note that the SWAP gate is identical with its inverse, namely

\[
\text{SWAP} \circ \text{SWAP} = I_4,
\]  

(15.1.26)

which implies the characteristic equation

\[
\text{SWAP}^2 - I_4 = 0.
\]  

(15.1.27)

Thus, we can see the eigenvalue of the SWAP gate is \( \pm 1 \). In two electrons system, i.e., two spin-\( \frac{1}{2} \) system, triplet states are eigenstate of SWAP gate with eigenvalue 1, expressed as

\[
\text{SWAP} |\phi^+\rangle = |\phi^+\rangle, \\
|\phi^+\rangle = |\phi^+\rangle,
\]  

(15.1.28)

For singlet state, it is the eigenstate of SWAP gate with eigenvalue -1,

\[
\text{SWAP} |\psi^-\rangle = -|\psi^-\rangle.
\]  

(15.1.29)

Therefore, rewrite the SWAP gate as

\[
\text{SWAP} = |\phi^+\rangle \langle \phi^+| + |\phi^-\rangle \langle \phi^-| + |\psi^+\rangle \langle \psi^+| - |\psi^-\rangle \langle \psi^-| = I_4 - |\psi^-\rangle \langle \psi^-|.
\]  

(15.1.30)

And the partial transpose of Werner state \( \rho(\lambda) \) can be calculated as

\[
[\rho(\lambda)]^{\text{PT}} = (I_2 \otimes T)\rho(\lambda)
\]

\[
= (I_2 \otimes T)\left[\frac{1}{4}(1 - \lambda)I_4 + \lambda|\phi^+\rangle \langle \phi^+|\right]
\]

\[
= \frac{1}{4} (1 - \lambda)I_4 + \frac{\lambda}{2} \text{SWAP}
\]

\[
= \frac{1}{4} (1 - \lambda)I_4 + \frac{1}{2} \lambda I_4 - \lambda |\psi^-\rangle \langle \psi^-|
\]

\[
= \frac{1}{4} (1 + \lambda)I_4 - \lambda |\psi^-\rangle \langle \psi^-|.
\]  

(15.1.31)
And the eigenvalues can be obtained by the way
\[
\begin{align*}
\left[ \rho(\lambda) \right]^{\mathrm{PT}} |\psi^-\rangle &= \frac{1}{4} (1 + \lambda - \lambda) |\psi^-\rangle = \frac{1 - 3\lambda}{4} |\psi^-\rangle, \\
\left[ \rho(\lambda) \right]^{\mathrm{PT}} |\phi^+\rangle &= \frac{1}{4} (1 + \lambda) \begin{pmatrix} |\phi^+\rangle \\ |\phi^-\rangle \\ |\psi^-\rangle \\ |\psi^+\rangle \end{pmatrix}.
\end{align*}
\]
(15.1.32)
(15.1.33)

Therefore, \( \rho(\lambda) \) has the eigenvalues \( \left\{ \frac{1}{4} (1 + \lambda), \frac{1}{4} (1 - 3\lambda) \right\} \), and the eigenstates
\[
\left\{ \begin{array}{c}
|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, \\
\frac{1}{4} (1 + \lambda), \text{triplet state}
\end{array} \right\}, \quad \left\{ \begin{array}{c}
|\psi^-\rangle, \\
\frac{1}{4} (1 - 3\lambda), \text{singlet state}
\end{array} \right\}.
\]
(15.1.34)

For the partial transpose criterion, we have the constrain for the parameter \( \lambda \)
\[
\left[ \rho(\lambda) \right]^{\mathrm{PT}} \geq 0 \iff \left\{ \begin{array}{l}
\frac{1}{4} (1 + \lambda) \geq 0 \iff -1 \leq \lambda \leq \frac{1}{3}, \\
\frac{1}{4} (1 - 3\lambda) \geq 0 \iff \frac{1}{3} \leq \lambda \leq 1.
\end{array} \right.
\]
(15.1.35)

From inequality (15.1.14) and inequality (15.1.35), we can infer that
\[
\rho(\lambda) \geq 0 \iff -\frac{1}{3} \leq \lambda \leq 1,
\]
\[
\left[ \rho(\lambda) \right]^{\mathrm{PT}} \geq 0 \iff -1 \leq \lambda \leq \frac{1}{3}.
\]

That means
\[
\rho(\lambda) \text{ is separable } \iff -\frac{1}{3} \leq \lambda \leq \frac{1}{3},
\]
\[
\rho(\lambda) \text{ is entangled } \iff \frac{1}{3} < \lambda \leq 1.
\]

15.1.5 Example: the Werner state and the CHSH inequality

Now let’s consider the CHSH inequality for the bipartite Werner states. We may set the four observables to be \( A = \sigma_A \cdot \hat{a}_1 \) and \( B = \sigma_A \cdot \hat{a}_2 \) in subsystem Alice, and \( C = \sigma_B \cdot \hat{b}_1 \) and \( D = \sigma_B \cdot \hat{b}_2 \) in subsystem Bob.

Firstly, we can calculate
\[
\rho(\lambda)^{\mathrm{PT}} = \text{tr}_A \left[ (\sigma_A \cdot \hat{a}) (\sigma_B \cdot \hat{b}) \rho(\lambda) \right] = \text{tr}_A \left[ (\sigma_A \cdot \hat{a}) (\sigma_B \cdot \hat{b}) \left( \frac{1}{4} (1 - \lambda) |\phi^+\rangle \langle \phi^+| \right) \right] = \frac{1 - \lambda}{4} \text{tr}_A (\sigma_A \cdot \hat{a}) \text{tr}_B (\sigma_B \cdot \hat{b}) + \lambda \langle \phi^+ | (\sigma_A \cdot \hat{a}) (\sigma_B \cdot \hat{b}) |\phi^+ \rangle = \lambda \langle \phi^+ | (\sigma_A \cdot \hat{a}) (\sigma_B \cdot \hat{b}) |\phi^+ \rangle = \lambda \hat{a} \cdot \hat{b}.
\]
(15.1.36)

Therefore, we can get
\[
| \langle AC \rangle + \langle BC \rangle + \langle AD \rangle - \langle BD \rangle | = | \lambda | \left| \hat{a}_1 \cdot \hat{b}_1 + \hat{a}_2 \cdot \hat{b}_1 + \hat{a}_1 \cdot \hat{b}_2 - \hat{a}_2 \cdot \hat{b}_2 \right| = | \lambda | \left| \langle \hat{a}_1 + \hat{a}_2 \rangle \cdot \hat{b}_1 + \langle \hat{a}_1 - \hat{a}_2 \rangle \cdot \hat{b}_2 \right|.
\]
(15.1.37)
Since \((\bar{a}'_1 + \bar{a}'_2)\) and \((\bar{a}'_1 - \bar{a}'_2)\) are mutually orthogonal, therefore
\[
0 \leq |(\bar{a}'_1 + \bar{a}'_2)\hat{b}_1 + (\bar{a}'_1 - \bar{a}'_2)\hat{b}_2| \leq 2\sqrt{2}. 
\] (15.1.38)

To ensure the CHSH inequality
\[
|\langle AC \rangle + \langle BC \rangle + \langle AD \rangle - \langle BD \rangle| \leq 2 
\] (15.1.39)
being correct, we only need
\[
|\langle AC \rangle + \langle BC \rangle + \langle AD \rangle - \langle BD \rangle|_{\text{max}} \leq 2, 
\] (15.1.40)
i.e.,
\[
2\sqrt{2}|\lambda| \leq 2, \quad \iff \quad -\frac{1}{\sqrt{2}} \leq \lambda \leq \frac{1}{\sqrt{2}}. 
\] (15.1.41)

Combine above constrain of parameter \(\lambda\) with the inequality \((15.1.14)\) which ensures \(\rho(\lambda)\) nonnegative namely a density operator, and we can get
\[
-\frac{1}{3} \leq \lambda \leq \frac{1}{\sqrt{2}}. 
\] (15.1.42)

As we shall see that with \(\lambda \in [-1/3, 1]\), the operator \(\rho(\lambda)\) is ensured to be a possible density operator, and

- if \(\lambda \in (1/3, 1]\), the Werner state \(\rho(\lambda)\) is entangled;
- if \(\lambda \in [-1/3, 1/\sqrt{2}]\), the Werner state \(\rho(\lambda)\) obeys the CHSH inequality;
- if \(\lambda \in [1/3, 1/\sqrt{2}]\), the Werner state \(\rho(\lambda)\) is entangled but compatible with the CHSH inequality.

Since the CHSH inequality is actually a kind of Bell’s inequality, which is the criterion of local hidden variable theorem, we can conclude that entanglement is not always contradicted with the local hidden variable theorem, namely the entanglement can exist in the local hidden variable theory in some case.

### 15.2 Multi-partite entanglement

#### 15.2.1 Definition

**Def 15.2.1.** We define the multi-partite separability and entanglement in the following.

1° The \(n\)-partite state is fully separable iff

- **Pure state:**
  \[
  |\psi\rangle_{A_1A_2...A_n} = |\psi\rangle_{A_1} \otimes |\psi\rangle_{A_2} \otimes ... \otimes |\psi\rangle_{A_n}; 
  \] (15.2.1)

- **Mixed state:**
  \[
  \rho_{A_1A_2...A_n} = \sum_{i=1}^{n} p_i \rho_{A_1}^i \otimes \rho_{A_2}^i \otimes ... \otimes \rho_{A_n}^i. 
  \] (15.2.2)

2° The \(n\)-partite state is fully entangled iff it is not fully separable.

3° The \(n\)-partite state is genuinely entangled iff all bipartite partitions are entangled.

**Example:**
1° A fully separable tripartite state is shown in Figure 15.1 where the dashed lines represent separable states and $A_i$ denotes the $i$-th particle, with $i = 1, 2, 3$.

2° As shown in Figure 15.2, there are three cases of the fully entangled tripartite state, where the solid lines represent the entangled states.

3° See Figure 15.3 for diagrammatical representation of genuinely entangled tripartite state.

15.2.2 The GHZ state

Def 15.2.2 ($n$-qubit GHZ states). The state defined as

$$|\text{GHZ}_x\rangle_n = \frac{1}{\sqrt{2^k}} (|x\rangle \pm |\bar{x}\rangle)$$

(15.2.3)

is called $n$-qubit GHZ state, where $x$ is an $n$-bit string of 0 and 1 with $x + \bar{x} = \left(\frac{1}{n}\right)_{\text{binary}}$.

From this definition we can see the examples as

1° $n = 1$:

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle \pm |1\rangle),$$

are eigenstates of $X$ matrix with eigenvalues $\pm 1$, and form an orthogonal basis of Hilbert space $H_2$. 

153
2° $n = 2$:
\[
\begin{align*}
|\psi(0, 0)\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \\
|\psi(0, 1)\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle), \\
|\psi(1, 0)\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle), \\
|\psi(1, 1)\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)
\end{align*}
\]

≡ Bell states \{\ket{\phi^\pm}, \ket{\psi^\mp}\},

are eigenvectors of the parity-bit operator \(Z \otimes Z\) and the phase-bit operator \(X \otimes X\),

and form an orthogonal basis of Hilbert space \(\mathcal{H}_2 \otimes \mathcal{H}_2\).

3° $n = 3$:
\[
\begin{align*}
|\Phi_1\rangle &= \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle), \\
|\Phi_2\rangle &= \frac{1}{\sqrt{2}} (|001\rangle + |110\rangle), \\
|\Phi_3\rangle &= \frac{1}{\sqrt{2}} (|010\rangle + |101\rangle), \\
|\Phi_4\rangle &= \frac{1}{\sqrt{2}} (|011\rangle + |100\rangle), \\
|\Phi_5\rangle &= \frac{1}{\sqrt{2}} (|111\rangle - |100\rangle), \\
|\Phi_6\rangle &= \frac{1}{\sqrt{2}} (|101\rangle - |010\rangle), \\
|\Phi_7\rangle &= \frac{1}{\sqrt{2}} (|100\rangle - |011\rangle).
\end{align*}
\]

as the eigenstates of the phase-bit operator \(X \otimes X \otimes X\) and

• 1st parity-bit operator: \(Z \otimes Z \otimes I_2\),
• 2nd parity-bit operator: \(I_2 \otimes Z \otimes Z\).

4° $n$-qubit: \(\ket{\text{GHZ}}_n = \frac{1}{\sqrt{2}} \left( \frac{|00\cdots0\rangle + |11\cdots1\rangle}{\sqrt{n}} \right)\), with \(n\) observables

\[
\begin{align*}
\text{phase-bit operator: } & \underbrace{\underbrace{X \otimes X \otimes \cdots \otimes X}}_n, \\
\text{1st parity-bit operator: } & \underbrace{Z \otimes Z \otimes I_2 \otimes \cdots \otimes I_2}_{n-2}, \\
\text{2nd parity-bit operator: } & \underbrace{I_2 \otimes Z \otimes Z \otimes I_2 \otimes \cdots \otimes I_2}_{n-3}, \\
& \vdots \\
\text{(n - 1)-th parity-bit operator: } & \underbrace{I_2 \otimes I_2 \otimes \cdots \otimes I_2 \otimes Z \otimes Z}_{n-2}.
\end{align*}
\]

\[
\begin{align*}
(X_1 \otimes X_2 \otimes \cdots \otimes X_n) |\text{GHZ}_n\rangle &= (X_1 \otimes X_2 \otimes \cdots \otimes X_n) \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right] \\
&= \frac{1}{\sqrt{2}} \left[ |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle + (-1)^{x_1} |0x_2x_3\cdots x_n\rangle \right] \\
&= (-1)^{x_1} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right] \\
&= (-1)^{x_1} |\text{GHZ}_n\rangle.
\end{align*}
\]
\[
\left[ Z_1 \otimes Z_2 \otimes I_2 \otimes \cdots \otimes I_2 \right]_{n-2} \left| \text{GHZ} \right>_n
= \left[ Z_1 \otimes Z_2 \otimes I_2 \otimes \cdots \otimes I_2 \right]_{n-2} \frac{1}{\sqrt{2}} \left( |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right)
= \frac{1}{\sqrt{2}} \left[ (-1)^{x_2} |0x_2x_3\cdots x_n\rangle + (-1)^{x_1}(-1)^{x_2+1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_2} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_2} |\text{GHZ}\rangle_n.
\]

\[
\left( I_2 \otimes Z_2 \otimes I_2 \otimes I_2 \right)_{n-3} \left| \text{GHZ} \right>_n
= \left( I_2 \otimes Z_2 \otimes I_2 \otimes I_2 \right)_{n-3} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= \frac{1}{\sqrt{2}} \left[ (-1)^{x_2+x_3} |0x_2x_3\cdots x_n\rangle + (-1)^{x_1}(-1)^{x_2+1}(-1)^{x_3+1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_2+x_3} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_2+x_3} |\text{GHZ}\rangle_n.
\]

\[
\left( I_2 \otimes I_2 \otimes Z_3 \otimes I_2 \otimes \cdots I_2 \right)_{n-4} \left| \text{GHZ} \right>_n
= \left( I_2 \otimes I_2 \otimes Z_3 \otimes I_2 \otimes \cdots I_2 \right)_{n-4} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= \frac{1}{\sqrt{2}} \left[ (-1)^{x_3+x_4} |0x_2x_3\cdots x_n\rangle + (-1)^{x_1}(-1)^{x_3+1}(-1)^{x_4+1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_3+x_4} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_3+x_4} |\text{GHZ}\rangle_n.
\]

\[
\left( I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n \right)_{n-2} \left| \text{GHZ} \right>_n
= \left( I_2 \otimes \cdots \otimes I_2 \otimes Z_{n-1} \otimes Z_n \right)_{n-2} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= \frac{1}{\sqrt{2}} \left[ (-1)^{x_{n-1}+x_n} |0x_2x_3\cdots x_n\rangle + (-1)^{x_1}(-1)^{x_{n-1}+1}(-1)^{x_n+1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_{n-1}+x_n} \frac{1}{\sqrt{2}} \left[ |0x_2x_3\cdots x_n\rangle + (-1)^{x_1} |1\bar{x}_2\bar{x}_3\cdots \bar{x}_n\rangle \right]
= (-1)^{x_{n-1}+x_n} |\text{GHZ}\rangle_n.
\]
We can then conclude that the \( n \)-qubit GHZ state, i.e., state \( |\text{GHZ}\rangle_n \), is the common eigenvector of the phase-bit operator, the 1st parity-bit operator, the 2nd parity-bit operator, \( \cdots \) and the \( (n-1) \)-th parity bit operator, with the corresponding eigenvalues \( (x_1, x_2, x_2 + x_3, \cdots, x_{n-1} + x_n) \).

5° As \( n \to \) large number, the \( n \)-qubit GHZ state defined as

\[
|\text{GHZ}\rangle_n = \frac{1}{\sqrt{2}} (|\text{live}\rangle + |\text{dead}\rangle)
\]

is also called the cat state, which is the macroscopically distinguished states used in Schrödinger’s cat experiment.

15.2.3 Properties of GHZ states

1° For the GHZ states \( |\text{GHZ}\rangle \), each qubit is maximally entangled with the other \( n-1 \) qubits, e.g.

\[
\rho_{A_1} := \text{tr}_{A_2A_3\cdots A_n} |\text{GHZ}\rangle_n \langle \text{GHZ}| = \frac{1}{2} \text{tr}_{A_2A_3\cdots A_n} \left( |0y\rangle_n \pm |1\bar{y}\rangle_n \right) \left( \langle 0y|_n \pm \langle 1\bar{y}|_n \right)
\]

with \( y \) as a \((n-1)\)-bit string of 0 and 1, thus

\[
\rho_{A_1} = \frac{1}{2} \sum_{y'} A_{A_3\cdots A_n} \left( |y'\rangle \left( |0y\rangle_n \pm |1\bar{y}\rangle_n \right) \left( \langle 0y|_n \pm \langle 1\bar{y}|_n \right) |y'\rangle_{A_2A_3\cdots A_n}
\]

\[
= \frac{1}{2} \sum_{y'} \left( |0\rangle \delta_{y'y} \pm |1\rangle \delta_{y'y} \right) \left( |0\rangle \delta_{yy'} \pm |1\rangle \delta_{yy'} \right)
\]

\[
= \frac{1}{2} \left( |0\rangle + \langle 1| \right)
\]

i.e.,

\[
\rho_{A_1} = \frac{1}{2} I_2. \tag{15.2.4}
\]

2° Cutting one qubit gives rise to a separable \((n-1)\)-qubit mixed state with rank 2.

\[
\rho_{A_2A_3\cdots A_n} = \text{tr}_{A_1} |\text{GHZ}\rangle_n \langle \text{GHZ}| = \frac{1}{2} \sum_{i=0}^{1} A_{i} \left( |iy\rangle_n \pm |i\bar{y}\rangle_n \right) \left( \langle iy|_n + \langle i\bar{y}|_n \right) |i\rangle_{A_1}
\]

\[
= \frac{1}{2} \left( |y\rangle_{n-1} \langle y| + |\bar{y}\rangle_{n-1} \langle \bar{y}| \right)
\]

from which we can find that \( \rho_{A_2A_3\cdots A_n} \) can be rewritten in the form as illustrated in E.Q. (15.2.2), namely \( \rho_{A_2A_3\cdots A_n} \) is separable. Note that we can apply the PPT criterion namely

\[
\rho_{A_2A_3\cdots A_n}^{PT} = (I_d \otimes T) \rho_{A_2A_3\cdots A_n} = \rho_{A_2A_3\cdots A_n} \geq 0, \tag{15.2.5}
\]

which is the necessary condition for quantum separability.

3° \( n \)-qubit GHZ states \( |\text{GHZ}\rangle_n \) form an orthonormal basis of \( n \)-qubit Hilber space, see Table 15.1

4° Given an \( n \)-qubit \( |\text{GHZ}\rangle_n \) state, other \((2^n - 1)\) \( |\text{GHZ}\rangle_n \) states can be generated via LOCC, e.g. in the case of \( n = 2 \), the other Bell states can be obtained from \( |\phi^+\rangle \) state via LOCC, as shown in Figure 15.4

156
Table 15.1: $|\text{GHZ}\pm\rangle_n$ states form an orthonormal basis of $n$-qubit Hilbert space.

<table>
<thead>
<tr>
<th>$n$</th>
<th>GHZ states (orthonormal basis)</th>
<th>$n$-qubit Hilbert space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$</td>
<td>$</td>
<td>\pm\rangle$</td>
</tr>
<tr>
<td>$n = 2$</td>
<td>$</td>
<td>\psi(i, j)\rangle$, $i, j = 0, 1$</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>$</td>
<td>\Phi_i\rangle$, $i = 1, 2, \ldots, 8$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>$</td>
<td>\text{GHZ}\pm\rangle_n$</td>
</tr>
</tbody>
</table>

Figure 15.4: Bell states generated from $|\phi^+\rangle$
Chapter 16

Entanglement Measures and Entropy: Bi-Partite System

The main problem is that we do not understand fully what entanglement is.

—Horodecki

Then the usefulness of entanglement emerges because it allows us to overcome a particular constraint that will be the LOCC constraint.

—Plenio and Virmani

It is the constraint to LOCC operations that entanglement to the status of a resource.

—Plenio and Virmani

We will simply accept the non-uniqueness of entanglement measures as an expression of the fact that they correspond to different operational tasks under which different forms of entanglement may have different degree of usefulness.

—Plenio and Virmani

Reference:

• [Preskill] New Chapter 4: Quantum entanglement;

16.1 Pure states and mixed states

1° A quantum system is described by the density matrix which is

(1) linear & non-negative operators, i.e., $\rho \geq 0$;

(2) Hermitian: $\rho^\dagger = \rho$;

(3) Trace 1: $\text{tr}(\rho) = 1$.

2° $\rho$ is called pure state iff $\rho^2 = \rho$ or $\text{tr}(\rho^2) = 1$.

Note: $\rho$ is a pure state $\iff \rho = |\psi\rangle\langle\psi|$. 
3° \( \rho \) is called mixed state iff \( \rho^2 \neq \rho \) or \( \text{tr}(\rho^2) \leq 1 \).

**Note**: The mixed state has the formalism

\[
\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|\]  \hspace{1cm} (16.1.1)

with \( i \geq 2, p_i \geq 0 \) and \( \sum_i p_i = 1 \). It stands for a state ensemble \( \{p_i, |\psi_i\rangle\} \). Every mixed state density matrix has infinity number of state ensembles, which is called the ambiguity of the density matrix.

### 16.2 Bipartite entangled pure states

**Def 16.2.1.** The bipartite pure state \( |\chi\rangle_{AB} \) is called separable or factorizable if it can be written as a product of two states, namely

\[
|\chi\rangle_{AB} = |\varphi\rangle_A \otimes |\psi\rangle_B. \] \hspace{1cm} (16.2.1)

**Def 16.2.2.** The bipartite pure state \( |\chi\rangle_{AB} \) is entangled if it is not separable.

**Thm 16.2.0.1.** With the Schmidt decomposition of \( |\chi\rangle_{AB} \), the Schmidt number is 1, then \( |\chi\rangle_{AB} \) is separable, and the Schmidt number is greater than 1, then \( |\chi\rangle_{AB} \) is entangled.

**Thm 16.2.0.2.** The density matrix of bipartite system has the formalism expressed as

\[
\rho_{AB} := |\chi\rangle_{AB} \langle \chi|. \]

1° When \( \rho_A = \text{tr}_B \rho_{AB} \) or \( \rho_B = \text{tr}_A \rho_{AB} \) is pure state, then \( |\chi\rangle_{AB} \) is separable;

2° When \( \rho_A \) or \( \rho_B \) is mixed state, then \( |\chi\rangle_{AB} \) is entangled;

3° When \( \rho_A = \rho_B = \frac{1}{2} \text{Id} \), then \( |\chi\rangle_{AB} \) is maximal entangled.

**Proof.**

1° When

\[
|\chi\rangle_{AB} := |\varphi\rangle_A \otimes |\psi\rangle_B, \] \hspace{1cm} (16.2.2)

then

\[
\rho_{AB} = |\varphi\rangle_A \langle \varphi| \otimes |\psi\rangle_B \langle \psi|. \] \hspace{1cm} (16.2.3)

Therefore

\[
\rho_A = |\varphi\rangle_A \langle \varphi|, \quad \rho_B = |\psi\rangle_B \langle \psi|, \] \hspace{1cm} (16.2.4)

which are pure states.

2° With the Schmidt decomposition of state \( |\chi\rangle_{AB} \), density matrix \( \rho_A \) or \( \rho_B \) is mixed state which is directly associated with the Schmidt decomposition number of state \( |\chi\rangle_{AB} \) greater than 1.

3° Maximally entangled state means our maximal ignorance on subsystem, i.e., \( \rho \propto \text{Id} \).

**e.g.1.** Bell state:

\[
|\Psi\rangle_{AB} := \frac{1}{\sqrt{2}} \left( |00\rangle + |11\rangle \right), \] \hspace{1cm} (16.2.5)

\[
\rho_A = \rho_B = \frac{1}{2} I_2. \] \hspace{1cm} (16.2.6)
Higher dimensional Bell state:

\[ |\Psi\rangle_{AB} := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle, \quad (16.2.7) \]

\[ \rho_A = \rho_B = \frac{1}{d} I_d. \quad (16.2.8) \]

**Def 16.2.3 (Von-Neumann entropy).** The von-Neumann entropy is a distinguished measure of bipartite entangled pure state \( \rho_{AB} \) defined as

\[ S(\rho_A) := -\text{tr}(\rho_A \log \rho_A), \quad (16.2.9) \]

\[ S(\rho_B) := -\text{tr}(\rho_B \log \rho_B). \quad (16.2.10) \]

**Properties:**

1° \( S(\rho_A) = 0 \) implies \( \rho_A \) is a pure state.

2° \( S(\rho_A) \) is in the domain \( 0 \leq S(\rho_A) \leq \log d \) where \( d = \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B) \).

3° In the diagonalized form of density matrix \( \rho \),

\[ \rho = \sum \lambda_i |\psi_i\rangle \langle \psi_i|, \quad (16.2.11) \]

von-Neumann entropy has the formulation

\[ S(\rho) = -\text{tr}(\rho \log \rho) = -\sum \lambda_i \log \lambda_i. \quad (16.2.12) \]

**Note 1:** The von-Neumann entropy is defined only for pure states not for mixed states.

**Note 2:** The von-Neumann entropy is a measure of our ignorance of the quantum state, which is associated with the Shannon entropy in information theory or thermodynamics.

**Examples:**

e.g.1. For the separable pure state

\[ \rho_A := |\varphi\rangle_A \langle \varphi| = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \quad (16.2.13) \]

the von-Neumann entropy is

\[ S(\rho_A) = -\text{tr} \left( \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \log \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \right) = 0, \quad (16.2.14) \]

where the relations

\[ 1 \cdot \log 1 = 0, \quad 0 \cdot \log 0 = 0 \quad (16.2.15) \]

have been applied.
e.g.2. For the higher dimensional Bell state:

\[ |\Psi\rangle_{AB} := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle, \quad (16.2.16) \]

\[ \rho_A = \frac{1}{d} I_d = \left( \begin{array}{cccc}
\frac{1}{d} & & & \\
& \frac{1}{d} & & \\
& & \ddots & \\
& & & \frac{1}{d}
\end{array} \right), \quad (16.2.17) \]

\[
S(\rho_A) = -\text{tr}\left( \left( \begin{array}{cccc}
\frac{1}{d} & & & \\
& \frac{1}{d} & & \\
& & \ddots & \\
& & & \frac{1}{d}
\end{array} \right) \log \left( \begin{array}{cccc}
\frac{1}{d} & & & \\
& \frac{1}{d} & & \\
& & \ddots & \\
& & & \frac{1}{d}
\end{array} \right) \right)
= -\text{tr}\left( \left( \begin{array}{cccc}
\frac{1}{d} \log \frac{1}{d} & & & \\
& \frac{1}{d} \log \frac{1}{d} & & \\
& & \ddots & \\
& & & \frac{1}{d} \log \frac{1}{d}
\end{array} \right) \right)
= -d \cdot \frac{1}{d} \log \frac{1}{d} = \log d. \quad (16.2.18)\]

So the higher dimensional Bell state is the maximal entangled states.

**Def 16.2.4.** *Arbitrary two-qubit pure state has the formalism*

\[ |\psi\rangle := \alpha_1|00\rangle + \alpha_2|01\rangle + \alpha_3|10\rangle + \alpha_4|11\rangle \quad (16.2.19) \]

*with the matrix formulation expressed as*

\[ |\psi\rangle := \left( \begin{array}{c}
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\alpha_4
\end{array} \right). \quad (16.2.20) \]

*The Pauli \( \sigma_y \) is defined as*

\[ \sigma_y := \left( \begin{array}{cc}
0 & -i \\
i & 0
\end{array} \right). \quad (16.2.21) \]

*And we have*

\[ |\tilde{\psi}\rangle := (\sigma_y \otimes \sigma_y) |\psi^*\rangle 
= \left( \begin{array}{cccc}
1 & & & -1 \\
& 1 & & \\
& & \ddots & \\
& & & 1
\end{array} \right) \left( \begin{array}{c}
\alpha_1^* \\
\alpha_2^* \\
\alpha_3^* \\
\alpha_4^*
\end{array} \right)
= \left( \begin{array}{c}
-\alpha_1^* \\
\alpha_3^* \\
\alpha_2^* \\
-\alpha_4^*
\end{array} \right), \quad (16.2.22) \]

*where state \( |\psi^*\rangle \) is denoted as the complex conjugate of the state \( |\psi\rangle \).*
The concurrence of two-qubit pure state is defined as

\[
C(\langle \psi | \rangle) := |\langle \psi | \bar{\psi} \rangle|
\]

\[
= |\begin{pmatrix} \alpha_1^* & \alpha_2^* & \alpha_3^* & \alpha_4^* \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{pmatrix}|
\]

\[
= 2|\alpha_2\alpha_3 - \alpha_1\alpha_4|.
\]  \hspace{1cm} (16.2.23)

### 16.3 Example for two-qubit pure entangled state

For arbitrary two-qubit pure state

\[
|\psi\rangle_{AB} := \alpha_1|00\rangle + \alpha_2|01\rangle + \alpha_3|10\rangle + \alpha_4|11\rangle,
\]

the concurrence is

\[
C(|\psi\rangle_{AB}) = 2|\alpha_1\alpha_4 - \alpha_2\alpha_3|.
\]  \hspace{1cm} (16.3.1)

The density matrix of subsystem \(A\) is

\[
\rho_A = \text{tr}_B|\psi\rangle_{AB}\langle \psi |
\]

\[
= B\langle 0|\rho_{AB}|0\rangle_B + B\langle 1|\rho_{AB}|1\rangle_B
\]

\[
= \begin{pmatrix}
\alpha_1^* \alpha_1 + \alpha_2^* \alpha_2 & \alpha_3^* \alpha_1 + \alpha_4^* \alpha_2 \\
\alpha_1^* \alpha_3 + \alpha_2^* \alpha_4 & \alpha_3^* \alpha_3 + \alpha_4^* \alpha_4
\end{pmatrix}
\]

\hspace{1cm} (16.3.3)

with the eigenvalue

\[
\lambda_{\pm}(\rho_A) = \frac{1 \pm \sqrt{1 - 4(\alpha_1\alpha_4 - \alpha_2\alpha_3)(\alpha_1^*\alpha_4^* - \alpha_2^*\alpha_3^*)}}{2}.
\]  \hspace{1cm} (16.3.4)

According to the formula \[16.2.12\], the von-Neumann entropy of subsystem \(A\) is

\[
S(\rho_A) = -\lambda_+ \ln \lambda_+ - \lambda_- \ln \lambda_-.
\]  \hspace{1cm} (16.3.5)

1° Bell states.

\[
|\phi_{\pm}\rangle := \frac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right), \quad |\psi_{\pm}\rangle := \frac{1}{\sqrt{2}}\left(|01\rangle \pm |10\rangle\right).
\]  \hspace{1cm} (16.3.6)

Von-Neumann entropy:

\[
S(|\phi_{\pm}\rangle) = \ln 2, \quad S(|\psi_{\pm}\rangle) = \ln 2.
\]  \hspace{1cm} (16.3.7)

Concurrence:

\[
C(|\phi_{\pm}\rangle) = 1, \quad C(|\psi_{\pm}\rangle) = 1.
\]  \hspace{1cm} (16.3.8)

Bell states are maximal entangled states.

2° Separable states.

\[
|\psi\rangle_{AB} := (x_1|0\rangle + x_2|1\rangle) \otimes (x_3|0\rangle \otimes x_4|1\rangle).
\]  \hspace{1cm} (16.3.9)

Von-Neumann entropy:

\[
S(|\psi\rangle_{AB}) = 0.
\]  \hspace{1cm} (16.3.10)

Concurrence:

\[
C(|\psi\rangle_{AB}) = 0.
\]  \hspace{1cm} (16.3.11)
Evolution of Bell states.

\[ |\psi\rangle_{AB} := \cos\left(\frac{\pi}{4} - \theta\right)|00\rangle_{AB} - e^{i\varphi} \sin\left(\frac{\pi}{4} - \theta\right)|11\rangle_{AB}, \]  

(16.3.12)

where \(-\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}\) and \(0 \leq \varphi \leq 2\pi\).

Von-Neumann entropy:

The density matrix of subsystem A denoted as \(\rho_A\) has the eigenvalues

\[ \lambda_\pm = \frac{1 \pm \sin 2\theta}{2}. \]  

(16.3.13)

Thus

\[ S(\rho_A) = -\frac{1 + \sin 2\theta}{2} \ln \frac{1 + \sin 2\theta}{2} - \frac{1 - \sin 2\theta}{2} \ln \frac{1 - \sin 2\theta}{2}. \]  

(16.3.14)

When \(\theta = 0\), state \(|\psi\rangle_{AB}\) evolves into Bell states with the von-Neumann entropy

\[ S(\rho_A)|_{\theta=0} = -\frac{1}{2} \ln \frac{1}{2} - \frac{1}{2} \ln \frac{1}{2} = \ln 2. \]  

(16.3.15)

When \(\theta = \frac{\pi}{4}\), state \(|\psi\rangle_{AB}\) evolves into separable states with the von-Neumann entropy

\[ S(\rho_A)|_{\theta=\frac{\pi}{4}} = 0. \]  

(16.3.16)

Concurrence:

\[ C(|\psi\rangle_{AB}) = 2 \left| \cos\left(\frac{\pi}{4} - \theta\right) \sin\left(\frac{\pi}{4} - \theta\right) \right| = |\cos 2\theta|. \]  

(16.3.17)

**Remarks:** \(S(|\psi\rangle)\) and \(C(|\psi\rangle)\) are different measures.

1° Degree of entanglement measure of separable state is zero.

2° Bell states are maximally entangled states.

3° The configurations of \(C(|\psi\rangle)\) and \(S(|\psi\rangle)\) look the same in shape.

### 16.4 Axioms of entanglement measure

**Def 16.4.1.** An entanglement measure or degree of entanglement is a real valued function \(E(\rho)\) to quantify the amount of entanglement in a given quantum state \(\rho\).

**Necessary requirements on** \(E(\rho)\).

1° \(E(\rho)\) is a mapping from a bipartite state \(\rho\) to a positive valued real number, namely

\[ \rho \rightarrow E(\rho) \in \mathbb{R}^+. \]  

(16.4.1)

2° Normalization.

\[ E(\rho) \leq E(|\Psi\rangle_{AB}\langle\Psi|) = \ln d, \]  

(16.4.2)

where \(|\Psi\rangle\) is Bell state.
3° \( E(\rho) = 0 \) for separable states.

4° \( E(\langle \psi \rangle_{AB} \langle \psi \rangle) = S(\text{tr}_B \langle \psi \rangle_{AB} \langle \psi \rangle) \) for bipartite pure states.

5° \( E(\rho) \) is not increased by LOCC.

\[
\rho \xrightarrow{\text{LOCC}} \rho' = \sum_i A_i \rho A_i^\dagger, \quad \sum_i A_i A_i^\dagger = \text{Id}.
\]

(16.4.3)

\[
E(\rho) \geq E(\rho').
\]

(16.4.4)

6° Convexity.

\[
E(\lambda \rho + (1 - \lambda) \sigma) \leq \lambda E(\rho) + (1 - \lambda) E(\sigma).
\]

(16.4.5)

7° Additivity.

\[
E(\rho \otimes \sigma) \leq E(\rho) + E(\sigma).
\]

(16.4.6)

8° Subadditivity.

\[
E(\rho) - E(\sigma) \to 0, \quad \text{if } |\rho - \sigma| \to 0.
\]

(16.4.7)

9° Continuity.

\[
E(\rho) - E(\sigma) \to 0, \quad \text{if } |\rho - \sigma| \to 0.
\]

(16.4.8)

16.5 Example: entanglement of formation

**Def 16.5.1.** The entanglement of formation of the mixed state \( \rho = \sum_i p_i \langle \psi_i \rangle \langle \psi_i \rangle \) is defined as

\[
E_F(\rho) := \text{Inf} \left\{ \sum_i p_i E(\langle \psi_i \rangle \langle \psi_i \rangle) \right\},
\]

(16.5.1)

where Inf stands for Infimum, namely great lowest bound.

E.g. For a bipartite mixed state \( \rho \),

\[
E_F(\rho) = S \left( \frac{1 + \sqrt{1 - C^2(\rho)}}{2} \right),
\]

(16.5.2)

where \( C(\rho) \) is the concurrence and the function \( S \) is defined as

\[
S(x) = -x \log_2 x - (1 - x) \log_2 (1 - x).
\]

(16.5.3)

With the new defined density matrix denoted as \( \tilde{\rho} \)

\[
\tilde{\rho} := (\sigma_y \otimes \sigma_y) \rho (\sigma_y \otimes \sigma_y),
\]

(16.5.4)

the non-Hermitian matrix \( \rho \tilde{\rho} \) has the diagonized formalism

\[
\rho \tilde{\rho} = \sum_{i=1}^{4} \lambda_i^2 \langle \psi_i \rangle \langle \psi_i \rangle,
\]

(16.5.5)

where \( \lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 \geq 0 \).

\[
C(\rho) = \max \{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}.
\]

(16.5.6)

**Note:** Entanglement of formation is the entanglement measure for \( \rho \) neither the concurrence nor the von-Neumann entropy.
16.6 PPT criterion for quantum separability

PPT is the abbreviation of positive partial transpose criterion or Peres-Horodecki criterion for quantum separability.

Thm 16.6.0.3. For a quantum state $\rho_{AB}$, when $\rho_{AB}$ is separable, then

$$\rho_{AB}^T := (\text{Id} \otimes T) \rho_{AB} \geq 0,$$

which is only a necessary condition for separability.

Thm 16.6.0.4. When $\dim \mathcal{H}_A = 2$ and $\dim \mathcal{H}_B = 2$ or 3, then

$$\rho_{AB} \text{ is separable } \iff \rho_{AB}^T \geq 0.$$

Corollary 16.6.0.1. When $\rho_{AB}^T$ is not positive, then $\rho_A$ is entangled.

E.g. Werner state in dimension 2.

$$\rho(\lambda) := \frac{1}{4} (1 - \lambda) I_4 + \lambda |\phi^+\rangle \langle \phi^+|$$

with

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle).$$

Note:

$$\rho = \begin{pmatrix} A & C \\ C^\dagger & D \end{pmatrix}$$

where $A$, $B$, $C$ and $D$ are $2 \times 2$ matrix.

$$\rho_A^T := (T \otimes I_2) \rho = \begin{pmatrix} A & C^\dagger \\ C & B \end{pmatrix};$$

$$\rho_B^T := (I_2 \otimes T) \rho = \begin{pmatrix} A^\dagger & C^T \\ C^* & B^* \end{pmatrix}.$$

The Werner state has the matrix formalism

$$\rho(\lambda) = \begin{pmatrix} \frac{1}{4}(1 - \lambda) & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4}(1 - \lambda) & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{4}(1 - \lambda) \end{pmatrix},$$

and the its partial transpose matrix has the eigenvalues $\rho_A^T(\lambda) = \{ \frac{1}{4}(1 + \lambda), \frac{1}{2}(1 - 3\lambda) \}$. Therefore, when $\lambda > \frac{1}{3}$, according to the PPT criterion, Werner state is entangled. Detailed calculations refer to the section 15.1.4.

16.7 Example: Werner state in $d$-dimension

The Werner state in dimension $d$ is defined as

$$\rho(\lambda) := \frac{1}{d^2} (1 - \lambda) I_d^2 + \lambda |\Phi\rangle \langle \Phi|$$

with

$$|\Phi\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle,$$

which is a maximally entangled state, namely $S(|\Phi\rangle) = \ln d$. 165
16.7.1 \( \rho(\lambda) \) is a density matrix

1° Trace 1.
\[
\text{tr} \rho(\lambda) = \lambda + \frac{1}{d^2} (1 - \lambda) \text{tr} I_{d^2} = \lambda + 1 - \lambda = 1. \tag{16.7.3}
\]

2° Hermitian.
\[
\rho^\dagger(\lambda) = \frac{1}{d^2} (1 - \lambda) I_{d^2} + \lambda \Phi \langle \Phi \rangle = \rho(\lambda). \tag{16.7.4}
\]

3° Positivity.
\[
\rho(\lambda) |\Phi\rangle = (\lambda + \frac{1}{d^2} (1 - \lambda)) |\Phi\rangle. \tag{16.7.5}
\]
\[
\rho(\lambda) |\Phi\rangle = \frac{1}{d^2} (1 - \lambda) |\Phi\rangle, \tag{16.7.6}
\]

where state \(|\Phi\rangle\) is orthogonal to the state \(|\Phi\rangle\). \(\rho(\lambda)\) has the eigenstate \(|\Phi\rangle\) with the eigenvalue denoted as \(\tau_{\Phi} = \lambda + \frac{1}{d^2} (1 - \lambda)\).

\[
\rho(\lambda) |\Phi\rangle = \frac{1}{d^2} (1 - \lambda) |\Phi\rangle, \tag{16.7.7}
\]

\(\rho(\lambda)\) is a density matrix.

16.7.2 Partial transpose of \(\rho(\lambda)\)

\[
(|\Phi\rangle \langle \Phi|)^{TB} = (I_d \otimes T) \frac{1}{d} \sum_{i,j=0}^{d-1} |ij\rangle \langle j|i| \]
\[
= \frac{1}{d} \sum_{i,j=1}^{d} |ij\rangle \langle ji| \]
\[
= \frac{1}{d} \text{SWAP}. \tag{16.7.8}
\]

Construct the symmetric eigenstate of the SWAP gate denoted as \(|S(ij)\rangle\)
\[
|S(ij)\rangle := |ij\rangle + |ji\rangle, \tag{16.7.9}
\]

from which we have
\[
\text{SWAP} |S(ij)\rangle = |\Phi\rangle. \tag{16.7.10}
\]

Construct the anti-symmetric eigenstate of the SWAP gate denoted as \(|A(ij)\rangle\)
\[
|A(ij)\rangle := |ij\rangle - |ji\rangle, \tag{16.7.11}
\]

from which we have
\[
\text{SWAP} |A(ij)\rangle = -|A(ij)\rangle. \tag{16.7.12}
\]

With the new-defined operators denoted as \(E_{\text{sym}}\) and \(E_{\text{anti}}\) expressed as
\[
E_{\text{sym}} = \frac{1}{4} \sum_{i,j=1}^{d} |S(ij)\rangle \langle S(ij)|, \tag{16.7.13}
\]
\[
E_{\text{anti}} = \frac{1}{4} \sum_{i,j=1}^{d} |A(ij)\rangle \langle A(ij)|.
\]
the SWAP gate has the formalism
\[
\text{SWAP} = E_{\text{sym}} - E_{\text{anti}} = I_{d^2} - 2E_{\text{anti}}, \quad (16.7.14)
\]
where the relation \( E_{\text{sym}} + E_{\text{anti}} = I_{d^2} \) has been applied.

\[
\begin{align*}
\text{tr}E_{\text{sym}} &= \frac{1}{4}(d^2 + d + d^2 + d) = \frac{1}{2}d(d + 1), \quad (16.7.15) \\
\text{tr}E_{\text{anti}} &= \frac{1}{4}(d^2 - d + d^2 - d) = \frac{1}{2}d(d - 1), \quad (16.7.16) \\
\text{tr}E_{\text{sym}} + \text{tr}E_{\text{anti}} &= \frac{1}{2}d(d + 1) + \frac{1}{2}d(d - 1). \quad (16.7.17)
\end{align*}
\]

### 16.7.3 Positivity of partial transpose \( \rho(\lambda) \)

The partial transpose of the Werner state is denoted as
\[
(\rho(\lambda))^{T_B} = \frac{1}{d^2}(1 - \lambda)I_{d^2} + \lambda \frac{1}{d}(I_{d^2} - 2E_{\text{anti}})
\]
which has the eigenvalues denoted as \( \tau_{\text{sym}} \) and \( \tau_{\text{anti}} \)
\[
\begin{align*}
\tau_{\text{sym}} &= \frac{1}{d}\left(\frac{1 - \lambda}{d} + \lambda\right), \quad (16.7.19) \\
\tau_{\text{anti}} &= \frac{1}{d}\left(\frac{1 - \lambda}{d} + \lambda\right) - \frac{2\lambda}{d}. \quad (16.7.20)
\end{align*}
\]

The positivity requires the parameter \( \lambda \) in the region
\[
\frac{1}{1 - d} \leq \lambda \leq \frac{1}{1 + d}, \quad (16.7.21)
\]

therefore, with the constrain in E.Q. (16.7.7), we can infer that \( \rho(\lambda) \) is entangled when
\[
\frac{1}{1 + d} \leq \lambda \leq 1. \quad (16.7.22)
\]

**Note:** We do not know whether \( \rho(\lambda) \) is separable or not when \( \frac{1}{1 - d} \leq \lambda \leq \frac{1}{1 + d} \), since PPT criterion is only a necessary condition.

### 16.8 Example: the other form of the Werner state

\[
\rho_{\text{anti}}(\lambda) := \frac{1}{d^2}(1 - \lambda)I_{d^2} + \lambda \frac{1}{2d(d - 1)}E_{\text{anti}}. \quad (16.8.1)
\]

#### 16.8.1 \( \rho_{\text{anti}} \) is a density matrix

1° Trace 1.

\[
\begin{align*}
\text{tr}\rho_{\text{anti}}(\lambda) &= \frac{1}{d^2}(1 - \lambda)\text{tr}(I_{d^2}) + \lambda \frac{1}{2d(d - 1)}\text{tr}(E_{\text{anti}}) \\
&= \frac{1}{2}(1 - \lambda) + \lambda \frac{1}{2d} \text{tr}(E_{\text{anti}}) \\
&= 1 - \lambda + \lambda \\
&= 1. \quad (16.8.2)
\end{align*}
\]
2° Hermitian.

\[ \rho_{\text{anti}}^\dagger(\lambda) = \frac{1}{d^2}(1 - \lambda)I_{d^2} + \lambda \frac{1}{2d(d-1)}E_{\text{anti}}^\dagger \]

\[ = \frac{1}{d^2}(1 - \lambda)I_{d^2} + \lambda \frac{1}{2d(d-1)}E_{\text{anti}} \]

\[ = \rho_{\text{anti}}(\lambda). \quad (16.8.3) \]

3° Positivity.

\[ \tau_{\text{sym}} = \frac{1}{d^2}(1 - \lambda), \quad (16.8.4) \]

\[ \tau_{\text{anti}} = \lambda \frac{1}{2d(d-1)} + \frac{1}{d^2}(1 - \lambda). \quad (16.8.5) \]

\[ \rho_{\text{anti}} \geq 0 \iff \frac{1-d}{1+d} \leq \lambda \leq 1. \quad (16.8.6) \]

16.8.2 Partial transpose of \( \rho_{\text{anti}}(\lambda) \)

\[ (|\Phi\rangle\langle\Phi|)^{PT} = \frac{1}{d}(I_{d^2} - 2E_{\text{anti}}), \]

\[ |\Phi\rangle\langle\Phi| = \frac{1}{d}(I_{d^2} - 2E_{\text{anti}}^{PT}), \quad (16.8.7) \]

therefore

\[ E_{\text{anti}}^{PT} = \frac{1}{2}(I_{d^2} - d|\Phi\rangle\langle\Phi|). \quad (16.8.8) \]

\[ \rho_{\text{anti}}^{PT} = \frac{1}{d^2}(1 - \lambda)I_{d^2} + \frac{\lambda}{2d(d-1)}E_{\text{anti}}^{PT} \]

\[ = \frac{1}{d^2(d-1)}(\lambda d + (d-1)(1-\lambda))I_{d^2} - \frac{\lambda}{d-1}|\Phi\rangle\langle\Phi|, \quad (16.8.9) \]

which has the eigenvalues

\[ \tau_{\Phi} = \frac{1}{d^2(d-1)}(\lambda d + (d-1)(1-\lambda)) - \frac{\lambda}{d-1}, \]

\[ \tau_{\Phi^\perp} = \frac{1}{d^2(d-1)}(\lambda d + (d-1)(1-\lambda)). \quad (16.8.10) \]

\[ \rho_{\text{anti}}^{PT} \geq 0 \iff 1-d \leq \lambda \leq \frac{1}{1+d}. \quad (16.8.11) \]

When \( \frac{1-d}{1+d} \leq \lambda \leq 1 \), \( \rho_{\text{anti}} \) is entangled. However, we do not know \( \rho_{\text{anti}} \) is separable or not, when \( \frac{1-d}{1+d} \leq \lambda \leq \frac{1}{1+d} \).
Part IV

Quantum Open System and Quantum Error Correction Codes
Chapter 17

Quantum Mechanics (III): Quantum Open System

Real systems suffer from unwanted interactions with the outside world. These unwanted interactions show up as noise in quantum information processing systems. We need to understand and control such noise processes in order to build useful quantum information processing systems.

—Nielsen & Chuang

No quantum systems are ever perfectly closed, and especially not quantum computers, which must be delicately programmed by an external system to perform some desired set of operations.

—Nielsen & Chuang

An open system is nothing more than one which has interactions with some other environment system, whose dynamics we wish to neglect or average over.

—Nielsen & Chuang

The mathematical formalism of quantum operations is the key tool for our description of the dynamics of open quantum systems.

—Nielsen & Chuang

Another advantage of quantum operations in applications to quantum computation and quantum information is that they are especially well adapted to describe discrete state changes, that is, transformations between an initial state $\rho$ and $\rho'$, without explicit reference to the passage of time.

—Nielsen & Chuang

We will construct an artificial example of a system whose evolution is not described by a quantum operation, … It is an interesting problem for further study to study quantum information processing beyond the quantum operation formalism.

—Nielsen & Chuang

References:

- [Preskill] Chapter 2: Foundations II: measurement and evolution;
- [Nielsen & Chuang] Chapter 8: Quantum noise and quantum operation.
17.1 Introduction

17.1.1 Why we talk about quantum open system

**Reason 1:** Quantum Computer is an open system. Quantum computer will interact with the environment (noise). The observers (control) will also interact with the Quantum Computer.

**Reason 2:** Quantum Computer is a many-qubit system, and it consists of a set of sub-systems which interact with each other.

17.1.2 Closed system and open system

<table>
<thead>
<tr>
<th>closed system</th>
<th>open system</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>pure state/Ray</td>
</tr>
<tr>
<td>observable</td>
<td>self-adjoint</td>
</tr>
<tr>
<td>measurement</td>
<td>projective measurement</td>
</tr>
<tr>
<td>evolution</td>
<td>Shrôdinger Equation</td>
</tr>
</tbody>
</table>

Table 17.2: State and quantum gate in closed and open system

<table>
<thead>
<tr>
<th>state</th>
<th>quantum Gate</th>
</tr>
</thead>
<tbody>
<tr>
<td>closed system</td>
<td>pure state: $</td>
</tr>
<tr>
<td>open system</td>
<td>mixed state: $\rho = \frac{1}{2} (1 + \vec{p} \cdot \vec{\sigma})$</td>
</tr>
</tbody>
</table>

17.2 Projective measurement

The following three points are equivalent:

1° An arbitrary observable $\mathcal{O}$ can be expressed as

$$\mathcal{O} = \sum_{n} a_n \Pi_n,$$

(17.2.1)

where

$$\Pi_n := |n\rangle \langle n|$$

(17.2.2)

with $a_n$ being the eigenvalue of the observable $\mathcal{O}$ associated with the eigenvector $|n\rangle$. 
The mean value of the observable is

\[
\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle = \langle \psi | \sum_n a_n \Pi_n | \psi \rangle = \sum_n a_n \langle \psi | n \rangle \langle n | \psi \rangle = \sum_n a_n P_n, \tag{17.2.3}
\]

with

\[
P_n := | \langle \psi | n \rangle |^2. \tag{17.2.4}
\]

2° In the orthonormal basis \( \{ |n \rangle \} \), observables can be determined by the basis.

The probability, that the post-measurement state is \( |n \rangle \), should be

\[
\text{Prob}(n) = | \langle n \rangle \langle n | \psi \rangle |^2 = \langle \psi | n \rangle \langle n | \psi \rangle = \langle \psi | \Pi_n | \psi \rangle = \text{tr}(\rho \Pi_n). \tag{17.2.5}
\]

The post-measurement state is

\[
| \psi \rangle \rightarrow |n \rangle \frac{\langle n \rangle \langle \psi |}{\sqrt{\text{Prob}(n)}}, \tag{17.2.6}
\]

or

\[
\rho \rightarrow \rho' = \sum_n \text{Prob}(n) \frac{\Pi_n \rho \Pi_n}{\text{Prob}(n)} = \Pi_n \rho \Pi_n. \tag{17.2.7}
\]

The expression is for the pure state.

3° The complete set of orthogonal projectors \( \{ \Pi_n \} \) are defined as

\[
\begin{align*}
\Pi_n \Pi_m &= \delta_{mn}, \text{ Orthonormal;} \\
\Pi_n^\dagger &= \Pi_n, \text{ Hermition;} \\
\Pi_n^2 &= \Pi_n, \text{ Projection;} \\
\sum_n \Pi_n &= \text{Id}, \text{ Completeness.}
\end{align*} \tag{17.2.8}
\]

Now, if we consider a mixed state described by the density matrix, then after the measurement,

\[
\rho \rightarrow \rho' = \sum_n \text{Prob}(n) \frac{\Pi_n \rho \Pi_n}{\text{Prob}(n)} = \sum_n \Pi_n \rho \Pi_n \tag{17.2.9}
\]

which gives rise to the ensemble

\[
\rho' = \{ \text{Prob}(n) \frac{\Pi_n \rho \Pi_n}{\text{Prob}(n)} \}.
\]

Example: A single-qubit density matrix is defined as

\[
\rho(\vec{p}) := \frac{1}{2} (1 + \hat{p} \cdot \vec{\sigma}), \quad \text{with } \vec{p} \in \mathbb{R}^3, \text{ and } 0 \leq |\vec{p}| \leq 1. \tag{17.2.10}
\]
We can verify the projectors defined as

\[
\begin{align*}
\Pi_1 &:= \frac{1}{2} (1 + \hat{n} \cdot \hat{\sigma}) = |\hat{n}\rangle \langle \hat{n}|, \\
\Pi_2 &:= \frac{1}{2} (1 - \hat{n} \cdot \hat{\sigma}) = |\hat{n}\rangle \langle -\hat{n}|,
\end{align*}
\]

satisfying the definition

\[
\begin{align*}
\Pi_1 + \Pi_2 &= \text{Id}, \\
\Pi_i \Pi_j &= \delta_{ij} \text{Id}, \\
\Pi_i^\dagger &= \Pi_i.
\end{align*}
\]

The probabilities that the post-measurement state is $|\hat{n}\rangle$ and $|\hat{n}\rangle$ can be calculated respectively

\[
\begin{align*}
P_1 &= \text{tr}(\rho \Pi_1) = \frac{1}{2} (1 + \hat{p} \cdot \hat{n}), \\
P_2 &= \text{tr}(\rho \Pi_2) = \frac{1}{2} (1 - \hat{p} \cdot \hat{n}),
\end{align*}
\]

### 17.3 General measurement theory

1° **Measurement operators**

- measurement operators $\{M_m\}$;
- the measurements can be represented by measurement operators;
- measurement operator $M_m$ is labeled by $m$;
- completeness relation $\sum_m M_m^\dagger M_m = \text{Id}$.

The projector is defined as

\[
\Pi_n := |n\rangle \langle n|.
\]

The measurement operator is defined as

\[
M_n := \Pi_n.
\]

Therefore,

\[
M_n^\dagger M_n = \Pi_n,
\]

denotes

\[
\sum_n M_n^\dagger M_n = \sum_n \Pi_n \\
= \sum_n |n\rangle \langle n| \\
= \text{Id}.
\]

2° **State**

- pure state $|\psi\rangle$;
- mixed state $\rho$.

3° **Measurement statistics**
• pure state: \( \text{Prob}(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle \);

• mixed state: \( \text{Prob}(m) = \text{tr} \left( M_m^\dagger M_m \rho \right) \).

4° Total probability:

• pure state:

\[
\sum_m \text{Prob}(m) = \sum_m \langle \psi | M_m^\dagger M_m | \psi \rangle \\
= \langle \psi | \sum_m M_m^\dagger M_m | \psi \rangle \\
= \langle \psi | \psi \rangle \\
= 1; \quad (17.3.5)
\]

• mixed state:

\[
\sum_m \text{Prob}(m) = \sum_m \text{tr} \left( M_m^\dagger M_m \rho \right) \\
= \text{tr} \left( \sum_m M_m^\dagger M_m \rho \right) \\
= \text{tr} \rho \\
= 1. \quad (17.3.6)
\]

5° Post-measurement state:

• pure state:

\[ |\psi\rangle \rightarrow \frac{M_m |\psi\rangle}{\sqrt{\text{Prob}(m)}}; \]

• mixed state:

\[ \rho \rightarrow \frac{M_m \rho M_m^\dagger}{\text{Prob}(m)}. \]

17.4 Definition of POVM

POVM is the abbreviation of “Positive Operator-Valued Measurement”. Usually, we do not care for the post measurement state, because the post measurement states may be destroyed in experiments immediately. Therefore, only the measurement probability matters.

Def 17.4.1. General Measurement with negligible post-measurement state,

\[
F_a := M_a^\dagger M_a, \quad (17.4.1)
\]

with

\[
F_a^\dagger = M_a^\dagger (M_a^\dagger)^\dagger \\
= M_a^\dagger M_a \\
= F_a, \quad (17.4.2)
\]

\[
F_a \geq 0, \quad (17.4.3)
\]
\[ \sum_a F_a = \text{Id}. \quad (17.4.4) \]

\( \{ F_a \mid \forall a \} \) is the set of positive operator-valued measurements, and \( \{ M_a \mid \forall a \} \) is the set of the general measurements.

The probability is determined by \( F_a \), and \( M_a \) is connected with the post measurement state, which we have no interest.

**Examples:**

1. Projective measurement is the special case of POVM.

   \[ F_a = \Pi_a \implies \begin{cases} F_a \geq 0, \\ \sum_a F_a = \text{Id}. \end{cases} \quad (17.4.5) \]

2. Construct the non-orthogonal projective measurement denoted as

   \[ \Pi_a := |\phi_a\rangle\langle\phi_a| \quad (17.4.6) \]

   with

   \[ \langle\phi_a|\phi_a\rangle = 1, \quad \langle\phi_a|\phi_b\rangle \neq \delta_{ab}. \quad (17.4.7) \]

   Construct the POVM denoted as

   \[ F_a := \lambda_a|\phi_a\rangle\langle\phi_a| = |\tilde{\phi}_a\rangle\langle\tilde{\phi}_a|, \quad (17.4.8) \]

   where \( \lambda_a \) is a real number with \( 0 < \lambda_a \leq 1 \) and

   \[ |\tilde{\phi}_a\rangle := \sqrt{\lambda_a}|\phi_a\rangle. \quad (17.4.9) \]

   Note that the measurement operator \( F_a \) satisfies the relations

   \[ \begin{cases} F_a^2 = F_a, \\ F_a \cdot F_b = 0 \text{ with } a \neq b. \end{cases} \quad (17.4.10) \]

   The probability of measurement result associated with the POVM \( F_a \) is

   \[ P_a = \text{tr}(F_a \rho) = \lambda_a \langle\phi_a|\rho|\phi_a\rangle = \langle\tilde{\phi}_a|\rho|\tilde{\phi}_a\rangle \geq 0. \quad (17.4.11) \]

   The post-measurement state is

   \[ \rho \rightarrow \rho' = \sum_a \sqrt{F_a} \rho \sqrt{F_a} = \sum_a P_a |\phi_a\rangle\langle\phi_a|, \quad (17.4.12) \]

   where

   \[ \sqrt{F_a} := \sqrt{\lambda_a}|\phi_a\rangle. \quad (17.4.13) \]

3. Construct the POVM denoted as \( \{ F_1, F_2, F_3 \} \), expressed as

   \[ \begin{cases} F_1 := \frac{2}{3} |\hat{n}_1\rangle \langle\hat{n}_1| := \frac{1}{3} (1 + \hat{n}_1 \hat{\sigma}), \\ F_2 := \frac{2}{3} |\hat{n}_2\rangle \langle\hat{n}_2| := \frac{1}{3} (1 + \hat{n}_2 \hat{\sigma}), \\ F_3 := \frac{1}{3} |\hat{n}_3\rangle \langle\hat{n}_3| := \frac{1}{3} (1 + \hat{n}_3 \hat{\sigma}). \end{cases} \quad (17.4.14a) \]

   \[ \begin{cases} F_1 := \frac{2}{3} |\hat{n}_1\rangle \langle\hat{n}_1| := \frac{1}{3} (1 + \hat{n}_1 \hat{\sigma}), \\ F_2 := \frac{2}{3} |\hat{n}_2\rangle \langle\hat{n}_2| := \frac{1}{3} (1 + \hat{n}_2 \hat{\sigma}), \\ F_3 := \frac{1}{3} |\hat{n}_3\rangle \langle\hat{n}_3| := \frac{1}{3} (1 + \hat{n}_3 \hat{\sigma}). \end{cases} \quad (17.4.14b) \]

   \[ \begin{cases} F_1 := \frac{2}{3} |\hat{n}_1\rangle \langle\hat{n}_1| := \frac{1}{3} (1 + \hat{n}_1 \hat{\sigma}), \\ F_2 := \frac{2}{3} |\hat{n}_2\rangle \langle\hat{n}_2| := \frac{1}{3} (1 + \hat{n}_2 \hat{\sigma}), \\ F_3 := \frac{1}{3} |\hat{n}_3\rangle \langle\hat{n}_3| := \frac{1}{3} (1 + \hat{n}_3 \hat{\sigma}). \end{cases} \quad (17.4.14c) \]
where \( \vec{n}_1, \vec{n}_2 \) and \( \vec{n}_3 \) are three coplanar unit vectors and the angle between any pair of them is \( \frac{2\pi}{3} \) in \( \mathbb{R}^3 \), as shown in Figure 17.1.

And we can see \( \vec{F}_i \) with \( i = 1, 2, 3 \) satisfy the relations

\[
\begin{align*}
\sum_{i=1}^{3} \vec{F}_i &= \text{Id}; \\
\vec{F}_a^2 &\neq \vec{F}_a, \quad \text{with } a = 1, 2, 3; \\
\vec{F}_a \vec{F}_b &= \frac{4}{9} \langle \vec{n}_a | \vec{n}_b \rangle \langle \vec{n}_b | \vec{n}_a \rangle \neq 0, \text{ with } a, b = 1, 2, 3.
\end{align*}
\]

We can verify these relations one by one.

a)

\[
\begin{align*}
\sum_{i=1}^{3} \vec{F}_i &= \frac{1}{3} (1 + \vec{n}_1 \hat{\sigma}) + \frac{1}{3} (1 + \vec{n}_2 \hat{\sigma}) + \frac{1}{3} (1 + \vec{n}_3 \hat{\sigma}) \\
&= \frac{1}{3} [3 + (\vec{n}_1 + \vec{n}_2 + \vec{n}_3) \hat{\sigma}] \\
&= \frac{1}{3} (3 + 0) \\
&= \text{Id}.
\end{align*}
\]

b)

\[
\begin{align*}
\vec{F}_a \vec{F}_a &= \left( \frac{2}{3} \right)^2 |\vec{n}_a\rangle \langle \vec{n}_a| \vec{n}_a| \\
&= \left( \frac{2}{3} \right)^2 |\vec{n}_a\rangle \langle \vec{n}_a| \\
&= \frac{2}{3} \vec{F}_a.
\end{align*}
\]

c)

\[
\begin{align*}
\vec{F}_a \vec{F}_b &= \frac{1}{9} (1 + \vec{n}_a \hat{\sigma}) (1 + \vec{n}_b \hat{\sigma}) \\
&= \frac{1}{9} \left[ 1 + (\vec{n}_a + \vec{n}_b) \hat{\sigma} + (\vec{n}_a \hat{\sigma})(\vec{n}_b \hat{\sigma}) \right] \\
&= \frac{1}{9} \left[ 1 + (\vec{n}_a + \vec{n}_b) \hat{\sigma} + \vec{n}_a \vec{n}_b \right],
\end{align*}
\]

Figure 17.1: Diagrammatic representation of unit vectors \( \vec{n}_1, \vec{n}_2 \) and \( \vec{n}_3 \).
therefore

\[ F_a F_b \neq 0, \quad (17.4.19) \]

since \( F_a F_b \) vanishes iff \( \bar{n}_b = -\bar{n}_a \).

(17.4.20)

\[ \text{17.5 More on POVM} \]

### 17.5.1 Dimensional analysis

For a Hilbert space \( \mathcal{H} \), the number of projective measurements should be

\[ \# \{ \Pi_n \} = \dim \mathcal{H}, \quad (17.5.1) \]

while the number of POVM should satisfy

\[ \# \{ \Phi_n \} \geq \dim \mathcal{H}. \quad (17.5.2) \]

For example, in the case of qubit, i.e., \( \# \mathcal{H}_2 = 2 \), we have

\[ \# \{ \Pi_1, \Pi_2 \} = 2, \]

\[ \# \{ \Phi_1, \Phi_2, \Phi_3 \} = 3. \quad (17.5.3) \]

### 17.5.2 POVM on subsystem can be viewed as projective measurements on the entire system

POVM can be viewed as projective measurement on a larger Hilbert space. With \( \{ \Phi_A^a \} \) being the projective measurement on the original Hilbert space \( \mathcal{H}_A \) and \( \{ \Phi_A^a \} \) being the POVM on the \( \mathcal{H}_A \). There is a larger Hilbert space \( \mathcal{H}_A' \) with \( \mathcal{H}_A \subseteq \mathcal{H}_A' \), such that \( \{ \Phi_A^a \} \) on Hilbert space \( \mathcal{H}_A \) can be viewed as projective measurements \( \{ \Phi_A' + a \} \) on \( \mathcal{H}_A' \).

\[ \# \{ \Phi_A^a \} \leq \# \{ \Phi_A^a \} \leq \# \{ \Phi_A'' + a \} . \quad (17.5.4) \]

### 17.5.3 Tensor product realization of POVM

**Thm 17.5.3.1.** Given one dimension non-negative POVM \( \{ \Phi_A^a \} \) on \( \mathcal{H}_A \), there exists \( \mathcal{H}_B \) such that

\[ \text{Prob}(a) = \text{tr}_{AB} \left[ \Phi_A^a (\rho_A \otimes \rho_B) \right] = \text{tr}_A \left( \Phi_A^a \rho_A \right), \quad (17.5.5) \]

where \( \{ \Phi_A^a \} \) are the projective measurements on \( \mathcal{H}_A \otimes \mathcal{H}_B \).

We can construct \( \{ \Phi_A^a \} \) from \( \{ \Phi_A^a \} \), namely

\[ P_a = A \quad \rho_A \quad F_A^a \quad = \quad A \quad \rho_A \quad \otimes \quad E_a \quad \rho_B \quad , \quad (17.5.6) \]

where the circles mean traces.

Let’s assume that \( \{ |i\>_A \} = \{ |j\>_A \} \) is an orthonormal basis of the Hilbert space \( \mathcal{H}_A \), and \( \{ |\mu\>_B \} = \{ |\nu\>_B \} \) is an orthonormal basis of the Hilbert space \( \mathcal{H}_B \). Then \( \{ |\mu\>_A \} = \{ |\nu\>_A \} \) is an orthonormal basis of the Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_B \).
17.5.3.1 Operator formula of $F_a$

From

$$\text{tr}_A\left( F_a^A \rho_A \right) = \text{tr}_{AB} \left[ E_{a}^{AB}(\rho_A \otimes \rho_B) \right],$$

we can see

$$F_a^A = \text{tr}_B \left[ E_{a}^{AB}(I_A \otimes \rho_B) \right].$$

(17.5.8)

17.5.3.2 Matrix formalism of $F_a$

$$\text{tr}_A \left( F_a \rho_A \right) = \sum_j \langle j | F_a | j \rangle$$

$$= \sum_{i,j} \langle j | F_a | i \rangle \langle i | \rho_A | j \rangle, \quad \text{(17.5.9)}$$

$$\text{tr}_{AB} \left[ E_{a}^{AB}(\rho_A \otimes \rho_B) \right] = \sum_{j,\nu} \langle j | E_{a}^{AB} | \nu \rangle \langle \nu | \rho_A \otimes \rho_B | j \rangle$$

$$= \sum_{j,\nu} \langle j | E_{a}^{AB} | \nu \rangle \langle \nu | \rho_A | j \rangle \langle j | \rho_B | \nu \rangle. \quad \text{(17.5.10)}$$

Therefor, we can get

$$\langle j | F_a^A | i \rangle = \sum_{j,\nu} \langle j | E_{a}^{AB} | \nu \rangle \langle \nu | \rho_B | j \rangle \langle j | \rho_B | \nu \rangle. \quad \text{(17.5.11)}$$

17.5.3.3 The properties of $F_a^A$

1) Hermitian. With E.Q. \[17.5.8\] and E.Q. \[17.5.11\], from

$$(E_{a}^{AB})^\dagger = E_{a}^{AB} \quad \text{and} \quad \rho_B^\dagger = \rho_B,$$

we can see

$$F_a^\dagger = F_a. \quad \text{(17.5.12)}$$

2) Positivity, i.e., $F_a^A \geq 0$. From the diagonal form of $\rho_B$, i.e.,

$$\rho_B = \sum_{\mu} P_{\mu} | \mu \rangle_B \langle \mu |, \quad \text{with} \sum_{\mu} P_{\mu} = 1, \text{and} \ 0 \leq P_{\mu} \leq 1,$$

we get

$$A \left( \psi | F_a^A | \psi \right)_A = A \left( \psi | \text{tr}_B \left[ E_{a}^{AB}(I_A \otimes \rho_B) \right] | \psi \right)_A$$

$$= \sum_{\mu,\nu} A \left( \psi | B \langle \nu | E_{a}^{AB} | \mu \rangle_B | \psi \right)_A \langle \mu | \rho_B | \nu \rangle_B$$

$$= \sum_{\mu,\nu} A \left( \psi | B \langle \nu | E_{a}^{AB} | \mu \rangle_B | \psi \right)_A P_{\mu} \delta_{\mu \nu}$$

$$= \sum_{\mu} A \left( \psi | B \langle \mu | E_{a}^{AB} | \mu \rangle_B | \psi \right)_A P_{\mu}. \quad \text{(17.5.14)}$$

Thus

$$A \left( \psi | F_a^A | \psi \right)_A \geq 0, \quad \text{(17.5.15)}$$

since

$$E_{a}^{AB} \geq 0, \text{ and} P_{\mu} \geq 0.$$
3) Completeness.

\[ \sum_a F_a^A = \sum_a \text{tr}_B \left[ E_{a}^{AB}(I_A \otimes \rho_B) \right] \]
\[ = \text{tr}_B \left[ \sum_a E_{a}^{AB}(I_A \otimes \rho_B) \right] \]
\[ = \text{tr}_B( I_A \otimes \rho_B) \]
\[ = I_A \otimes \text{tr}_B \rho_B \]
\[ = I_A. \quad (17.5.16) \]

17.5.3.4 Example

We are going to talk about an example that has been discussed at the beginning of section 17.4. The POVM denoted as \{F_a^A \mid a = 1, 2, 3\}, have the explicit expression

\[ F_a^A = \frac{2}{3} |\tilde{n}_a\rangle \langle \tilde{n}_a|, \quad \text{with} \quad a = 1, 2, 3. \quad (17.5.17) \]

And the vectors \{\tilde{n}_a \mid a = 1, 2, 3\} are shown in Figure 17.1, from which we can see

\[ \tilde{n}_1 + \tilde{n}_2 + \tilde{n}_3 = 0. \quad (17.5.18) \]

Now, we introduce an auxiliary Hilbert space \( \mathcal{H}_B \) and the density matrix \( \rho_B \) on \( \mathcal{H}_B \) is defined as

\[ \rho_B := |0\rangle_B \langle 0| \quad (17.5.19) \]

Therefore the density matrix of the compound Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_B \) is

\[ \rho_{AB} := \rho_A \otimes \rho_B \]
\[ = \rho_A \otimes |0\rangle_B \langle 0|. \quad (17.5.20) \]

We can construct the two-qubit state as

\[ |\Phi_a\rangle_{AB} := \sqrt{\frac{2}{3}} |\tilde{n}_a\rangle_A |0\rangle_B + \sqrt{\frac{1}{3}} |0\rangle_A |1\rangle_B, \quad \text{with} \quad a = 1, 2, 3. \quad (17.5.21) \]

It can be easily verified that \{ |\Phi_a\rangle_{AB} \mid a = 1, 2, 3\} makes an orthonormal basis for the compound Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_B \).

\[ A\langle \tilde{n}_a' | \tilde{n}_a \rangle_A = \begin{cases} \frac{1}{2}, & \text{if } a \neq a'; \\ 1, & \text{if } a = a'. \end{cases} \quad (17.5.22) \]

With another two-qubit state defined as

\[ |\Phi_0\rangle_{AB} := |1\rangle_A \otimes |1\rangle_B, \quad (17.5.23) \]
we can get
\[\langle \Phi_{a'} | \Phi_a \rangle = \delta_{aa'}, \quad \forall a, a' \in \{0, 1, 2, 3\}. \quad (17.5.24)\]

With the orthonormal basis \(| \Phi_a \rangle = \{\Phi_a\}_{a = 0, 1, 2, 3}\), we can construct the projectors in the following manner

\[
E_a^{AB} = |\Phi_a\rangle_{AB} \langle \Phi_a| = \left(\sqrt{\frac{2}{3}} |\tilde{n}_a\rangle_A \langle 0|_B + \sqrt{\frac{1}{3}} |0\rangle_A |1|_B\right)\left(\sqrt{\frac{2}{3}} A |\tilde{n}_a|_B \langle 0| + \sqrt{\frac{1}{3}} |0\rangle_B \langle 1|\right)
\]

\[
= \frac{2}{3} |\tilde{n}_a\rangle_A |0\rangle_B A |\tilde{n}_a|_B \langle 0| + \frac{1}{3} |0\rangle_A |1|_B A |\tilde{n}_a|_B \langle 0| + \sqrt{\frac{2}{3}} |\tilde{n}_a\rangle_A |0\rangle_B \langle 0| + \sqrt{\frac{2}{3}} |0\rangle_A |1|_B A |\tilde{n}_a|_B \langle 0|,
\]

namely

\[
E_a^{AB} = \frac{2}{3} |\tilde{n}_a\rangle_A |\tilde{n}_a| \otimes |0\rangle_B \langle 0| + \frac{1}{3} |0\rangle_A |0| \otimes |1|_B \langle 1|
\]

\[
+ \frac{\sqrt{2}}{3} |\tilde{n}_a\rangle_A |0\rangle \otimes |0\rangle_B (1) + \frac{\sqrt{2}}{3} |0\rangle_A |1|_B A |\tilde{n}_a|_B \langle 0|,
\]

for \(a = 1, 2, 3\). And \(E_0^{AB}\) on the other hand, is defined as

\[
E_0^{AB} := |\Phi_0\rangle_{AB} \langle \Phi_0| = |1\rangle_A \langle 1| \otimes |0\rangle_B (1).
\]

Therefore,

\[
E_a^{AB} (I_A \otimes \rho_B)
\]

\[
= E_a^{AB} (|0\rangle_B \langle 0|)
\]

\[
= \frac{2}{3} |\tilde{n}_a\rangle_A \langle \tilde{n}_a| \otimes |0\rangle_B (0) + \frac{\sqrt{2}}{3} |0\rangle_A \langle \tilde{n}_a| \otimes |1|_B (0),
\]

for \(a = 1, 2, 3\), and

\[
E_a^{AB} (I_A \otimes \rho_B) = 0.
\]

Thus,

\[
F_a^A = \text{tr}_B \left[E_a^{AB} (I_A \otimes \rho_B)\right] = \frac{2}{3} |\tilde{n}_a\rangle_A \langle \tilde{n}_a|,
\]

for \(a = 1, 2, 3\), and

\[
F_0^A = 0.
\]

Therefore, the relation \[17.5.3] has been verified.
17.5.4 Direct-sum realization of POVM

**Thm 17.5.4.1** (Neumark’s theorem). Given one-dimensional non-negative POVM \( \{ F^A_a \} \). It can be realized by extending \( \mathcal{H}_A \) to \( \mathcal{H}'_A \), \( \mathcal{H}_A \subseteq \mathcal{H}'_A \), such that \( \mathcal{H}_A = \mathcal{H}_A \oplus \mathcal{H}'_A \), where \( \mathcal{H}'_A \) is orthogonal to \( \mathcal{H}_A \).

**Proof.** See Preskill’s lecture notes, Chapter 3.1.4. \( \square \)

17.5.5 POVM as quantum operation (superoperator)

**Thm 17.5.5.1.** Given one-dimensional non-negative POVM \( \{ F^A_a \} \) on \( \mathcal{H}_A \). It can be realized by a unitary transformation on \( \mathcal{H}_A \otimes \mathcal{H}_B \) denoted as \( U_{AB} \) followed with an orthogonal measurement \( \Pi^B_\mu \) on \( \mathcal{H}_B \), namely

\[
\text{tr}_A(\rho_A F^A_a) = \text{tr}_{AB}(\Pi^B_\mu \rho_{AB}^\prime),
\]

where

\[
\rho'_{AB} = U_{AB} \rho_{AB} U_{AB}^\dagger.
\]

**Proof.** The basis of the Hilbert space \( \mathcal{H}_B \) is denoted as \( \{ |\mu\rangle_B \} = \{ |\nu\rangle_B \} \).

\[
\text{tr}_B(\Pi^B_\mu \rho'_{AB}) = \sum_\mu B(\mu|\Pi^B_\mu |\rho'_{AB}|\mu)_B
\]

\[
= \sum_{\mu,\nu} B(\mu|\Pi^B_\mu |\nu \rangle_B \langle \nu |\rho'_{AB}|\mu\rangle_B)
\]

\[
= \sum_\nu \delta_{\mu\nu} B(\nu |\rho'_{AB}|\mu\rangle_B
\]

\[
= B(\mu |\rho'_{AB}|\mu\rangle_B
\]

\[
= B(\mu |U_{AB}(\rho_A \otimes |0\rangle_B \langle 0| U_{AB}^\dagger |\mu\rangle_B)
\]

\[
= B(\mu |U_{AB}|0\rangle_B (|\rho_A\rangle_B |0\rangle_{AB} U_{AB}^\dagger |\mu\rangle_B)
\]

\[
= M_\mu \rho_A M_\mu^\dagger,
\]

where

\[
M_\mu := B(\mu |U_{AB}|0\rangle_B.
\]

\[
\text{tr}_{AB}(\Pi^B_\mu \rho'_{AB}) = \text{tr}_A(M_\mu \rho_A M_\mu^\dagger)
\]

\[
= \text{tr}_A(F_A \rho_A),
\]

where the POVM is defined as

\[
F^A_a := M_\mu M_\mu.
\]

\[
\square
\]

181
Check: Due to $M_\mu^\dagger M_\mu \geq 0$, $F_\mu \geq 0$.

$$\sum_\mu B(0)|U_{AB}^\dagger|\mu B(\mu|U_{AB}|0)_B = B(0)|U_{AB}^\dagger U_{AB}|0)_B = B(0|I_{AB}|0)_B = I_A,$$  \hspace{1cm} (17.5.39)

therefore

$$\sum_\mu F_\mu = Id \iff \sum_\mu M_\mu^\dagger M_\mu = Id.$$  \hspace{1cm} (17.5.40)

Remark: Quantum operation.

$$\rho_\Lambda \xrightarrow{S} \rho'_\Lambda = S(\rho_\Lambda) = \sum_\mu M_\mu \rho_\Lambda M_\mu^\dagger,$$  \hspace{1cm} (17.5.41)

where

$$\sum_\mu M_\mu^\dagger M_\mu = Id.$$  \hspace{1cm} (17.5.42)

It is called the Kraus representation of quantum operation, and details see following section \[7.6.4\].

17.5.6 Realization of a POVM

Problem description\[7\]

Consider the POVM defined by the four positive operators

$$P_1 = \frac{1}{2} |\uparrow_z\rangle\langle\uparrow_z|, \quad P_2 = \frac{1}{2} |\downarrow_z\rangle\langle\downarrow_z|,$$

$$P_3 = \frac{1}{2} |\uparrow_x\rangle\langle\uparrow_x|, \quad P_4 = \frac{1}{2} |\downarrow_x\rangle\langle\downarrow_x|.$$  \hspace{1cm} (17.5.43)

Show how this POVM can be realized as an orthogonal measurement in a two-qubit Hilbert space, if one ancilla spin is introduced.

For convenience, we apply the notations

$$|0\rangle = |\uparrow_z\rangle, \quad |1\rangle = |\downarrow_z\rangle, \quad |+\rangle = |\uparrow_x\rangle, \quad |\rangle = |\downarrow_x\rangle,$$  \hspace{1cm} (17.5.44)

in the following discussions.

1° Check $\{P_i|i = 1, 2, 3, 4\}$ indeed as a POVM.

$$P_i^2 = \frac{1}{4} P_i$$  \hspace{1cm} (17.5.45)

implies that $P_i$ is not a projector.

With the new notation

$$P_k := \frac{1}{2} |k\rangle\langle k|,$$  \hspace{1cm} (17.5.46)

where $k = 0, 1, +, -$. For $\forall |\psi\rangle$, we have

$$\langle\psi|P_k|\psi\rangle = \frac{1}{2} |\langle\psi|k\rangle|^2 \geq 0.$$  \hspace{1cm} (17.5.47)

\[7\] Originated from the exercise 3.1, Chapter 3 of John Preskill’s online lecture notes.
\(P_i\) sum to the identity, namely

\[
P_1 + P_2 + P_3 + P_4 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1| + |+\rangle\langle +| + |-\rangle\langle -|)
= \frac{1}{2}I_2 + \frac{1}{2}I_2
= I_2.
\] (17.5.48)

2° Construct the projective measurement \(\{E_i|i = 1, 2, 3, 4\}\).

\[
E_1 = \frac{1}{2}|00\rangle\langle 00|, \quad E_2 = \frac{1}{2}|10\rangle\langle 10|,
E_3 = \frac{1}{2}|+1\rangle\langle +1|, \quad E_4 = \frac{1}{2}|-1\rangle\langle -1|.
\] (17.5.49)

Check \(\{E_i|i = 1, 2, 3, 4\}\) as projective measurement.

\[
E_i^2 = E_i; 
\]
(17.5.50)

\[
E_i^\dagger = E_i; 
\]
(17.5.51)

\[
E_i E_j = 0, \quad i \neq j; 
\]
(17.5.52)

\[
E_1 + E_2 + E_3 + E_4 = |00\rangle\langle 00| + |10\rangle\langle 10| + |+1\rangle\langle +1| + |-1\rangle\langle -1| = I_4. 
\] (17.5.53)

3° Realize POVM \(P_i\) on Hilbert space \(\mathcal{H}_A\) by projective measurement \(E_i\) on Hilbert space \(\mathcal{H}_A \otimes \mathcal{H}_B\).

In Hilbert space \(\mathcal{H}_A \otimes \mathcal{H}_B\), construct the density matrix \(\rho_{AB}\) as

\[
\rho_{AB} := \rho_A \otimes \frac{1}{2}I_B.
\] (17.5.54)

Check the relation \(\text{tr}_{AB}(E_i \rho_{AB}) = \text{tr}_A(P_i \rho_A)\) one by one.

\[
\text{tr}_{AB}(E_1 \rho_{AB}) = AB\langle 00|\rho_{AB}|00\rangle_{AB}
= AB\langle 00|\rho_A \otimes \frac{1}{2}I_B|00\rangle_{AB}
= \frac{1}{2}A\langle 0|\rho_A|0\rangle_A
= \text{tr}_A\left(\frac{1}{2}\langle 0|\rho_A\rangle\right)
= \text{tr}_A(P_1 \rho_A); 
\]
(17.5.55)

\[
\text{tr}_{AB}(E_2 \rho_{AB}) = AB\langle 10|\rho_A \otimes \frac{1}{2}I_B|10\rangle_{AB} = \text{tr}_A(P_2 \rho_A); 
\]
(17.5.56)

\[
\text{tr}_{AB}(E_{3,4} \rho_{AB}) = \frac{1}{2}A\langle \pm|\rho_A|\pm\rangle_A = \text{tr}_A(P_{3,4} \rho_A). 
\]
(17.5.57)
17.6 Quantum operation (superoperator)

17.6.1 Definition of the superoperator

Question: What’s the time evolution equation of density matrix?

As we know, for the pure state $|\psi\rangle$, $\rho = |\psi\rangle\langle\psi|$, $|\psi\rangle$ satisfies the Shr¨ödinger Equation, which implies it will evolve under unitary transformation, i.e.,

$$|\psi\rangle \xrightarrow{U} U |\psi\rangle,$$

$$\rho := |\psi\rangle\langle\psi| \xrightarrow{U} U |\psi\rangle\langle\psi| U^\dagger = U \rho U^\dagger. \quad (17.6.1)$$

In the case of mixed state described by $\rho$, we can assume $\rho \rightarrow \rho' = S(\rho)$, where $S$ is called “superoperator”. We are now going to explore some properties of the superoperator $S$.

As for sure both $\rho$ and $\rho'$ are density matrices, therefore $S$ is a mapping between density matrices. The supperoperator has the following constrains:

- a linear mapping on density matrices, i.e.,
  $$S(\lambda_1 \rho_1 + \lambda_2 \rho_2) = \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2), \quad \text{with } \lambda_1, \lambda_2 \in \mathbb{C}; \quad (17.6.4)$$

- trace-preserving mapping (TP), namely
  $$\text{tr}[S(\rho)] = \text{tr}(\rho) = 1; \quad (17.6.5)$$

- positive mapping (P), i.e.,
  $$S(\rho) \geq 0. \quad (17.6.6)$$

Therefore, $S$ should be a linear positive trace-preserving operator.

17.6.2 The wonderful theorem

The following statements are equivalent.

1) $S(\rho)$ is a CPTP mapping (Complete-Positive-trace-preserving).

$$\begin{cases} S(\rho) \geq 0, \quad \text{Positive}; \\ (S \otimes \text{Id})\rho_{AB} \geq 0, \quad \text{Complete Positive.} \end{cases} \quad (17.6.7)$$

2) The Kraus representation theorem.

**Thm 17.6.2.1 (Kraus representation).** A CPTP mapping $S(\rho)$ has the operator sum representation,

$$S(\rho) = \sum_{\mu} M_\mu \rho M_\mu^\dagger, \quad (17.6.8)$$

with the Kraus operators $M_\mu$ satisfying

$$\sum_{\mu} M_\mu^\dagger M_\mu = \text{Id}. \quad (17.6.9)$$
3) The Stinespring representation

\[
S(\rho_A) = \text{tr}_B \left[ U_{AB}(\rho_A \otimes |0\rangle_B \langle 0|) U_{AB}^\dagger \right] 
\]

(17.6.10)

with the diagram representation

\[
\begin{array}{c}
\text{A:} \quad \rho_A \quad \text{SSS} \\
\text{B:} \quad |0\rangle_B \langle 0| \\
\end{array}
\]

\[
\xrightarrow{\text{U}_{AB}} \\
S(\rho_A) 
\]

(17.6.11)

The symbol \(\text{recycle bin}\) stands for the recycle bin, which means that the corresponding output qubit is out of our interest.

Math tool: Function Analysis.

4) The Choi-Jamiolkowski representation

\[
\rho'_{AB} := (S \otimes \text{Id})(|\Psi\rangle_{AB}(\langle \Psi|), 
\]

(17.6.12)

where \(|\Psi\rangle_{AB}\) is the maximal bipartite entangled state defined as

\[
|\Psi\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_i |ii\rangle_{AB} 
\]

(17.6.13)

with \(N\) as the dimension of the Hilbert space \(\mathcal{H}\). The superoperator can be uniquely determined in the way

\[
S(\rho_A) = N\text{tr}_B \left[ \rho'_{AB}(I_A \otimes \rho_A^T) \right], 
\]

(17.6.14)

where \(\rho_A^T\) is defined as the transposition of the density matrix \(\rho_A\).

17.6.3 The CPTP mapping

\(S(\rho_A)\) is positive, and \((S \otimes I_B)\rho_{AB}\) is also positive.

Remarks:

- Separable system \(\rho_{AB} = \rho_A \otimes \rho_B\).

\[
\begin{array}{c}
\rho_A \\
\xrightarrow{S} \\
S(\rho_A) \geq 0 \\
\rho_B \\
\end{array}
\]

(17.6.15)

- Entangled state.

\[
\begin{array}{c}
\rho_{AB} \\
\xrightarrow{S} \\
S(\rho_A) \geq 0 \\
\rho_B \\
\end{array}
\]

(17.6.16)

The CPTP is required for the entangled system after the superoperator acting on the subsystem to ensure the validity of the density matrix.

- Difference between CPTP mapping and PTP mapping. For example, transpose is PTP mapping, but not CPTP mapping. We can see that

\[
T: \quad \rho \geq 0 \quad \Rightarrow \quad \rho^T \geq 0, 
\]

(17.6.17)

185
i.e., transpose is PTP mapping. On the other hand, we can consider a density matrix \( \rho_{AB} \),

\[
\rho_{AB} = \frac{1}{N} \sum_{i,j} |ii\rangle_{AB} \langle jj|.
\]

(17.6.18)

If we denote the transpose as \( T \), then \( T \otimes I_B \) means the partial transpose acting on subsystem A.

\[
\rho'_{AB} = (T \otimes I_B) \rho_{AB} = \frac{1}{N} \sum_{i,j} (T \otimes I_B) |ii\rangle_{AB} \langle jj|. 
\]

(17.6.19)

As we know that \((\text{SWAP})^2 = \text{Id} \Rightarrow (\text{SWAP} - \text{Id})(\text{SWAP} + \text{Id}) = 0\),

(17.6.20)

thus \( \text{SWAP} \) has eigenvalues 1 and \(-1\). It suggests \( \text{SWAP} \) is not positive, namely \( \rho'_{AB} = (T \otimes I_B) \rho_{AB} \) is not positive. Therefore, \( T \otimes I_B \) is not CPTP mapping.

17.6.4 The Kraus representation

\[
\rho' := S(\rho) = \sum_{\mu} M_\mu \rho M_\mu^\dagger,
\]

(17.6.21)

with

\[
\sum_{\mu} M_\mu M_\mu^\dagger = \text{Id}.
\]

(17.6.22)

Examples:

(i) Unitary evolution of closed system,

\[
\rho \rightarrow U \rho U^\dagger,
\]

(17.6.23)

with

\[
\begin{cases} 
M_1 = U, & M_1^\dagger = U^\dagger; \\
M_2 = 0, & M_2^\dagger = 0.
\end{cases}
\]

(17.6.24)

It can be easily verified that

\[
M_1^\dagger M_1 + M_2^\dagger M_2 = \text{Id}.
\]

(17.6.25)

(ii) Projective measurement.

\[
\Pi_n = |n\rangle \langle n|; \quad \sum_n \Pi_n = \text{Id}.
\]

(17.6.26)

\[
\rho := |\psi\rangle \langle \psi| \xrightarrow{\Pi_n} \rho' = S(\rho) = \sum_n \Pi_n \rho \Pi_n^\dagger.
\]

(17.6.27)

(iii) POVM with one-dimensional operators \( \{F_a\} \).

\[
\rho \rightarrow \rho' = \sum_a \sqrt{F_a} \rho \sqrt{F_a}, \quad \text{with } F_a \geq 0, \text{ and } \sum_a F_a = 1.
\]

(17.6.28)
17.6.5 The Stinespring representation

\[ \begin{align*}
A &: \quad \rho_A - \xrightarrow{\rho_{AB}} - U_{AB} - S(\rho_A)
\end{align*} \]

\[ (17.6.29) \]

The Stinespring representation is actually equivalent with Kraus representation.

\[ \rho_{AB} := \rho_A \otimes |0\rangle_B \langle 0| \quad \Rightarrow \quad \rho_{AB}' := U_{AB} \rho_{AB} U_{AB}^\dagger, \]

\[ (17.6.30) \]

therefore

\[ \begin{align*}
\rho_A' &= \text{tr}_B \rho_{AB}' \\
&= \text{tr}_B \left( U_{AB} \rho_{AB} U_{AB}^\dagger \right) \\
&= \sum_\mu B \{ \mu | U_{AB} \rho_{AB} U_{AB}^\dagger | \mu \}_B \\
&= \sum_\mu B \{ \mu | U_{AB} (\rho_A \otimes |0\rangle_B \langle 0|) U_{AB}^\dagger | \mu \}_B \\
&= \sum_\mu B \{ \mu | U_{AB} |0\rangle_B \langle 0| U_{AB}^\dagger | \mu \}_B \\
&= \sum_\mu M_\mu \rho_A M_\mu^\dagger,
\end{align*} \]

\[ (17.6.31) \]

where

\[ M_\mu = B \{ \mu | U_{AB} |0\rangle_B \}, \quad M_\mu^\dagger = B \{ 0 | U_{AB}^\dagger | \mu \}_B. \]

\[ (17.6.32) \]

Then

\[ \begin{align*}
\sum_\mu M_\mu^\dagger M_\mu &= \sum_\mu B \{ 0 | U_{AB}^\dagger | \mu \}_B \langle \mu | U_{AB} |0\rangle_B \\
&= B \langle 0 | U_{AB}^\dagger U_{AB} |0\rangle_B \\
&= I_A.
\end{align*} \]

\[ (17.6.33) \]

17.6.6 Remarks on quantum operation

Remarks:

1° The Kraus representation is not unique and is base-dependent. In the construction of POVM as quantum operation, it has the Kraus representation

\[ F^A_a := M_\mu^\dagger M_\mu, \]

\[ (17.6.34) \]

with

\[ M_\mu = B \{ \mu | U_{AB} |0\rangle_B \}. \]

\[ (17.6.35) \]

With the new basis defined as

\[ B \{ \nu \} = \sum_\mu U_{\nu \mu} B \{ \mu \} = \sum_\mu U_{\nu \mu} M_\mu, \]

\[ (17.6.36) \]

the superoperator has the Kraus representation

\[ S(\rho_A) = \sum_\nu N_\nu \rho_A N_\nu^\dagger, \]

\[ (17.6.37) \]
where
\[ N_\nu = B(\nu|U_{AB}|0)_B = \sum_\mu U_{\nu\mu} M_\mu. \] (17.6.38)

The Kraus representations of the superoperator $S$, namely $M_\mu$ and $N_\nu$, are equivalent, i.e., \( \{M_\mu\} \cong \{N_\nu\} \).

2° The dimension of Hilbert space $\mathcal{H}_A$ is $N$, namely $\dim \mathcal{H}_A = N$. Kraus operator as the mapping from Hilbert space $\mathcal{H}_A$ to $\mathcal{H}_A$, there at most $N^2$ Kraus operators, i.e., $\#\{M_\mu\} = N^2$.

3°
\[ \rho \xrightarrow{S} \rho' = S(\rho). \] (17.6.39)

In general, $S^{-1}(\rho')$ dose not exist! Open system changes usually for decoherence or information loss.

**Note:** In closed system, $S^{-1}(\rho')$ exists, since
\[ \rho' = U\rho U\dagger, \quad \rho = U\dagger\rho' U. \] (17.6.40)

### 17.7 Quantum channel

Example for Quantum Operation:
\[ \rho \xrightarrow{S} \rho' = S(\rho), \] (17.7.1)
where the superoperator $S$ is also called a quantum channel.

#### 17.7.1 The bit-flip channel

The quantum bit-flip channel:
\[ |0\rangle \xrightarrow{X} |1\rangle, \quad |1\rangle \xrightarrow{X} |0\rangle, \quad \text{with} \quad X = \sigma_x. \] (17.7.2)

It stands for the quantum error-model:
\[ \begin{array}{c c c}
|0\rangle & \xrightarrow{p} & |1\rangle \\
1 - p & \xrightarrow{X} & 1 - p
\end{array} \quad \begin{array}{c c c}
|1\rangle & \xrightarrow{p} & |0\rangle \\
1 - p & \xrightarrow{X} & 1 - p
\end{array}, \] (17.7.3)

where $p$ is the error-probability, $0 \leq p \leq 1$, namely

\[
\begin{cases}
S|0\rangle \{0\} = (1 - p)|0\rangle \{0\} + p|1\rangle \{1\}; \\
S|1\rangle \{1\} = p|0\rangle \{0\} + (1 - p)|1\rangle \{1\}.
\end{cases}
\] (17.7.4a, 17.7.4b)

Therefore
\[ \rho \rightarrow \rho' := S(\rho) = (1 - p)\rho + pX\rho X \] (17.7.5)

We can then find the Kraus operators
\[
\begin{cases}
M_0 = \sqrt{1 - p}I = \sqrt{1 - p} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \\
M_1 = \sqrt{p}X = \sqrt{p} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
\end{cases}
\] (17.7.6)
and verify that

\[
M_0^\dagger M_0 + M_1^\dagger M_1 = (1 - p)(I_2)^2 + pX^2 = (1 - p)I_2 + pI_2 = I_2.
\]  

(17.7.7)

1) Unitary Representation (Stinespring representation).

\[
\begin{array}{ccc}
A: & |\psi_A\rangle & U_{AB} \rightarrow S(|\psi_A\rangle \langle \psi|) \\
B: & |0_B\rangle & \\
\end{array}
\]

(17.7.8)

where

\[
U_{AB}(|\psi_A\rangle \otimes |0_B\rangle) := (M_0 \otimes I_2 + iM_1 \otimes X)|\psi_A\rangle \otimes |0_B\rangle
= \sqrt{1 - p}|\psi_A\rangle \otimes |0_B\rangle + i\sqrt{p}X|\psi_A\rangle \otimes |1_B\rangle.
\]  

(17.7.9)

Then

\[
\rho_A = S(\rho_A),
\]

\[
= \text{tr}_B \left( U_{AB} (|\psi_A\rangle \langle \psi| \otimes |0_B\rangle \langle 0|) U_{AB}^\dagger \right)
= \text{tr}_B \left( (\sqrt{1 - p}|\psi_A\rangle \otimes |0_B\rangle + i\sqrt{p}X|\psi_A\rangle \otimes |1_B\rangle) \right.
\]

\[
\left. (\sqrt{1 - p}|\psi_A\rangle \otimes |0_B\rangle - i\sqrt{p}A|\psi|X \otimes 1) \langle 1| \right)
= \text{tr}_B \left[ (1 - p)|\psi_A\rangle \langle \psi| \otimes |0_B\rangle \langle 0| + pX |\psi_A\rangle \langle \psi| X \otimes |1_B\rangle \langle 1| \right]
+i\sqrt{p(1 - p)}\text{tr}_B \left[ -|\psi_A\rangle \langle \psi| X \otimes |0_B\rangle \langle 1| + X |\psi_A\rangle \langle \psi| \otimes |1_B\rangle \langle 0| \right]
= (1 - p)|\psi_A\rangle \langle \psi| + pX |\psi_A\rangle \langle \psi| X,
\]  

(17.7.10)

and

\[
U_{AB}^\dagger U_{AB} = (M_0 \otimes I_2 + iM_1 \otimes X)(M_0^\dagger \otimes I_2 + iM_1^\dagger \otimes X)
= M_0^\dagger M_0 \otimes I_2 + M_1^\dagger M_1 \otimes X^2 - i(M_1^\dagger M_0 \otimes X - M_0^\dagger M_1 \otimes X)
= (M_0^\dagger M_0 + M_1^\dagger M_1) \otimes I_2 - i(M_1M_0 \otimes X - M_0M_1 \otimes X)
= I_2 \otimes I_2 - i(M_1M_0 \otimes X - M_0M_1 \otimes X)
= I_4.
\]  

(17.7.11)

where we have applied the relations of the Kraus operators expressed as

\[
\begin{align*}
M_0^\dagger &= M_0, \\
M_1^\dagger &= M_1,
\end{align*}
\]  

(17.7.12a)

(17.7.12b)

as is shown in E.Q. (17.7.6). Thus, we can also verify

\[
U_{AB} U_{AB}^\dagger = I_4.
\]  

(17.7.13)

Therefore,

\[
U_{AB} := M_0 \otimes I_2 + iM_1 \otimes X
\]  

(17.7.14)

is unitary.
2) The Bloch sphere representation.

\[ \rho = \frac{1}{2} (\mathbf{I} + p_3 \mathbf{Z}) \]  

(17.7.15)

with

\[ |p_3| \leq 1, \quad \hat{p} = p_3 \hat{e}_3, \]  

(17.7.16)

and

\[ Z = \sigma_3. \]  

(17.7.17)

Therefore,

\[ S(\rho) = (1 - p)\rho + pX\rho X \]

\[ = \frac{1}{2} (\mathbf{I}d + p'_3 Z), \]  

(17.7.18)

with

\[ p'_3 := (1 - 2p)p_3, \]  

(17.7.19)

since

\[ (1 - p)\rho + pX\rho X = \frac{1 - p}{2} (\mathbf{I}d + p_3 \mathbf{Z}) + \frac{p}{2} X (\mathbf{I}d + p_3 \mathbf{Z}) X \]

\[ = \frac{1 - p}{2} (\mathbf{I}d + p_3 \mathbf{Z}) + \frac{p}{2} (\mathbf{I}d - p_3 \mathbf{Z}) \]

\[ = \frac{1}{2} [\mathbf{I}d + (1 - 2p)p_3 \mathbf{Z}]. \]  

(17.7.20)

Qubit-flip error model means contraction in the Bloch ball instead of inflation, i.e., \( p'_3 \leq p_3 \). We may conclude here that the inflation is not permitted here in Quantum Mechanics.

### 17.7.2 The phase-flip channel

Next, we are going to discuss the phase-flip error, which exists in Quantum Mechanics and has no classical analogy.

Phase-flip error:

\[ |0\rangle \rightarrow |0\rangle, \quad |1\rangle \rightarrow -|1\rangle. \]  

(17.7.21)

Therefore, the phase-flip channel should have the form

\[ S(\rho) = (1 - p)\rho + pZ\rho Z. \]  

(17.7.22)

And we can define the bit-phase-flip error as

\[ S(\rho) = (1 - p)\rho + pY\rho Y, \quad \text{with} \quad Y = \sigma_y. \]  

(17.7.23)

### 17.7.3 Depolarizing channel

Depolarization suggests the process

\[ \text{Polarization}(\rho) \rightarrow \text{Non-polarization}(\rho) \]  

(17.7.24)

with the formalism

\[ S(\rho) = (1 - p)\rho + \frac{p}{3} (X\rho X + Y\rho Y + Z\rho Z), \]  

(17.7.25)
and the Kraus operators defined as
\[
M_0 := \sqrt{1-p} I_2, \quad M_1 := \sqrt{\frac{p}{3}} X, \quad M_2 := \sqrt{\frac{p}{3}} Y, \quad M_3 := \sqrt{\frac{p}{3}} Z. \tag{17.7.26}
\]
We can verify that
\[
\sum_{i=0}^{3} M_i^\dagger M_i = M_0^\dagger M_0 + M_1^\dagger M_1 + M_2^\dagger M_2 + M_3^\dagger M_3
\]
\[
= (1-p) I_2 + \frac{p}{3} I_2 + \frac{p}{3} I_2 + \frac{p}{3} I_2
\]
\[
= I_2.
\]

The Bloch Sphere representation.
\[
\rho = \frac{1}{2} (\text{Id} + \vec{p} \cdot \vec{\sigma}), \quad \text{with } |\vec{p}| \leq 1 \text{ and } \vec{p} \in \mathbb{R}^3. \tag{17.7.27}
\]
Therefore,
\[
S(\rho) = \frac{1}{2} \left[ S(1 + \vec{p} \cdot \vec{\sigma}) \right]
\]
\[
= \frac{1}{2} \left[ S(I_2) + S(\vec{p} \cdot \vec{\sigma}) \right]
\]
\[
= \frac{1}{2} \left[ I_2 + (1-p) \vec{p} \cdot \vec{\sigma} + \frac{p}{3} (X \vec{p} \cdot \vec{\sigma} X + Y \vec{p} \cdot \vec{\sigma} Y + Z \vec{p} \cdot \vec{\sigma} Z) \right]
\]
\[
= \frac{1}{2} \left[ I_2 + (1-p) \vec{p} \cdot \vec{\sigma} + \frac{p}{3} \left( (p_1 - p_1 - p_1) X + (-p_2 + p_2 - p_2) Y + (-p_3 - p_3 + p_3) Z \right) \right]
\]
\[
= \frac{1}{2} \left[ I_2 + (1-p) \vec{p} \cdot \vec{\sigma} - \frac{2p}{3} \vec{p} \cdot \vec{\sigma} \right]
\]
\[
= \frac{1}{2} \left( I_2 + \vec{p}' \cdot \vec{\sigma} \right), \tag{17.7.28}
\]
with
\[
\vec{p}' := \left( 1 - \frac{4p}{3} \right) \vec{p}. \tag{17.7.29}
\]
Depolarization is also a contraction process, due to $|\vec{p}'| < |\vec{p}|$. When $p = \frac{3}{4}$,
\[
S(\rho) = \frac{1}{2} I_2, \tag{17.7.30}
\]
indicated no-polarization at all.

17.7.4 The phase-damping channel
\[
\rho \xrightarrow{S} S(\rho) = \sum_{i=0}^{2} M_i \rho M_i^\dagger, \tag{17.7.31}
\]
with
\[
M_0 := \sqrt{1-p} I_2, \quad M_1 := \sqrt{p} |0\rangle \langle 0|, \quad M_2 := \sqrt{p} |1\rangle \langle 1|. \tag{17.7.32}
\]
Therefore, with the matrix expression of the density matrix $\rho$,
\[
\rho := \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}, \tag{17.7.33}
\]
191
we can get
\[ S(\rho) = (1 - p)\rho + p \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix} = \begin{pmatrix} \rho_{00} & (1 - p)\rho_{01} \\ (1 - p)\rho_{10} & \rho_{11} \end{pmatrix}. \] (17.7.34)

As we shall see that from \( \rho \) to \( S(\rho) \),

- the diagonal terms are unchanged;
- the non-diagonal terms decay with the factor \((1 - p)\).

Therefore, in the manner shown as
\[ S^n(\rho) := S^nS^nS^n\ldots S^n\rho \]

with the new parameters defined as
\[ p = \Gamma \Delta t, \quad t = n\Delta t, \] (17.7.36)

which have the formalism
\[ (1 - p)^n = (1 - \Gamma \Delta t)^{t/\Delta t} \overset{\Delta t \to 0}{\longrightarrow} e^{-\Gamma t}, \] (17.7.37)

we can find immediately that with \( n \to \infty \), namely \( t \to \infty \),
\[ S^n(\rho) = \begin{pmatrix} \rho_{00} & e^{-\Gamma t} \\ e^{-\Gamma t} & \rho_{11} \end{pmatrix} \overset{t \to \infty}{\longrightarrow} \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix}, \] (17.7.38)

i.e., the non-diagonal terms vanish or the coherent terms vanish. The geometric picture can be described as the Bloch ball contracted into a line on the poles, namely the contraction has the preferred bases \( |0\rangle \) and \( |1\rangle \).

**Note:** In the Kraus operators of phase-damping channel \( \{ M_i | i = 1, 2, 3 \} \), we can rewrite the operators as
\[ M_0 := \sqrt{1 - p}I_2, \quad M_1 := \frac{\sqrt{p}}{2}(I_2 + Z), \quad M_2 := \frac{\sqrt{p}}{2}(I_2 - Z). \] (17.7.39)

Then
\[ S(\rho) = \sum_{i=0}^{2} M_i \rho M_i^\dagger \]
\[ = (1 - p)\rho + \frac{p}{4}(I_2 + Z)\rho(I_2 + Z) + \frac{p}{4}(I_2 - Z)\rho(I_2 - Z) \]
\[ = (1 - \frac{1}{2}p)\rho + \frac{1}{2}pZ\rho Z. \] (17.7.40)

And now we can redefine the Karus formalism of phase-damping channel as
\[ N_0 := \sqrt{1 - \frac{1}{2}p}I_2, \quad N_1 := \sqrt{\frac{1}{2}p}Z. \] (17.7.41)

The Karus operator is not unique, \( \{ N_0, N_1 \} \cong \{ M_0, M_1, M_2 \} \).
17.7.5 The amplitude-damping channel

\[ \rho \xrightarrow{\mathcal{S}} S(\rho) = M_0 \rho M_0^\dagger + M_1 \rho M_1^\dagger, \]  
with

\[ M_0 := \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\rho} \end{pmatrix}, \quad M_1 := \begin{pmatrix} 0 & \sqrt{\rho} \\ 0 & 0 \end{pmatrix}. \]

We can verify that

\[ M_0^\dagger M_0 + M_1^\dagger M_0 = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) + \left( \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\rho} & 0 \\ 0 & \sqrt{\rho} & 0 & 0 \\ \sqrt{\rho} & 0 & 0 & 0 \end{array} \right) \]

\[ = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) = I_2. \]

And the amplitude-damping channel works as

\[ S(\rho) = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \end{array} \right) \left( \begin{array}{cccc} \rho_{00} & \rho_{01} & 0 & 0 \\ \rho_{10} & \rho_{11} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) + \left( \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\rho} & 0 \\ 0 & \sqrt{\rho} & 0 & 0 \\ \sqrt{\rho} & 0 & 0 & 0 \end{array} \right) \left( \begin{array}{cccc} \rho_{00} & 0 & 0 & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) \]

\[ = \left( \begin{array}{cccc} \rho_{00} & \rho_{01} & 0 & 0 \\ \rho_{10} & \rho_{11} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) + \left( \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\rho} & 0 \\ 0 & \sqrt{\rho} & 0 & 0 \\ \sqrt{\rho} & 0 & 0 & 0 \end{array} \right) \left( \begin{array}{cccc} \rho_{00} & 0 & 0 & 0 \\ 0 & \sqrt{1-\rho} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right) \]

\[ = \left( \begin{array}{cccc} \rho_{00} + \rho_{11} - (1-\rho)\rho_{11} & \sqrt{1-\rho}\rho_{01} & 0 & 0 \\ \sqrt{1-\rho}\rho_{10} & (1-\rho)\rho_{11} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right). \]

i.e.,

\[ S(\rho) = \left( \begin{array}{cccc} \rho_{00} + \rho_{11} - (1-\rho)\rho_{11} & \sqrt{1-\rho}\rho_{01} & 0 & 0 \\ \sqrt{1-\rho}\rho_{10} & (1-\rho)\rho_{11} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right). \]

Thus

\[ S^n(\rho) = \left( \begin{array}{cccc} \rho_{00} + \rho_{11} - (1-\rho)^n\rho_{11} & (1-\rho)^{n/2}\rho_{01} & 0 & 0 \\ (1-\rho)^{n/2}\rho_{10} & (1-\rho)^n\rho_{11} & 0 & 0 \\ 0 & 0 & \sqrt{1-\rho} & 0 \\ 0 & 0 & 0 & \sqrt{\rho} \end{array} \right), \]

with \( n \to \infty, (1-\rho)^n \to 0 \), i.e.,

\[ S^n(\rho) \xrightarrow{n \to \infty} \left( \begin{array}{cccc} \rho_{00} + \rho_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right). \]

**Example:** If

\[ \rho := |a|^2|0\rangle \langle 0| + |b|^2|1\rangle \langle 1|, \quad \text{with} \quad |a|^2 + |b|^2 = 1, \]

we can get

\[ S^n(\rho) \xrightarrow{n \to \infty} \left( \begin{array}{cccc} |a|^2 + |b|^2 & 0 \\ 0 & 0 \end{array} \right) = \left( \begin{array}{cccc} 1 & 0 \\ 0 & 0 \end{array} \right) = |0\rangle \langle 0|. \]

Therefore, we can get the pure state \(|0\rangle \langle 0|\) from the mixed state \( \rho = |a|^2|0\rangle \langle 0| + |b|^2|1\rangle \langle 1| \) through the amplitude-damping channel. The geometric picture can be described as the Bloch ball contracted toward to the point \(|0\rangle \langle 0|\).
17.7.6 The two-Pauli channel

\[ M_0 := \sqrt{1-p}I_2, \quad M_1 := \sqrt{p}/2 X, \quad M_2 := \sqrt{p}/2 Z. \]  

In the Bloch sphere representation of the single-qubit density matrix

\[ \rho := \frac{1}{2}(I_2 + \mathbf{p} \cdot \mathbf{\sigma}), \]  

after the two-Pauli channel,

\[ S(\rho) = (1-p)\rho + \frac{p}{2} X\rho X + \frac{p}{2} Z\rho Z \]

\[ = \frac{1}{2}(I_2 + (1-p)p_1 X + (1-2p)p_2 Y + (1-p)p_3 Z). \]  

And the parameter vector \( \mathbf{p} \) changes in the way

\[ p_1^{(1)} = (1-p)p_1, \quad p_2^{(1)} = (1-2p)p_2, \quad p_3^{(1)} = (1-p)p_3, \]  

thus we can see

\[ S^{(n)}(\rho) \to \frac{1}{2} I_2 \]  

as \( n \to +\infty \). The polarized state is transformed into unpolarized state, which implies that the information is completely lost.

17.8 The master equation

The master Equation is also called Lindblud’s Equation. In general, we apply the CPTP mapping to describe the quantum operation. In special cases, we perform the approximation and view the evolution as Markovin process:

\[ t_{m-1} \quad t_m \quad t_{m+1} \]

where state at time \( t_{i+1} \) only depends on state at time \( t_i \). Therefore, in Makovin process, \( \rho_A(t+dt) \) depends only on \( \rho_A(t) \). From Figure 17.2 we can see there are three time scales, i.e., \( \Delta t^A, \Delta t^O \) and \( \Delta t^E \). In Markovin process, we assume

\[ \Delta t^A \gg \Delta t^O \gg \Delta t^E, \]  

which means that for the observer, the physical system is changing very slowly while the environment is changing very fast. This ensures that \( \rho_A(t+dt) \) depends on \( \rho_A(t) \) only, so the differential equation is possible.
1° First-order approximation.

\[ \rho(dt) = \rho(0) + O(dt) = \rho(0) + d\rho(0). \quad (17.8.3) \]

2° Kraus representation.

\[ \rho(dt) = S(\rho(0)) = \sum_{\mu} M_{\mu}(dt)\rho(0)M_{\mu}^\dagger(dt), \quad (17.8.4) \]

with

\[ \sum_{\mu} M_{\mu}^\dagger(dt)M_{\mu}(dt) = \text{Id}. \]

3° Assume

\[
\begin{align*}
M_0 &= \text{Id} + (-iH + K)dt, \quad \text{with } H^\dagger = H, K^\dagger = K; \\
M_k &= L_k\sqrt{dt}, \quad \text{with } k = 1, 2, 3, \ldots
\end{align*} \quad (17.8.5a, 17.8.5b)
\]

4°

\[ \rho(0) + d\rho(0) = [\text{Id} + (-iH + K)dt] \rho(0) [\text{Id} + (iH + K)dt] + \sum_k L_k^\dagger \rho(0) L_k dt, \quad (17.8.6) \]

then

\[ d\rho = (-iH + K)\rho dt + \rho(iH + K) dt + \sum_{\nu=1} L_{\nu}^\dagger \rho L_{\nu} dt, \quad (17.8.7) \]

i.e.,

\[ \frac{d\rho}{dt} = \frac{1}{i}[H, \rho] + \{K, \rho\} + \sum_{\nu} L_{\nu}^\dagger \rho L_{\nu}. \quad (17.8.8) \]

With the requirement relation

\[ \sum_{\mu} M_{\mu}^\dagger M_{\mu} = \text{Id}, \quad (17.8.9) \]

we have the constrain relation

\[ (-iH + K) + (iH + K) + \sum_{\nu} L_{\nu}^\dagger L_{\nu} = 0, \quad (17.8.10) \]

i.e.,

\[ K = -\frac{1}{2} \sum_{\nu} L_{\nu}^\dagger L_{\nu}. \quad (17.8.11) \]

Hence, we can obtain

\[ \frac{d\rho}{dt} = \frac{1}{i}[H, \rho] + \sum_{\nu} \left( L_{\nu}^\dagger \rho L_{\nu} - \frac{1}{2} (L_{\nu}^\dagger L_{\nu}, \rho) \right). \quad (17.8.12) \]

**Note:** For the special case of \( K = 0 \) and \( L_k = 0 \), \( k = 1, 2, 3, \ldots \), we can get

\[ \frac{d\rho}{dt} = \frac{1}{i}[H, \rho], \quad (17.8.13) \]

which is the time evolution of closed system governed by the Shrödinger Equation.
Chapter 18

Notes on Finite Group Theory
Chapter 19

Notes on Stabilizer Formalism of Quantum Error Correction Codes
Part V

Fault-tolerant Quantum Computation
Chapter 20

Measurement-Based Quantum Computation
Chapter 21

Topological Quantum Computation
Chapter 22

Integrable Quantum Computation

Interested readers are invited to refer to two documents on the same homepage: the one is the published paper with the title of “Integrable Quantum Computation”; and the other one is the associated .ppt file of interpreting the published one.
Part VI

Research Projects
Chapter 23

Research Projects