

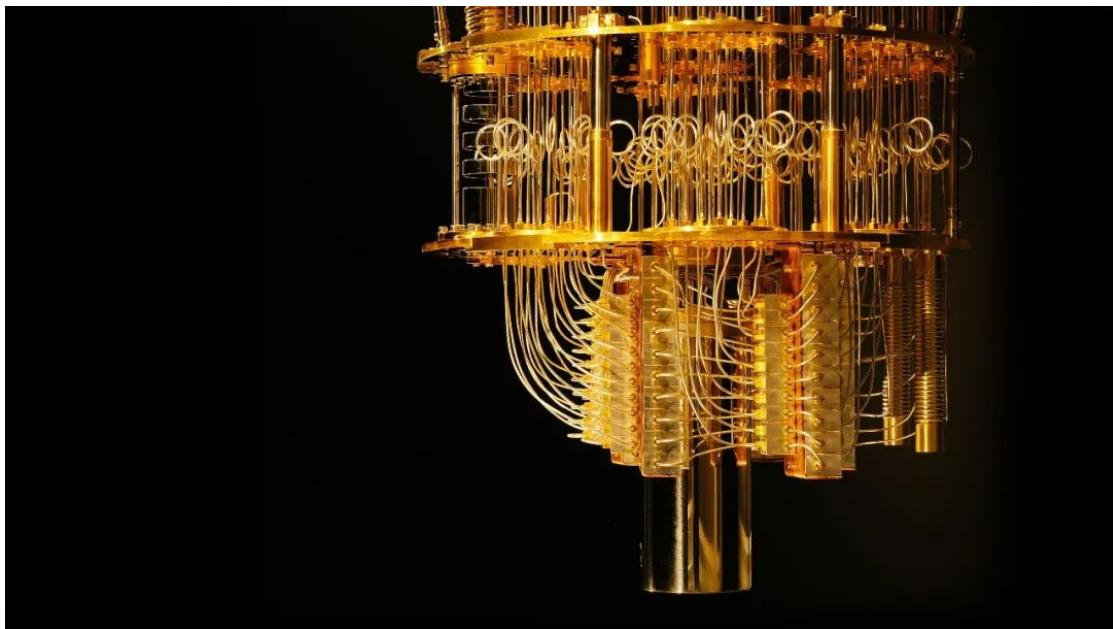
# Lecture Notes for Quantum Computing

Chao Liang

cliang@whu.edu.cn

NATIONAL ENGINEERING RESEARCH CENTRE FOR MULTIMEDIA SOFTWARE (NERCMS)

SCHOOL OF COMPUTER SCIENCE, WUHAN UNIVERSITY



Spring, 2024

(Last updated on May 29, 2024)

# Contents

<b>1</b>	<b>Introduction and Complex Number</b>	<b>4</b>
1.1	Introduction to Quantum Computing . . . . .	4
1.1.1	A Brief History . . . . .	4
1.1.2	Prof. Andrew Chi-Chih Yao's Talk in Micius Salon . . . . .	5
1.2	Complex Numbers . . . . .	6
1.2.1	Definitions . . . . .	7
1.2.2	The Algebra of Complex Numbers . . . . .	7
1.2.3	The Geometry of Complex Numbers . . . . .	10
<b>2</b>	<b>Complex Vector Space</b>	<b>13</b>
2.1	Complex Vector Space . . . . .	13
2.1.1	Unary Operations . . . . .	14
2.1.2	Matrix Multiplication . . . . .	15
2.1.3	Linear Map . . . . .	15
2.2	Basis and Dimension . . . . .	16
2.2.1	Basis . . . . .	16
2.2.2	Dimension . . . . .	16
2.3	Inner Product and Hilbert Space . . . . .	18
2.3.1	Inner Product . . . . .	18
2.3.2	Hilbert Space . . . . .	20
2.4	Eigenvalue and Eigenvector . . . . .	21
2.5	Hermitian and Unitary Matrices . . . . .	21
2.5.1	Hermitian Matrix . . . . .	21
2.5.2	Unitary Matrix . . . . .	23
<b>3</b>	<b>The Leap from Classic to Quantum</b>	<b>26</b>
3.1	Classic Deterministic Systems . . . . .	26
3.2	Probabilistic Systems . . . . .	27
3.3	Quantum Systems . . . . .	30
<b>4</b>	<b>Quantum Mechanics</b>	<b>35</b>
4.1	Quantum States . . . . .	36
4.1.1	Case 1: position on a line . . . . .	36
4.1.2	Case 2: single-particle spin system . . . . .	37
4.2	Observables and Measuring . . . . .	41
4.2.1	Basic concepts . . . . .	41

4.2.2	The principles . . . . .	42
4.2.3	The expected value of observing . . . . .	44
4.2.4	Multiple-step observing . . . . .	44
4.3	Dynamics . . . . .	45
<b>5</b>	<b>Quantum Gates</b>	<b>46</b>
5.1	Bits and Qubits . . . . .	46
5.2	Classic Gates . . . . .	48
5.3	Reversible Gates . . . . .	51
5.3.1	CNOT gate . . . . .	52
5.3.2	Toffoli gate . . . . .	53
5.3.3	Fredkin gate . . . . .	54
5.4	Quantum Gates . . . . .	54
<b>6</b>	<b>Quantum Cryptography</b>	<b>58</b>
6.1	Classic Cryptography . . . . .	58
6.1.1	Caesar cipher . . . . .	58
6.1.2	One-Time-Pad protocol . . . . .	59
6.1.3	Diffie-Hellman key exchange . . . . .	60
6.1.4	Public- and private-key cryptography . . . . .	61
6.2	Quantum Key Exchange . . . . .	62
6.2.1	The BB84 protocol . . . . .	63
6.2.2	The B92 protocol . . . . .	66
6.2.3	The EPR protocol . . . . .	67
6.3	Quantum Teleportation . . . . .	68

# 1 Introduction and Complex Number

## 1.1 Introduction to Quantum Computing

### 1.1.1 A Brief History

Quantum Mechanics as a branch of physics began with a set of scientific discoveries in the late 19th Century and has been in active development ever since. Most people will point to the 1980s as the start of physicists actively looking at computing with quantum systems<sup>1</sup>:

- **1982:** History of quantum computing starts with Richard Feynman lectures on the potential advantages of computing with quantum systems.
- **1985:** David Deutsch publishes the idea of a “universal quantum computer”.
- **1994:** Peter Shor presents an algorithm that can efficiently find the factors of large numbers, significantly outperforming the best classical algorithm and theoretically putting the underpinning of modern encryption at risk (referred to now as Shors algorithm).
- **1996:** Lov Grover presents an algorithm for quantum computers that would be more efficient for searching databases (referred to now as Groves search algorithm).
- **1996:** Seth Lloyd proposes a quantum algorithm which can simulate quantum-mechanical systems.
- **1999:** D-Wave Systems founded by Geordie Rose.
- **2000:** Eddie Farhi at MIT develops idea for adiabatic quantum computing.
- **2001:** IBM and Stanford University publish the first implementation of Shors algorithm, factoring 15 into its prime factors on a 7-qubit processor.
- **2010:** D-Wave One: first commercial quantum computer released (annealer).
- **2016:** IBM makes quantum computing available on IBM Cloud.
- **2019:** Google claims the achievement of quantum supremacy. Quantum Supremacy was termed by John Preskill in 2012 to describe when quantum systems could perform tasks surpassing those in the classical world.

A more complete history comes from the quantumpedia<sup>2</sup>, where the development of quantum computing is divided into five distinct periods (Figure 1.1):

---

<sup>1</sup><https://thequantuminsider.com/2020/05/26/history-of-quantum-computing/>

<sup>2</sup><https://quantumpedia.uk/a-brief-history-of-quantum-computing-e0bbd05893d0>

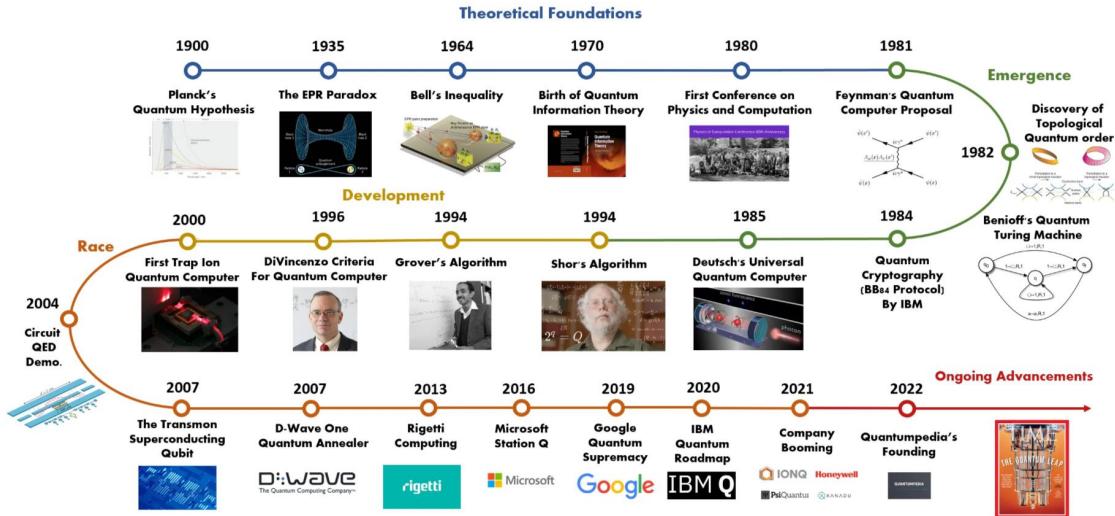


Figure 1.1: A Brief History of Quantum Computing (Copyright: Quantumpedia)

- **1900–1980:** The Theoretical Foundations of Quantum Computing.
- **1980–1994:** The Emergence of Quantum Computing.
- **1994–2000:** The Development of Quantum Algorithms.
- **2000–2021:** The Race to Build Quantum Computers.
- **2021–present:** Ongoing Advancements.

### 1.1.2 Prof. Andrew Chi-Chih Yao’s Talk in Micius Salon

Prof. Yao gave a talk entitled “The Advent of Quantum Computing” in Micius Salon in 2018<sup>3</sup>. Here are some key points:

- Two key topics: (1) what is the nature of quantum computer?; and (2) where does quantum computer gets its power from?
- The particle-wave duality plays the starting role in making it possible for us to do quantum computing faster than classic computing under certain circumstances
- Richard Feynman’s question: can quantum physics be simulated efficiently? Answer: unlikely by a classic computer, but hopefully by a quantum computer.
- The comparison of classic computer and quantum computers (Figure 1.2). Classic computers manipulate classic bits  $0110 \dots$  with Boolean operations in  $\{0, 1\}^n$ , while quantum computers manipulate quantum bits  $|0101 \dots\rangle$  with “rotations” in  $\mathbb{C}^{2^n}$

<sup>3</sup><https://www.bilibili.com/video/BV1Ct411Z7BQ/>

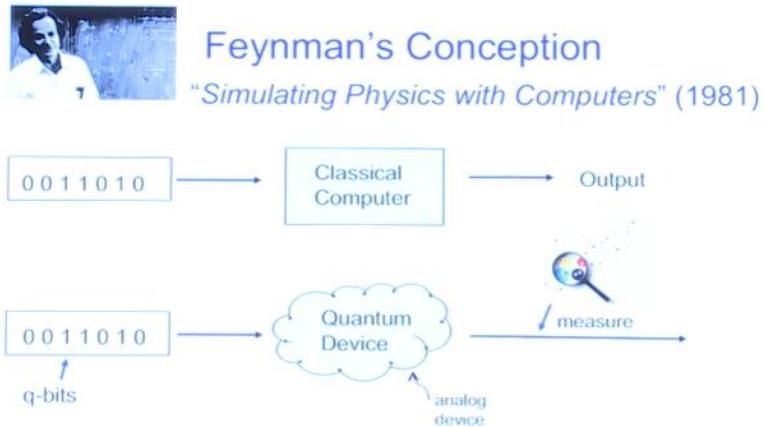


Figure 1.2: The comparison of classic and quantum computers.

- The **parallel superposition** is brought by the fact that each quantum bit represents not a single state, but a “probabilistic distribution” of states as shown in Figure 1.3<sup>4</sup>.



Figure 1.3: Some examples of parallel superposition.

- Parallelism could speed up computational tasks by parallel search.

## 1.2 Complex Numbers

The original motivation for the introduction of complex numbers was seeking solutions of polynomial equations. Here is the simplest example:

$$x^2 + 1 = 0 \quad (1.1)$$

Obviously, we cannot find its solution in the set of real numbers. To solve this problem, Mathematics introduces following definitions.

The fundamental reason we review complex numbers first in this course is that physics has recognized that quantum mechanics must be complex in nature<sup>5</sup> (Figure 1.4).

<sup>4</sup>Strictly speaking, Figure 1.3a is more accurate than the other two because states in quantum computing are mutually exclusive rather than similar to each other.

<sup>5</sup><https://physics.aps.org/articles/v15/7>

The screenshot shows a journal article from the Physics website. The title is "Quantum Mechanics Must Be Complex" by Alessio Avella. The article discusses independent studies demonstrating that quantum mechanics involving complex numbers is necessary to reproduce experimental results. It includes a conceptual sketch of a three-party game with two sources distributing entangled qubits to three observers (A, B, C) and mathematical equations for the entangled states.

Figure 1.4: Quantum mechanics must be complex (source: APS)

### 1.2.1 Definitions

**Definition 1.1 (Imaginary Number).** An *imaginary number* is a real number multiplied by the imaginary unit  $i$ , which is defined by its property  $i^2 = -1$ <sup>6</sup> or  $i = \sqrt{-1}$ .

**Definition 1.2 (Complex Number).** A *complex number* is a hybrid entity which adds a real number with an imaginary number, for instance,

$$c = a + b \times i = a + bi \quad (1.2)$$

where  $a, b$  are two real numbers,  $a$  is called the *real part* of  $c$ , whereas  $b$  is its *imaginary part*. The set of all complex numbers will be denoted as  $\mathbb{C}$ . When the  $\times$  is understood, we shall omit it.

**Proposition 1 (Fundamental Theorem of Algebra).** Every polynomial equation of one variable with complex coefficients has a complex solution.

### 1.2.2 The Algebra of Complex Numbers

**Definition 1.3 (Ordered Pair Representation).** Ordered pair representation defines a complex number as an ordered pair of reals:

$$c = a + bi^7 \mapsto (a, b) \quad (1.3)$$

<sup>6</sup>Thanks for Ziyi Ding's correction of  $1 \rightarrow -1$ .

<sup>7</sup>Thanks for Xin Shu's correction of  $a + b \rightarrow a + bi$ .

Hence, ordinary real numbers can be identified with pairs  $(a, 0)$

$$a \mapsto (a, 0)^8 \quad (1.4)$$

whereas imaginary numbers can be identified with pairs  $(0, b)$ . In particular,

$$i \mapsto (0, 1) \quad (1.5)$$

The four **arithmetic operations** between two complex numbers can be expressed as:

- Addition:

$$c_1 + c_2 = (a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \quad (1.6)$$

- Subtraction:

$$c_1 - c_2 = (a_1, b_1) - (a_2, b_2) = (a_1 - a_2, b_1 - b_2) \quad (1.7)$$

- Multiplication:

$$c_1 \times c_2 = (a_1, b_1) \times (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1) \quad (1.8)$$

- Subdivision:

$$\frac{c_1}{c_2} = \frac{(a_1, b_1)}{(a_2, b_2)} = \left( \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2}, \frac{a_2 b_1 - a_1 b_2}{a_2^2 + b_2^2} \right) \quad (1.9)$$

With the addition and multiplication operations, we can re-write a complex number as

$$c = a + bi = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0) \times (0, 1) \quad (1.10)$$

and from the denominator in the quotient formula in Eq.(1.9), we can define the **modulus** of a complex number as:

$$|c| = |a + bi| = +\sqrt{a^2 + b^2} \quad (1.11)$$

which has two useful properties:

- Property 1:  $\forall c_1, c_2 \in \mathbb{C}, |c_1||c_2| = |c_1 c_2|.$
- Property 2:  $\forall c_1, c_2 \in \mathbb{C}, |c_1 + c_2| \leq |c_1| + |c_2|.$

where the second property is also called triangular inequality of modulus operation.

Based on the above basic operations, it is easy to verify that complex numbers have the following **algebraic properties**:

- Addition has an identity called **additive identity**:  $(0, 0)$ , such that

$$\forall c \in \mathbb{C}, c + (0, 0) = c \quad (1.12)$$

---

<sup>8</sup>Thanks for Yipan Wei's correction of  $(a, b) \rightarrow (a, 0)$ .

- Multiplication has an identity called **multiplicative identity**:  $(1, 0)$ , such that

$$\forall c \in \mathbb{C}, c \times (1, 0) = (1, 0) \times c = c \quad (1.13)$$

- Both addition and multiplication are commutative:

$$\begin{cases} c_1 + c_2 = c_2 + c_1 \\ c_1 \times c_2 = c_2 \times c_1 \end{cases} \quad (1.14)$$

- Both addition and multiplication are associative:

$$\begin{cases} (c_1 + c_2) + c_3 = c_1 + (c_2 + c_3) \\ (c_1 \times c_2) \times c_3 = c_1 \times (c_2 \times c_3) \end{cases} \quad (1.15)$$

- Multiplication distributes with respect to addition:

$$c_1 \times (c_2 + c_3) = c_1 \times c_2 + c_1 \times c_3 \quad (1.16)$$

- Subtraction is defined everywhere.
- Division is defined everywhere except when the divisor is zero.

**Definition 1.4 (Conjugation).** *Ordered pair representation defines a complex number as an ordered pair of reals:*

$$c = a + bi^9 \mapsto (a, b) \quad (1.17)$$

Besides basic arithmetic operations and modulus operation, complex numbers have a unique operation called **conjugation**. If  $c = a + bi$  is an arbitrary complex number, then the conjugate of  $c$  is  $\bar{c} = a - bi$ . Two numbers related by conjugation are said to be **complex conjugates** of each other. The conjugation operation has several basic properties:

- Property 1: Conjugate respects addition  $\overline{c_1 + c_2} = \bar{c}_1 + \bar{c}_2$ .
- Property 2: Conjugate respects multiplication  $\overline{c_1 \times c_2} = \bar{c}_1 \times \bar{c}_2$ .
- Property 3: Conjugate  $c \mapsto \bar{c}$  is bijective.
- Property 4: The modulus squared of a complex number is obtained by multiplying the number with its conjugate  $c \times \bar{c} = |c|^2$ .

---

<sup>9</sup>Thanks for Xin Shu's correction of  $a + b \rightarrow a + bi$ .

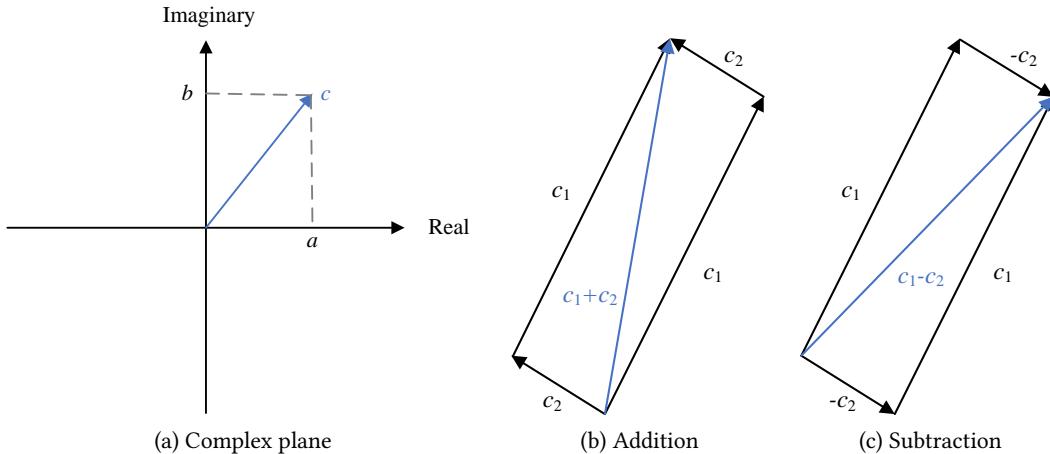


Figure 1.5: The complex plane (a) and parallelogram rule for (b) addition and (c) subtraction.

### 1.2.3 The Geometry of Complex Numbers

**Definition 1.5 (Complex Plane or Argand Plane).** *The complex plane is the plane formed by the complex numbers, with a Cartesian coordinate system such that the horizontal  $x$ -axis, called the real axis, is formed by the real numbers, and the vertical  $y$ -axis, called the imaginary axis, is formed by the imaginary numbers.*

In the complex plane (Figure 1.5a), we can easily find that the modulus is nothing more than the length of the vector. Indeed, the length of a vector, via Pythagoras theorem, is the square root of the sum of the squares of its edges, which is precisely the modulus, as defined in the previous section.

Next comes addition: vectors can be added using the so-called **parallelogram rule** illustrated by Figure 1.5b. In words, draw the parallelogram whose parallel edges are the two vectors to be added; their sum is the diagonal.

Subtraction too has a clear geometric meaning: subtracting  $c_2$  from  $c_1$  is the same as adding the negation of  $c_2$ , i.e.,  $-c_2$ , to  $c_1$  (Figure 1.5c).

To give a simple geometrical meaning to multiplication, we need to develop yet another characterization of complex numbers.

**Definition 1.6 (Polar Coordinate System).** *The polar coordinate system is a two-dimensional coordinate system in which each point on a plane is determined by a distance  $\rho$  from a reference point and an angle  $\theta$  from a reference direction.*

Similar to the previous **Cartesian representation**  $(a, b)$ , the **polar representation**  $(\rho, \theta)$  is capable to uniquely determine a complex number because these two representations can be mutually converted:

$$(a, b) \mapsto (\rho, \theta) \quad (1.18)$$

where  $\rho$  is the modulus

$$\rho = \sqrt{a^2 + b^2} \quad (1.19)$$

and  $\theta$  is also easy, via trigonometry<sup>10</sup>

$$\theta = \text{atan2}(b, a) \in (-\pi, \pi] \quad (1.21)$$

$$(\rho, \theta) \mapsto (a, b) \quad (1.22)$$

where  $a$  is the real part

$$a = \rho \cos(\theta) \quad (1.23)$$

and  $b$  is the imaginary part

$$b = \rho \sin(\theta) \quad (1.24)$$

In physics and engineering, angle  $\theta$  is also known as **phase** and distance  $\rho$  is also known as **magnitude**. Hence, we have another definition of a complex number

**Definition 1.7** (Complex Number). *A complex number is a magnitude and a phase.*

We are now ready for multiplication: given two complex numbers in polar coordinates,  $c_1 = (\rho_1, \theta_1)$  and  $c_2 = (\rho_2, \theta_2)$ , their product can be obtained by simply multiplying their magnitude and adding their phase:

$$c_1 \times c_2 = (\rho_1, \theta_1) \times (\rho_2, \theta_2) = (\rho_1 \rho_2, \theta_1 + \theta_2) \quad (1.25)$$

Now that we are armed with a geometric way of looking at multiplication, we can tackle division as well. After all, division is nothing more than the inverse operation of multiplication:

$$\frac{c_1}{c_2} = \left( \frac{\rho_1}{\rho_2}, \theta_1 - \theta_2 \right) \quad (1.26)$$

On this basis, we can further derive fast  $n$ -order power (Figure 1.6a) and root (Figure 1.6b) calculations about a complex number  $c = (\rho, \theta)$

$$c^n = (\rho^n, n\theta) \quad (1.27)$$

and<sup>11</sup>

$$c^{\frac{1}{n}} = \left( \rho^{\frac{1}{n}}, \frac{1}{n}(\theta + k2\pi) \right), \text{ where } k = 0, 1, \dots, n-1 \quad (1.28)$$

Instructor: Chao Liang

---

<sup>10</sup>The function atan2 computes the principal value of the argument function applied to the complex number  $a + bi$ .

$$a + b \quad (1.20)$$

For more information please refer to <https://en.wikipedia.org/wiki/Atan2>.

<sup>11</sup>Thanks for Ziyi Ding's correction of  $n \rightarrow \theta$  in Eq.(1.28).

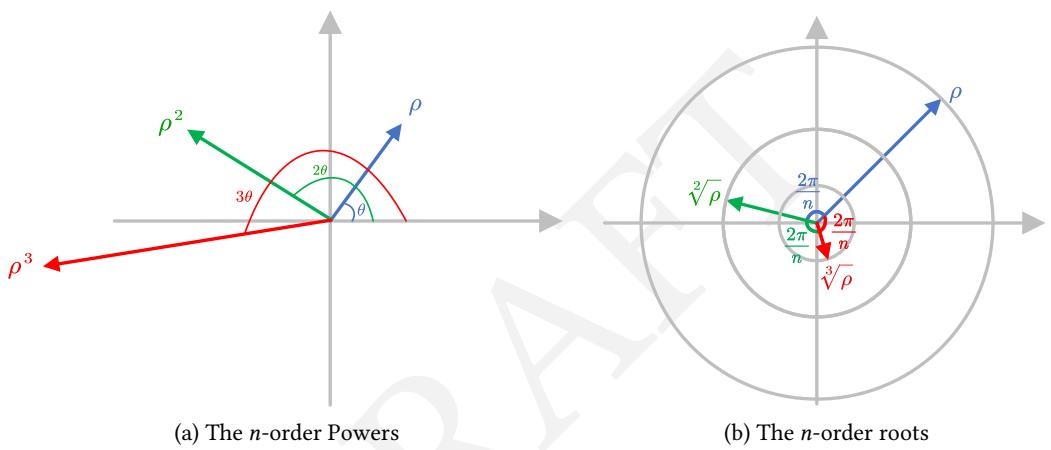


Figure 1.6: The  $n$ -order powers (a) and roots (b) of a complex number.

## 2 Complex Vector Space

### 2.1 Complex Vector Space

**Definition 2.1 (Complex Vector Space).** A complex vector space is a nonempty set  $\mathbb{V}$ , whose elements we shall call vectors, with three operations

- Addition:  $+: \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{V}$
- Negation:  $-: \mathbb{V} \rightarrow \mathbb{V}$
- Scalar multiplication:  $\cdot: \mathbb{C} \times \mathbb{V} \rightarrow \mathbb{V}$

and a distinguished element called the **zero vector**  $\mathbf{0} \in \mathbb{V}$  in the set. These operations and zero must satisfy the following properties:  $\forall v, w, x \in \mathbb{V}$  and for all  $c, c_1, c_2 \in \mathbb{C}$ ,

- i. Commutativity of addition:  $v + w = w + v$ ,
- ii. Associativity of addition:  $(v + w) + x = v + (w + x)$ ,
- iii. Additive identity:  $v + \mathbf{0} = v = \mathbf{0} + v$ ,
- iv. Additive inverse:  $v + (-v) = \mathbf{0} = (-v) + v$ ,
- v. Multiplication identity:  $1 \cdot v = v$ ,
- vi. Scalar multiplication distributes over addition:  $c \cdot (v + w) = c \cdot v + c \cdot w$ ,
- vii. Scalar multiplication distributes over complex addition:  $(c_1 + c_2) \cdot v = c_1 \cdot v + c_2 \cdot v$ ,

#### Example 2.1: $\mathbb{C}^n$

$\mathbb{C}^n$ , the set of vectors of length  $n$  with complex entries, is a complex vector space.

#### Example 2.2: $\mathbb{C}^{m \times n}$

$\mathbb{C}^{m \times n}$ , the set of all  $m$ -by- $n$  matrices (two-dimensional arrays) with complex entries, is a complex vector space.

### 2.1.1 Unary Operations

Three **unary operations** for  $\forall \mathbf{A} \in \mathbb{C}^{m \times n}$

- Transpose:

$$\mathbf{A}^\top \in \mathbb{C}^{n \times m} \text{ such that } \mathbf{A}^\top(j, k) = \mathbf{A}(k, j)^1 \quad (2.1)$$

- Conjugate:

$$\overline{\mathbf{A}} \in \mathbb{C}^{m \times n} \text{ such that } \overline{\mathbf{A}}(j, k) = \overline{\mathbf{A}(j, k)} \quad (2.2)$$

- Ajoint:

$$\mathbf{A}^\dagger \in \mathbb{C}^{n \times m} \text{ such that } \mathbf{A}^\dagger(j, k) = \overline{\mathbf{A}(k, j)} \quad (2.3)$$

**Property 1** (Properties of Transpose).  $\forall c \in \mathbb{C}$  and  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$

- Transpose is idempotent:

$$(\mathbf{A}^\top)^\top = \mathbf{A} \quad (2.4)$$

- Transpose respects addition:

$$(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top \quad (2.5)$$

- Transpose respects scalar multiplication:

$$(c \cdot \mathbf{A})^\top = c \cdot \mathbf{A}^\top \quad (2.6)$$

**Property 2** (Properties of Conjugate).  $\forall c \in \mathbb{C}$  and  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$

- Conjugate is idempotent:

$$\overline{\overline{\mathbf{A}}} = \mathbf{A} \quad (2.7)$$

- Conjugate respects addition:

$$\overline{\mathbf{A} + \mathbf{B}} = \overline{\mathbf{A}} + \overline{\mathbf{B}} \quad (2.8)$$

- Conjugate respects scalar multiplication:

$$\overline{c \cdot \mathbf{A}} = \bar{c} \cdot \overline{\mathbf{A}} \quad (2.9)$$

**Property 3** (Properties of Adjoint).  $\forall c \in \mathbb{C}$  and  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$

- Adjoint is idempotent:

$$(\mathbf{A}^\dagger)^\dagger = \mathbf{A} \quad (2.10)$$

- Adjoint respects addition:

$$(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger \quad (2.11)$$

- Conjugate respects scalar multiplication:

$$(c \cdot \mathbf{A})^\dagger = \bar{c} \cdot \mathbf{A}^\dagger \quad (2.12)$$

---

<sup>1</sup>Thanks for Ziyi Ding's correction of removing the redundant transpose mark.

### 2.1.2 Matrix Multiplication

**Property 4** (Properties of Matrix Multiplication).  $\forall \mathbf{A} \in \mathbb{C}^{m \times n}, \mathbf{B} \in \mathbb{C}^{n \times p}, \mathbf{C} \in \mathbb{C}^{n \times p}, \text{ and } \mathbf{D} \in \mathbb{C}^{p \times q}$ ,

- *Matrix multiplication distributes over addition:*

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C}) \quad (2.13)$$

$$(\mathbf{B} + \mathbf{C}) \times \mathbf{D} = (\mathbf{B} \times \mathbf{D}) + (\mathbf{C} \times \mathbf{D}) \quad (2.14)$$

- *Matrix multiplication respects scalar multiplication:*

$$c \cdot (\mathbf{A} \times \mathbf{B}) = (c \cdot \mathbf{A}) \times \mathbf{B} = \mathbf{A} \times (c \cdot \mathbf{B}) \quad (2.15)$$

- *Matrix multiplication relates to the transpose:*

$$(\mathbf{A} \times \mathbf{B})^\top = \mathbf{B}^\top \times \mathbf{A}^\top \quad (2.16)$$

- *Matrix multiplication respects to the conjugate:*

$$\overline{\mathbf{A} \times \mathbf{B}} = \overline{\mathbf{A}} \times \overline{\mathbf{B}} \quad (2.17)$$

- *Matrix multiplication relates to the adjoint:*

$$(\mathbf{A} \times \mathbf{B})^\dagger = \mathbf{B}^\dagger \times \mathbf{A}^\dagger \quad (2.18)$$

**The physical explanation.** The elements of  $\mathbb{C}^n$  are the ways of describing the states of a quantum system. Some suitable elements of  $\mathbb{C}^{nn}$  will correspond to the changes that occur to the states of a quantum system. Given a state  $\mathbf{x} \in \mathbb{C}^n$  and a matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , we shall form another state of the system  $\mathbf{A} \times \mathbf{x}$  which is an element of  $\mathbb{C}^n$ . Formally,  $\times$  in this case is a function  $\times : \mathbb{C}^{n \times n} \times \mathbb{C}^n \rightarrow \mathbb{C}^n$ . We say that the algebra of matrices “**acts**” on the vectors to yield new vectors.

### 2.1.3 Linear Map

**Definition 2.2** (Linear Map). A linear map from  $\mathbb{V}$  to  $\mathbb{V}'$  is a function  $f : \mathbb{V} \rightarrow \mathbb{V}', \forall \mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ , and  $c \in \mathbb{C}$  where

- *f respects the addition:*

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) \quad (2.19)$$

- *f respects the scalar multiplication:*

$$f(c \cdot \mathbf{v}) = c \cdot f(\mathbf{v}) \quad (2.20)$$

**The physical explanation.** We shall call any linear map from a complex vector space to itself an **operator**. If  $F : \mathbb{C}^n \rightarrow \mathbb{C}^n$  is an operator on  $\mathbb{C}^n$  and  $\mathbf{A}$  is an  $n$ -by- $n$  matrix such that for all  $\mathbf{v}$  we have  $F(\mathbf{v}) = \mathbf{A} \times \mathbf{v}$ , then we say that  $F$  is **represented** by  $\mathbf{A}$ . Several different matrices might represent the same operator.

## 2.2 Basis and Dimension

### 2.2.1 Basis

**Definition 2.3** (Linear Combination). Let  $\mathbb{V}$  be a complex (real) vector space.  $\mathbf{v} \in \mathbb{V}$  is a linear combination of the vectors  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}$  in  $\mathbb{V}$  if  $\mathbf{v}$  can be written as

$$\mathbf{v} = c_0 \cdot \mathbf{v}_0 + c_1 \cdot \mathbf{v}_1 + \dots + c_{n-1} \cdot \mathbf{v}_{n-1} \quad (2.21)$$

for some  $c_0, c_1, \dots, c_{n-1}$  in  $\mathbb{C}(\mathbb{R})$ .

**Definition 2.4** (Linearly Independent). A set  $\{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}\}$  of vectors in  $\mathbb{V}$  is called linearly independent if

$$\mathbf{0} = c_0 \cdot \mathbf{v}_0 + c_1 \cdot \mathbf{v}_1 + \dots + c_{n-1} \cdot \mathbf{v}_{n-1} \quad (2.22)$$

implies that  $c_0 = c_1 = \dots = c_{n-1} = 0$ . This means that the only way that a linear combination of the vectors can be the zero vector is if all the  $c_j$  are zero.

**Corollary 1.** For any  $\mathbf{v}_{i|i=0,1,\dots,n-1}$ , cannot be written as a combination of the others  $\{\mathbf{v}_j\}_{j=0,j \neq i}^{n-1}$

**Corollary 2.** For any  $\mathbf{0} \neq \mathbf{v} \in \mathbb{V}$ , unique coefficients  $\{c_i\}_{i=0}^{n-1}$

**Definition 2.5** (Basis). A set  $\mathcal{B} = \{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}\} \subseteq \mathbb{V}$  of vectors is called a basis of a (complex) vector space  $\mathbb{V}$  if both

- $\forall \mathbf{v} \in \mathbb{V}, \mathbf{v} = c_0 \cdot \mathbf{v}_0 + c_1 \cdot \mathbf{v}_1 + \dots + c_{n-1} \cdot \mathbf{v}_{n-1}$
- $\{\mathbf{v}_i | \mathbf{v}_i \in \mathbb{V}\}_{i=0}^{n-1}$  is linearly independent

### 2.2.2 Dimension

**Definition 2.6** (Dimension). The dimension of a (complex) vector space is the number of elements in a basis of the vector space.

**Definition 2.7** (Transition Matrix). A change of basis matrix or a transition matrix from basis  $\mathcal{B}$  to basis  $\mathcal{D}$  is a matrix  $\mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}}$  such that their coefficients satisfy

$$\mathbf{v}_{\mathcal{D}} = \mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}} \times \mathbf{v}_{\mathcal{B}} \quad (2.23)$$

In other words,  $\mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}}$  is a way of getting the coefficients with respect to one basis from the coefficients with respect to another basis.

**Remark.** Utilities of Transition Matrix

- Operator re-representation in a new basis

$$\mathbf{A}_{\mathcal{D}} = \mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}}^{-1} \times \mathbf{A}_{\mathcal{B}} \times \mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}} \quad (2.24)$$

- State re-representation in a new basis

$$\mathbf{v}_{\mathcal{D}} = \mathbf{M}_{\mathcal{D} \leftarrow \mathcal{B}} \times \mathbf{v}_{\mathcal{B}} \quad (2.25)$$

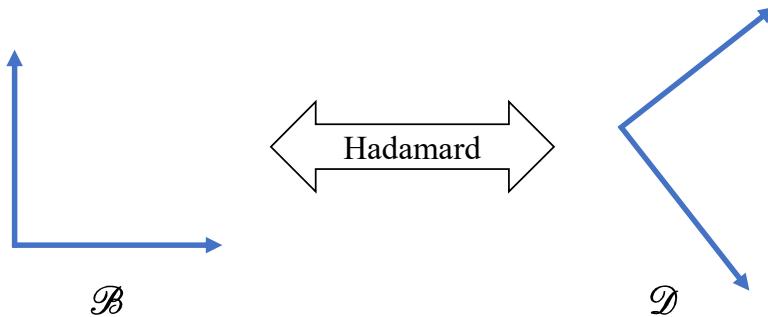


Figure 2.1: The Hadamard matrix for basis transition

### Example 2.3: Hadamard Matrix

In  $\mathbb{R}^2$ , the transition matrix from the canonical basis

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \text{ } ^a \quad (2.26)$$

to this other basis

$$\left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\} \quad (2.27)$$

is the Hadamard matrix:

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad (2.28)$$

as shown in Figure 2.1.

<sup>a</sup>Thanks for Xin Shu's correction of pointing out redundant basis vector.

**The motivation to change basis.** In physics, we are often faced with a problem in which it is easier to calculate something in a noncanonical basis. For example, consider a ball rolling down a ramp as depicted in Figure 2.2a.

The ball will not be moving in the direction of the canonical basis. Rather it will be rolling downward in the direction of +45, 45 basis. Suppose we wish to calculate when this ball will reach the bottom of the ramp or what is the speed of the ball. To do this, we change the problem from one in the canonical basis to one in the other basis. In this other basis, the motion is easier to deal with. Once we have completed the calculations, we change our results into the more understandable canonical basis and produce the desired answer. We might envision this as the flowchart shown in Figure 2.2b.

Throughout this course, we shall go from one basis to another basis, perform some calcu-

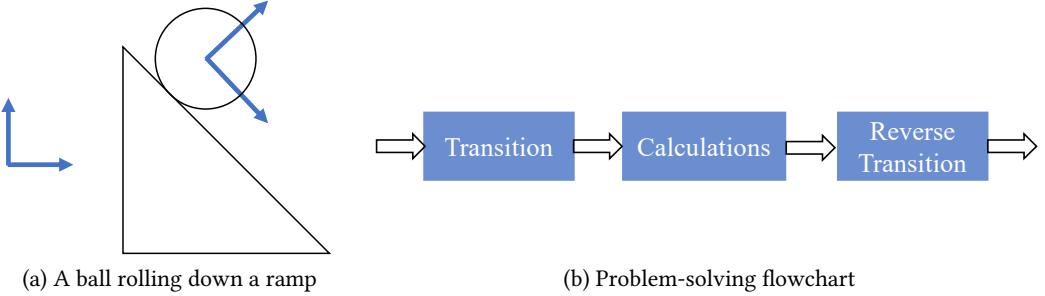


Figure 2.2: Basis transition example (a) and flowchart (b).

lations, and finally revert to the original basis. The Hadamard matrix will frequently be the means by which we change the basis.

## 2.3 Inner Product and Hilbert Space

### 2.3.1 Inner Product

**Definition 2.8** (Inner Product). *An inner product (also called a dot product or scalar product) on a complex vector space  $\mathbb{V}$  is a function*

$$\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{C} \quad (2.29)$$

that satisfies the following conditions for all  $v, v_1, v_2$ , and  $v_3$  in  $\mathbb{V}$  and for  $a, c \in \mathbb{C}$ :

i. Nondegenerate:

$$\langle v, v \rangle \geq 0 \text{ and } \langle v, v \rangle = 0^2 \Leftrightarrow v = 0 \quad (2.30)$$

ii. Respects addition:

$$\langle v_1 + v_2, v_3 \rangle = \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle \quad (2.31)$$

$$\langle v_1, v_2 + v_3 \rangle = \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle \quad (2.32)$$

iii. Respects scalar multiplication:

$$\langle c \cdot v_1, v_2 \rangle = \bar{c} \times \langle v_1, v_2 \rangle \quad (2.33)$$

$$\langle v_1, c \cdot v_2 \rangle = c \times \langle v_1, v_2 \rangle \quad (2.34)$$

iv. Skew symmetric:

$$\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle} \quad (2.35)$$

**Definition 2.9** (Inner Product Space). *A vector space with an inner space.*

---

<sup>2</sup>Thanks for Xin Shu's correction of pointing out the missing 0.

**Example 2.4: Inner Product in  $\mathbb{R}^n$** 

$\mathbb{R}^n$ : The inner product is given as

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1^\top \times \mathbf{v}_2 \quad (2.36)$$

**Example 2.5: Inner Product in  $\mathbb{C}^n$** 

$\mathbb{C}^n$ : The inner product is given as

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1^\dagger \times \mathbf{v}_2 \quad (2.37)$$

**Example 2.6: Inner Product in  $\mathbb{R}^{n \times n}$** 

$\mathbb{R}^{n \times n}$  has an inner product given for matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$  as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A}^\top \times \mathbf{B}) \quad (2.38)$$

where the **trace** of a square matrix  $C$  is given as the sum of the diagonal elements. That is,

$$\text{Tr}(C) = \sum_{i=0}^{n-1} C[i, i] \quad (2.39)$$

**Example 2.7: Inner Product in  $\mathbb{C}^{n \times n}$** 

$\mathbb{C}^{n \times n}$  has an inner product given for matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$  as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A}^\dagger \times \mathbf{B}) \quad (2.40)$$

**Definition 2.10** (Norm). *Norm is a unary function derived from inner product*

$$|\cdot| : \mathbb{V} \rightarrow \mathbb{R} \quad (2.41)$$

defined as  $|\mathbf{v}| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ , which has the following properties

- Norm is nondegenerate:

$$|\mathbf{v}| > 0 \text{ if } \mathbf{v} \neq \mathbf{0} \text{ and } |\mathbf{0}| = 0 \quad (2.42)$$

- Norm satisfies the triangular inequality:

$$|\mathbf{v} + \mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}| \quad (2.43)$$

- Norm respects scalar multiplication:

$$|c \cdot v| = |c| \cdot |v| \quad (2.44)$$

**Definition 2.11** (Distance). *Distance is a binary function defined based on norm*

$$d(\cdot, \cdot) : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R} \quad (2.45)$$

defined as  $d(v_1, v_2) = |v_1 - v_2| = \sqrt{\langle v_1 - v_2, v_1 - v_2 \rangle}$ , which has the following properties

- Distance is nondegenerate:

$$d(v, w) > 0 \text{ if } v \neq w \text{ and } d(v, w) = 0 \Leftrightarrow v = w \quad (2.46)$$

- Distance satisfies the triangular inequality:

$$d(u, v) \leq d(u, w) + d(w, v) \quad (2.47)$$

- Distance is symmetric:

$$d(u, v) = d(v, u) \quad (2.48)$$

**Definition 2.12** (Orthonormal Basis). *A basis  $\mathcal{B} = \{v_0, v_1, \dots, v_{n-1}\}$  for an inner space*

$$\langle v_i, v_j \rangle = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \quad (2.49)$$

with the following property

- For  $\forall v \in \mathbb{V}$  and any orthonormal basis  $\{e_i\}_{i=0}^{n-1}$  we have

$$v = \sum_{i=0}^{n-1} \langle e_i, v \rangle e_i \quad (2.50)$$

**Note:** inner product defines geometry in the vector space (Figure 2.3).

### 2.3.2 Hilbert Space

**Definition 2.13** (Cauchy Sequence). *Within an inner product space  $\mathbb{V}, \langle \cdot, \cdot \rangle$  (with the derived norm and a distance function), a sequence of vectors  $v_0, v_1, \dots$  is called a Cauchy sequence if  $\forall \epsilon > 0$ , there exists an  $N_0 \in \mathbb{N}$  such that for all  $m, n \geq N_0$ ,  $d(v_m, v_n) \leq \epsilon$ .*

**Definition 2.14** (Complete). *For any Cauchy sequence  $v_0, v_1, \dots$ , it is complete if there exist a  $\bar{v} \in \mathbb{V}$ , such that  $\lim_{n \rightarrow \infty} d(v_n - \bar{v}) = 0$ .*

**Definition 2.15** (Hilbert Space). *A Hilbert space is a complex inner space that is complete.*

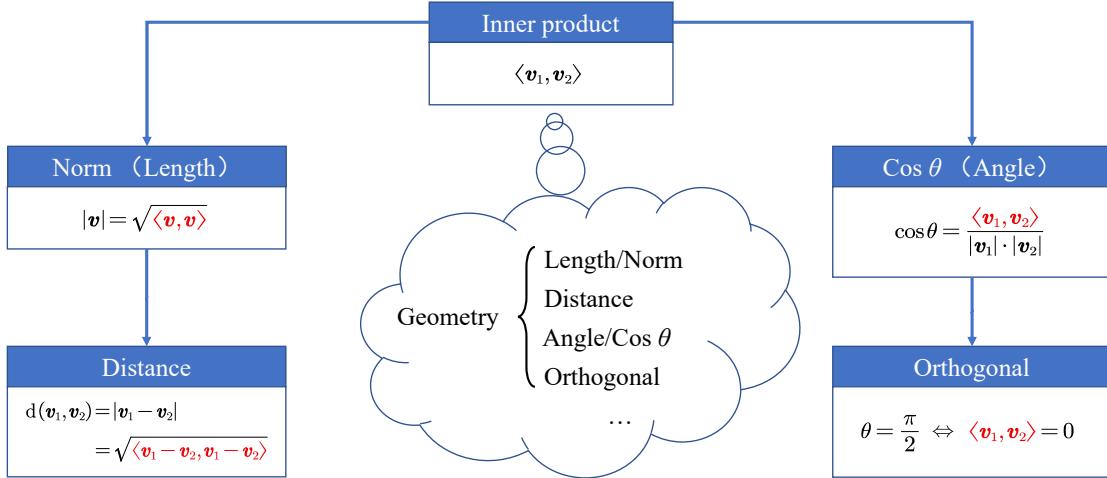


Figure 2.3: Inner product lays the geometric foundation in the vector space.

## 2.4 Eigenvalue and Eigenvector

**Definition 2.16** (Eigenvalue and Eigenvector). *For a matrix  $A \in \mathbb{C}^{n \times n}$ , if there is a number  $c \in \mathbb{C}$  and a vector  $0 \neq v \in \mathbb{C}^n$  such that*

$$Av = c \cdot v \quad (2.51)$$

*then  $c$  is called an eigenvalue of  $A$  and  $v$  is called an eigenvector of  $A$  associate with  $c$ .*

## 2.5 Hermitian and Unitary Matrices

### 2.5.1 Hermitian Matrix

**Definition 2.17** (Hermitian Matrix). *An  $n$ -by- $n$  matrix  $A$  is called hermitian if  $A^\dagger = A$ . In other words,  $A[j, k] = \overline{A[k, j]}$ .*

**Definition 2.18** (Self-Adjoint). *If  $A$  is a hermitian matrix then the operator that it represents is called self-adjoint.*

**Proposition 2.** *if  $A \in \mathbb{C}^{n \times n}$  is Hermitian,  $\forall v, w \in \mathbb{C}^n$  we have*

$$\langle Av, w \rangle = \langle v, Aw \rangle \quad (2.52)$$

*Proof.*

$$\langle \mathbf{A}\mathbf{v}, \mathbf{w} \rangle = (\mathbf{A}\mathbf{v})^\dagger \times \mathbf{w} \quad \% \text{ definition of inner product} \quad (2.53)$$

$$= \mathbf{v}^\dagger \times \mathbf{A}^\dagger \times \mathbf{w} \quad \% \text{ multiplication relates to the adjoint} \quad (2.54)$$

$$= \mathbf{v}^\dagger \times \mathbf{A} \times \mathbf{w} \quad \% \text{ definition of Hermitian matrices} \quad (2.55)$$

$$= \mathbf{v}^\dagger \times (\mathbf{A}\mathbf{w}) \quad \% \text{ multiplication is associative} \quad (2.56)$$

$$= \langle \mathbf{v}, \mathbf{A}\mathbf{w} \rangle \quad \% \text{ definition of inner product} \quad (2.57)$$

□

**Proposition 3.** *For a Hermitian matrix, its all eigenvalues are real.*

*Proof.* Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be a Hermitian matrix with an eigenvalue  $c \in \mathbb{C}$  and an eigenvector  $\mathbf{v} \in \mathbb{C}^n$

$$c\langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, c\mathbf{v} \rangle \quad \% \text{ inner product respects scalar multiplication} \quad (2.58)$$

$$= \langle \mathbf{v}, \mathbf{A}\mathbf{v} \rangle \quad \% \text{ definition of eigenvalue and eigenvector} \quad (2.59)$$

$$= \langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle \quad \% \text{ see Proposition 2} \quad (2.60)$$

$$= \langle c\mathbf{v}, \mathbf{v} \rangle \quad \% \text{ definition of eigenvalue and eigenvector} \quad (2.61)$$

$$= \bar{c}\langle \mathbf{v}, \mathbf{v} \rangle \quad \% \text{ inner product respects scalar multiplication} \quad (2.62)$$

(2.63)

□

**Proposition 4.** *For a Hermitian matrix, distinct eigenvectors that have distinct eigenvalues are orthogonal*

*Proof.* Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be a Hermitian matrix with two distinct eigenvectors  $\mathbf{v}_1 \neq \mathbf{v}_2 \in \mathbb{C}^n$  and their related eigenvalues  $c_1, c_2 \in \mathbb{C}$

$$c_2\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, c_2\mathbf{v}_2 \rangle \quad \% \text{ inner product respects scalar multiplication} \quad (2.64)$$

$$= \langle \mathbf{v}_1, \mathbf{A}\mathbf{v}_2 \rangle \quad \% \text{ definition of eigenvalue and eigenvector} \quad (2.65)$$

$$= \langle \mathbf{A}\mathbf{v}_1, \mathbf{v}_2 \rangle \quad \% \text{ see Proposition 2} \quad (2.66)$$

$$= \langle c_1\mathbf{v}_1, \mathbf{v}_2 \rangle \quad \% \text{ definition of eigenvalue and eigenvector} \quad (2.67)$$

$$= \bar{c}_1\langle \mathbf{v}_1, \mathbf{v}_2 \rangle \quad \% \text{ inner product respects scalar multiplication} \quad (2.68)$$

$$= c_1\langle \mathbf{v}_1, \mathbf{v}_2 \rangle \quad \% \text{ see proposition 3} \quad (2.69)$$

□

**Proposition 5** (The Spectral Theorem for Finite-Dimensional Self-Adjoint Operators.). *Every self-adjoint operator  $\mathbf{A}$  on a finite-dimensional complex vector space  $\mathbb{V}$  can be represented by a diagonal matrix whose diagonal entries are the eigenvalues of  $\mathbf{A}$ , and whose eigenvectors form an orthonormal basis for  $\mathbb{V}$  (we shall call this basis an eigenbasis).*

**Physical Meaning of Hermitian Matrix.** Hermitian matrices and their eigenbases will play a major role in our story. We shall see in the following lectures that associated with every physical observable of a quantum system there is a corresponding Hermitian matrix. Measurements of that observable always lead to a state that is represented by one of the eigenvectors of the associated Hermitian matrix.

### 2.5.2 Unitary Matrix

**Definition 2.19** (Unitary Matrix). *Given a reversible matrix  $\mathbf{U} \in \mathbb{C}^{n \times n}$  such that*

$$\mathbf{U} \times \mathbf{U}^\dagger = \mathbf{U}^\dagger \times \mathbf{U} = \mathbf{I}_n \quad (2.70)$$

*then  $\mathbf{U}$  is a unitary matrix.*

#### Example 2.8: Unitary Matrices

$$\mathbf{U}_1 = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ for any } \theta. \quad \mathbf{U}_2 = \begin{bmatrix} \frac{1+i}{2} & \frac{i}{\sqrt{3}} & \frac{3+i}{2\sqrt{15}} \\ \frac{-1}{2} & \frac{1}{\sqrt{3}} & \frac{4+3i}{2\sqrt{15}} \\ \frac{1}{2} & \frac{-i}{\sqrt{3}} & \frac{5i}{2\sqrt{15}} \end{bmatrix}$$

**Proposition 6** (Unitary Matrices Preserve Inner Products). *If  $\mathbf{U} \in \mathbb{C}^{n \times n}$  is unitary,  $\forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$  we have*

$$\langle \mathbf{U}\mathbf{v}, \mathbf{U}\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad (2.71)$$

*Proof.* Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be a Hermitian matrix with two distinct eigenvectors  $\mathbf{v}_1 \neq \mathbf{v}_2 \in \mathbb{C}^n$  and their related eigenvalues  $c_1, c_2 \in \mathbb{C}$

$$\langle \mathbf{U}\mathbf{v}, \mathbf{U}\mathbf{w} \rangle = (\mathbf{U}\mathbf{v})^\dagger \times (\mathbf{U}\mathbf{w}) \quad \% \text{ definition for inner product} \quad (2.72)$$

$$= \mathbf{v}^\dagger \mathbf{U}^\dagger \times \mathbf{U}\mathbf{w} \quad \% \text{ multiplication relates to adjoint} \quad (2.73)$$

$$= \mathbf{v}^\dagger \times \mathbf{I} \times \mathbf{w} \quad \% \text{ definition for unitary matrices} \quad (2.74)$$

$$= \langle \mathbf{v}, \mathbf{w} \rangle \quad \% \text{ definition for inner product} \quad (2.75)$$

□

**Proposition 7** (Unitary Matrices Preserve Norm). *If  $\mathbf{U} \in \mathbb{C}^{n \times n}$  is unitary,  $\forall \mathbf{v} \in \mathbb{C}^n$  we have*

$$|\mathbf{U}\mathbf{v}| = |\mathbf{v}| \quad (2.76)$$

*Proof.* Let  $A \in \mathbb{C}^{n \times n}$  be a Hermitian matrix with two distinct eigenvectors  $v_1 \neq v_2 \in \mathbb{C}^n$  and their related eigenvalues  $c_1, c_2 \in \mathbb{C}$ , we have<sup>3</sup>:

$$|Uv| = \sqrt{\langle Uv, Uv \rangle} \quad \% \text{ definition for norm} \quad (2.77)$$

$$= \sqrt{\langle v, v \rangle} \quad \% \text{ unitary matrices preserve inner product} \quad (2.78)$$

$$= |v| \quad \% \text{ definition for norm} \quad (2.79)$$

(2.80)

□

**Proposition 8** (Unitary Matrices Preserve Distance). *If  $U \in \mathbb{C}^{n \times n}$  is unitary,  $\forall v, w \in \mathbb{C}^n$  we have*

$$d(Uv, Uw) = d(v, w) \quad (2.81)$$

*Proof.*

$$d(Uv, Uw) = |Uv - Uw| \quad \% \text{ definition for distance} \quad (2.82)$$

$$= |U(v - w)| \quad \% \text{ multiplication distributes over addition} \quad (2.83)$$

$$= |v - w| \quad \% \text{ unitary matrices preserve norm} \quad (2.84)$$

$$= d(v, w) \quad \% \text{ definition of distance} \quad (2.85)$$

□

**Proposition 9.** *The modulus of eigenvalues of unitary matrix is 1.*

**Proposition 10.** *Unitary matrix is the transition matrix from an orthonormal basis to another orthonormal basis.*

**Physical meaning of unitary Matrix.** What does unitary really mean? As we saw, it means that it preserves the geometry. But it also means something else: If  $U$  is unitary and  $UV = V'$ , then we can easily form  $U^\dagger$  and multiply both sides of the equation by  $U^\dagger$  to get  $U^\dagger UV = U^\dagger V'$  or  $V = U^\dagger V'$ . In other words, because  $U$  is unitary, there is a related matrix that can undo the action that  $U$  performs.  $U^\dagger$  takes the result of  $U$ 's action and gets back the original vector. In the quantum world, all actions (that are not measurements) are undoable or reversible in such a manner.

**The roles of Hermitian and unitary matrices in quantum computing.** As shown in Figure 2.4, the Hermitian matrix plays an important role in the quantum measurement phrase, which decides the concrete basis to observe the final computational result  $|\psi^*\rangle$ . Once the basis ( $H_1$  or  $H_2$ ) is decided, the observation result must be probabilistically collapsed into one of the eigenvectors of the corresponding basis. The unitary matrix plays a role of action to change the state of the quantum computer. Considering its reversible property, all actions performed in quantum computing can be undone by performing an action described by  $U^\dagger$ .

The relations of identity, Hermitian, unitary, and square matrices are shown in Figure 2.5.

---

<sup>3</sup>Thanks for Yurui Wu's correction of removing the redundant left bracket in the following proof process.

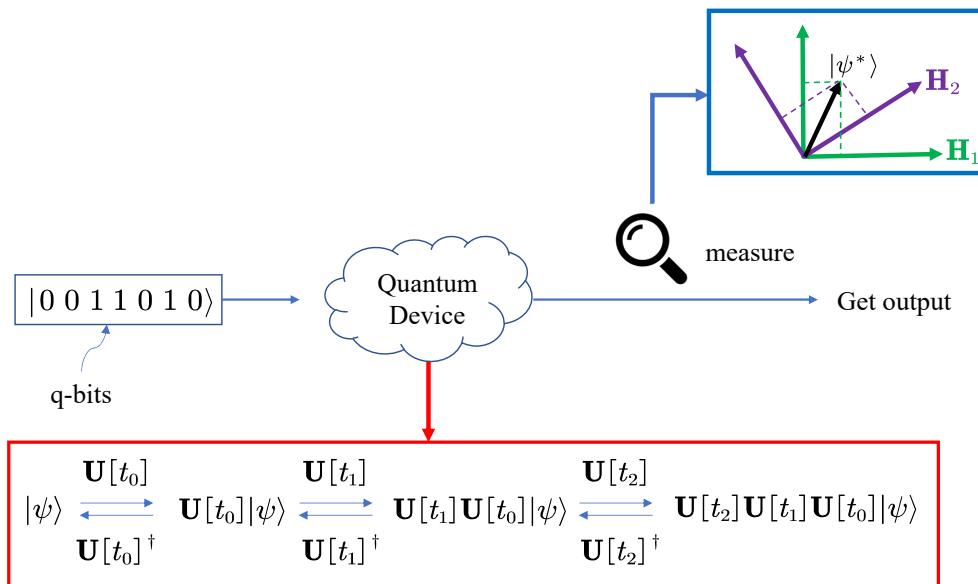


Figure 2.4: The role of Hermitian and unitary matrices.

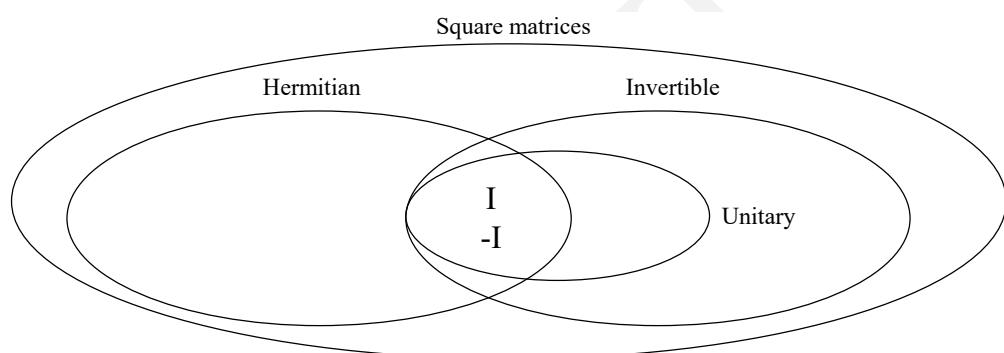


Figure 2.5: Types of matrices.

Instructor: Chao Liang

# 3 The Leap from Classic to Quantum

## 3.1 Classic Deterministic Systems

**Definition 3.1 (Discrete Dynamic System, Dynamics and States).** If  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a transformation and  $\dots, x_t, x_{t+1}, x_{t+2}, \dots$  is a sequence of vectors in  $\mathbb{R}^n$  such that  $x_{t+1} = f(x_t)$ , then we say that  $f$  and sequence  $\dots, x_t, x_{t+1}, x_{t+2}, \dots$  make up a discrete dynamical system, where function  $f$  is called dynamics and vectors  $\{x_t\}$  are called states.

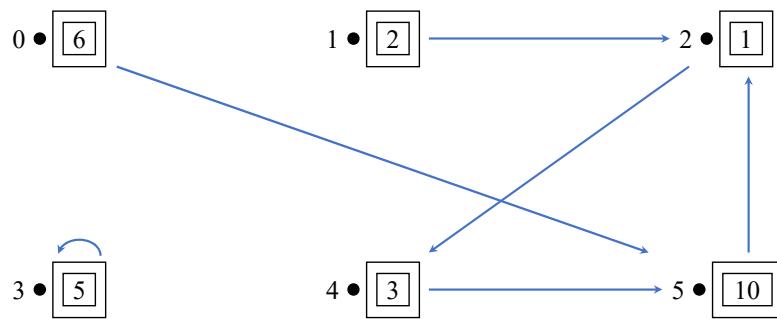


Figure 3.1: Classic billiards.

### Example 3.1: Classic Billiards

Let's consider a simple system described by a **simple (unweighted) directed graph** together with some toy marbles. There be 6 vertices in a graph and a total of 27 marbles. We might place 6 marbles on vertex 0, 2 marbles on vertex 1, and the rest as described by Figure 3.1.

We shall denote its **deterministic state** as  $x = [6, 2, 1, 5, 3, 10]^\top$ , and its **dynamics**

as a **Boolean adjacency matrix**  $M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$  where  $M(i, j) = 1$  if and only if there is an arrow from vertex  $j$  to vertex  $i$ .

The **state evolvement** can be represented as matrix multiplication:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t) = \mathbf{M}\mathbf{x}_t \quad (3.1)$$

The multiple step dynamics can be written as Boolean matrix multiplication:

$$\mathbf{M}^2(i, j) = \bigvee_{k=0}^{n-1} \mathbf{M}(i, k) \wedge \mathbf{M}(k, j) \quad (3.2)$$

where  $\vee$  and  $\wedge$  represent Boolean “OR” and “AND” operators, and  $\mathbf{M}^2(i, j) = 1$  if and only if there is a path of length 2 from vertex  $j$  to vertex  $i$  as shown in Figure 3.2.

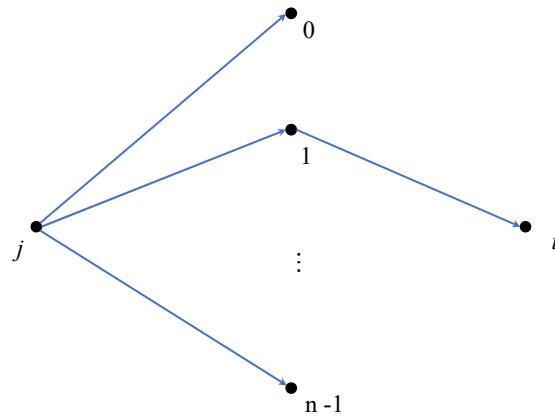


Figure 3.2: The 2-step state transition

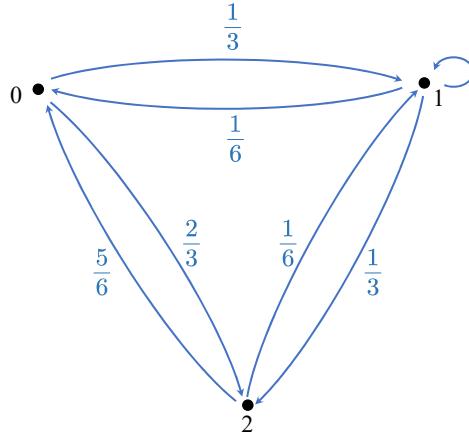
## 3.2 Probabilistic Systems

The **state** of a probabilistic system is composed of probabilistic entries, and the sum of all entries is 1.

### Example 3.2: A Three-Vertex Graph

$$\mathbf{x} = \left[ \frac{1}{5}, \frac{3}{10}, \frac{1}{2} \right]^\top$$

- one-fifth chance that the marble is on vertex 0;
- three-tenths chance that the marble is on vertex 1;
- half chance that the marble is on vertex 2.



The dynamics matrix for this graph is

$$M = \begin{bmatrix} 0 & \frac{1}{6} & \frac{5}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix}$$

Figure 3.3: Probabilistic system.

The **dynamics** of a probabilistic system is described by a directed (probabilistic) weighted graph, where several arrows shooting out of each vertex with real numbers between 0 and 1 as weights as shown in Figure 3.3. The corresponding matrix is called **doubly stochastic matrix**, which has the following two properties:

- The column sum, *i.e.*, the sum of all weights leaving a vertex, is 1;
- The row sum, *i.e.*, the sum of all weights entering a vertex, is 1.

The **state evolvement**. If we have  $x_t$  expressing the probability of the position of the marble at time  $t$  and  $M$  expressing the probability of the way the marble moves around, then  $x_{t+1} = Mx_t$  is expressing the probability of the marbles location at time  $t + 1$ .

The **multiple step dynamics** of probabilistic system is formulated with matrix multiplication with probability entries (*a.k.a.*, normal matrix multiplication). Figure 3.4 shows an example of the 2-step dynamics.

$$M^2(i, j) = \sum_{k=0}^{n-1} M(i, k)M(k, j)$$

where  $M^2(i, j)$  = the probability of going from vertex  $j$  to vertex  $i$  in 2 time clicks.

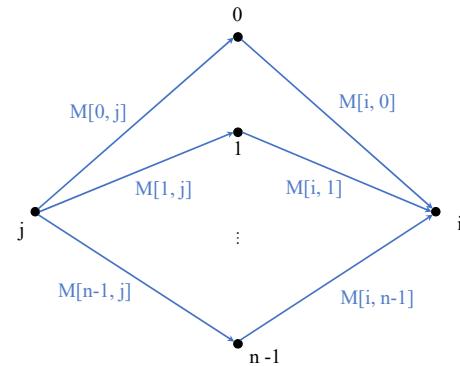
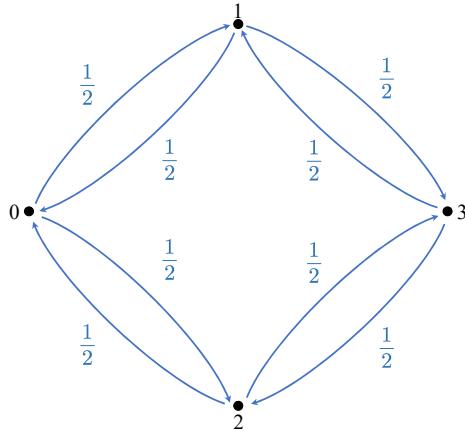


Figure 3.4: The 2-step dynamics in the probabilistic system.



The dynamics matrix for this graph is

$$\mathbf{M} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Figure 3.5: Stochastic billiard.

**Example 3.3: Stochastic Billiard**

Consider a stochastic billiard with dynamics shown in Figure 3.5 and initial state  $\mathbf{x}_0$ , its state evolvement procedure exhibits periodic cycles as follows:

$$\begin{aligned} \mathbf{x}_0 = [1 \ 0 \ 0 \ 0]^T &\xrightarrow{\mathbf{M}} \mathbf{x}_1 = [0 \ \frac{1}{2} \ \frac{1}{2} \ 0]^T \xrightarrow{\mathbf{M}} \mathbf{x}_2 = [\frac{1}{2} \ 0 \ 0 \ \frac{1}{2}]^T \\ &\xrightarrow{\mathbf{M}} \mathbf{x}_3 = \mathbf{x}_1 \xrightarrow{\mathbf{M}} \mathbf{x}_4 = \mathbf{x}_2 \xrightarrow{\mathbf{M}} \dots \end{aligned} \quad (3.3)$$

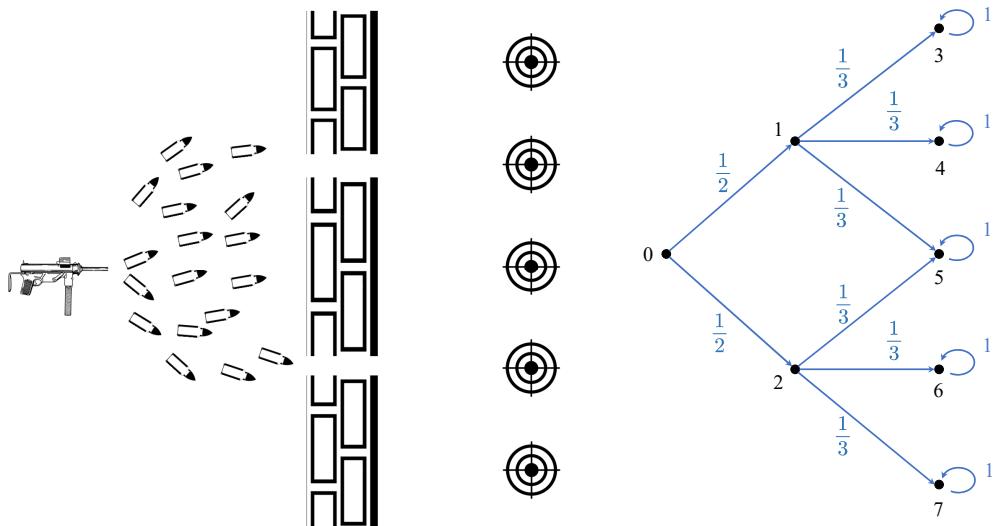


Figure 3.6: probabilistic double slit experiment

### Example 3.4: Probabilistic Double Silt Experiment

Assume a virtual double silt experiment as shown in Figure 3.6. The bullets are fired from the machine-gun, pass through two narrow slits in the wall, and eventually land on the targets behind the wall. Its dynamics matrix can be formulated as

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and accordingly, its 2-step dynamics can be computed by matrix multiplication:

$$\mathbf{M}^2 = \mathbf{M} \times \mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{6} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Hence, given an initial state  $\mathbf{x}_0 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^\top$ , its 2-step transition state  $\mathbf{x}_2 = \mathbf{M}^2 \mathbf{x}_0 = [0 \ 0 \ 0 \ \frac{1}{6} \ \frac{1}{6} \ \frac{1}{3} \ \frac{1}{6} \ \frac{1}{6}]^\top$ .

Note that the probability of the bullets landing on the middle target is the largest, i.e.,  $\frac{1}{3}$ . This is consistent with our knowledge because both routes can reach this target, meaning a summation of probabilities.

## 3.3 Quantum Systems

The **state** of a quantum system is composed of quantum entries (complex values), whose modulus square represents the probability, and the sum of modulus squared of all entries is 1.

**Example 3.5: The state of a quantum system**

Consider a complex vector  $\mathbf{x} = \left[ \frac{1}{\sqrt{3}}, \frac{2i}{\sqrt{15}}, \sqrt{\frac{2}{5}} \right]^\top$ , since  $\mathbf{x}^\dagger \mathbf{x} = \frac{1}{3} + \frac{4}{15} + \frac{2}{5} = 1$ , it is a qualified state vector of a quantum system.

The **dynamics** of a quantum system also has two representations. One is the graph form, which can be described by a directed (complex) weighted graph. The other is the matrix form, which corresponds to a special unitary matrix whose modulus square is a doubly stochastic matrix as exemplified in Figure 3.7.

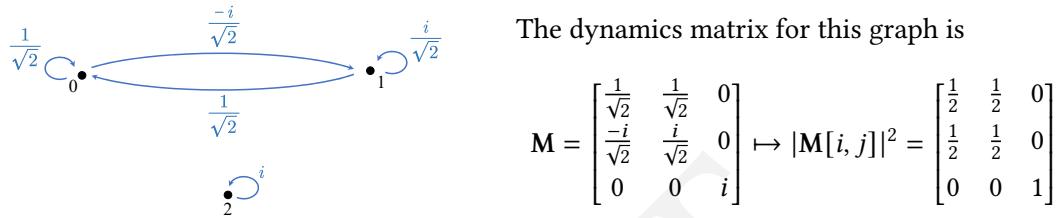


Figure 3.7: Quantum system.

Table 3.1 gives a detailed comparison of three systems in terms their states and dynamics. In particular, the dynamics is represented in two different forms, which are graph (Gra.) and matrix (Mat.), respectively.

Table 3.1: Comparison of three systems.

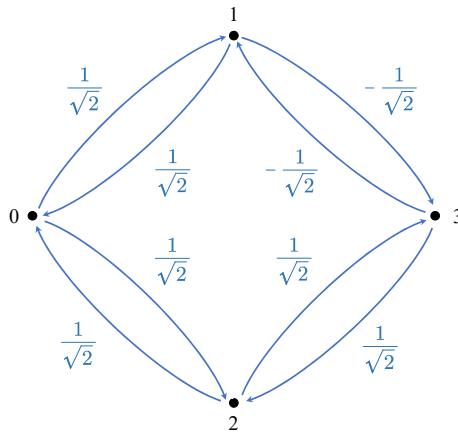
	Classic Deterministic System	Probabilistic System	Quantum System
State	$\mathbf{x} = [x_1, x_2, x_3]^\top$ $x_i \in \mathbb{N}$	$\mathbf{x} = [p_1, p_2, p_3]^\top$ $p_i \in [0, 1], \sum_i p_i = 1$	$\mathbf{x} = [c_1, c_2, c_3]^\top$ $c_i \in \mathbb{C}, \sum_i  c_i ^2 = 1$
Dynamics	Gra. Directed unweighted graph	Directed (probabilistic) weighted graph	Directed (complex) weighted graph
Mat.	Boolean adjacency matrix	Doubly stochastic matrix	Unitary matrix whose modulus squares is a doubly stochastic matrix

The **state evolvement** is formulated as matrix multiplication  $\mathbf{x}_{t+1} = \mathbf{M}\mathbf{x}_t$ .

The **forward dynamics** and **backward dynamics** can be represented as a matrix  $\mathbf{M}$  and its adjoint  $\mathbf{M}^\dagger$  as shown in Table 3.2. This means that if you perform some operation  $\mathbf{x} \mapsto \mathbf{M}\mathbf{x}$  and then undo the operation  $\mathbf{M}^\dagger \mathbf{M}\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$ , you will find yourself (with probability 1) in the same state with which you began.

Table 3.2: Comparison of forward and backward dynamics.

	Dynamics graph	Dynamics matrix
Forward dynamics		$\mathbf{M} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & i \end{bmatrix}$
Backward dynamics		$\mathbf{M} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & -i \end{bmatrix}$



The dynamics matrix for this graph is

$$\mathbf{M} = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

Figure 3.8: Quantum billiard.

### Example 3.6: Quantum Billiard

Consider a quantum billiard with dynamics shown in Figure 3.8 and initial state  $x_0$ , its state evolvement procedure exhibits periodic cycles as follows:

$$\begin{aligned}
 x_0 = [1 \ 0 \ 0 \ 0]^\top &\xrightarrow{\mathbf{M}} x_1 = [0 \ \frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} \ 0]^\top \\
 &\xrightarrow{\mathbf{M}} x_2 = \textcolor{red}{x}_0 \xrightarrow{\mathbf{M}} x_3 = \textcolor{blue}{x}_1 \xrightarrow{\mathbf{M}} \dots
 \end{aligned} \tag{3.4}$$

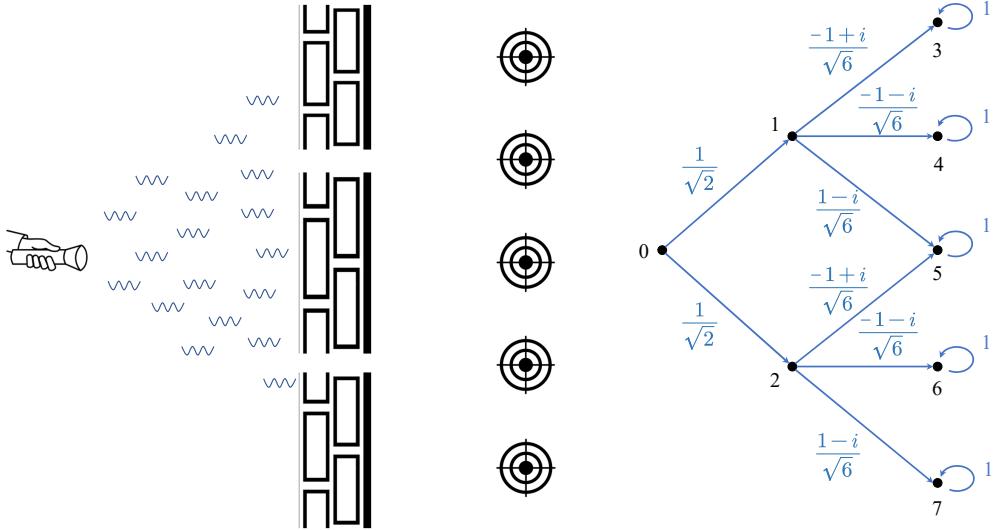


Figure 3.9: Double silt experiment

### Example 3.7: Double Silt Experiment

Given a double silt experiment as shown in Figure 3.9. The photons are ejected from the flashlight, pass through two narrow slits in the wall, and eventually land on the screens behind the wall. Its 1-step and 2-step dynamics matrices are respectively

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{-1+i}{\sqrt{6}} & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1-i}{\sqrt{6}} & \frac{-1+i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1-i}{\sqrt{6}} & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mapsto M^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-1+i}{\sqrt{12}} & \frac{-1+i}{\sqrt{6}} & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{-1-i}{\sqrt{12}} & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 & 0 \\ \textcolor{red}{0} & \frac{1-i}{\sqrt{6}} & \frac{-1+i}{\sqrt{6}} & 0 & 0 & 1 & 0 & 0 \\ \frac{-1-i}{\sqrt{12}} & 0 & \frac{-1-i}{\sqrt{6}} & 0 & 0 & 0 & 1 & 0 \\ \frac{-1+i}{\sqrt{12}} & 0 & \frac{1-i}{\sqrt{6}} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that  $M(5, 0) = \frac{1}{\sqrt{2}}(\frac{1-i}{\sqrt{6}}) + \frac{1}{\sqrt{2}}(\frac{-1+i}{\sqrt{6}}) = \frac{1-i}{\sqrt{12}} + \frac{-1+i}{\sqrt{12}} = \frac{0}{\sqrt{12}} = \textcolor{red}{0}$ , which is called **interference** phenomenon.

**Superposition.** Let the state of the system be given by  $x = [c_0, c_1, \dots, c_{n1}]^\top \in \mathbb{C}^n$ . It is incorrect to say that the probability of the photons being in position  $k$  is  $|c_k|^2$ . Rather, to be in

state  $x$  means that the particle is in some sense in *all* positions simultaneously. The photon passes through the top slit and the bottom slit simultaneously, and when it exits both slits, it can cancel itself out. A photon is not in a single position, rather it is in many positions, a superposition.

**Measurement and collapse.** The reason we see particles in one particular position is because we have performed a measurement. When we measure something at the quantum level, the quantum object that we have measured is no longer in a superposition of states, rather it collapses to a single classical state. So we have to redefine what the state of a quantum system is: a system is in state  $x$  means that *after measuring* it, it will be found in position  $i$  with probability  $|c_i|^2$ .

**Power of quantum computing.** It is exactly this superposition of states that is the real power behind quantum computing. Classical computers are in one state at every moment. Imagine putting a computer in many different classical states simultaneously and then processing with all the states at once. This is the ultimate in parallel processing!

Instructor: Chao Liang

## 4 Quantum Mechanics

**Double-slit experiment** [play the animation video<sup>1</sup> shown in Figure 4.1]: the double-slit interference experiment can be done with a single photon. Moreover, if we place a measurement device behind the double slits, the interference phenomenon will disappear (it seems the photon knows it is being watched). How is one to understand this phenomenon?

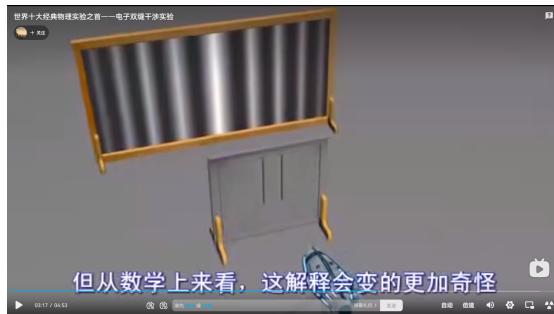


Figure 4.1: Double-slit experiment

**Explanations:** two important concepts can be derived from the above experiment:

- **Superposition.** Let the state of the quantum system be given by  $X = [c_0, c_1, \dots, c_{n-1}]^\top \in \mathbb{C}^n$ . It is incorrect to say that the probability of the photons being in position  $k$  is  $\|c_k\|^2$ . Rather, to be in state  $X$  means that the particle is in some sense *in all positions simultaneously*. To explain the above double-slit experiment, the photon passes through the top slit and the bottom slit simultaneously, and when it exists both slits, it can cancel itself out. A photon is not in a single position, rather it is in many positions, a superposition.
- **Measurement.** Seeing things existing in many positions simultaneously is counter-intuitive. Our daily-life experience tells us that things are in one position or (exclusive or!) another. How can this be? The reason we see particles in one particular position is because we have performed a measurement. When we measure something at the quantum level, the quantum object that we have measured is no longer in a superposition of states, rather it collapses to a single classical state. So we have to redefine what the state of a quantum system is: a system is in state  $X$  means that after measuring it, it will be found in position  $k$  with probability  $|c_k|^2$ .

In the following discussion, superposition and measurement are two important and fundamental concepts rooted in quantum mechanics.

<sup>1</sup> [https://www.bilibili.com/video/BV1Yx41127fG/?spm\\_id\\_from=333.337.search-card.all.click](https://www.bilibili.com/video/BV1Yx41127fG/?spm_id_from=333.337.search-card.all.click)

## 4.1 Quantum States

### 4.1.1 Case 1: position on a line

Consider a subatomic particle on a line where it can only be detected at one of the equally spaced points  $\{x_0, x_1, \dots, x_{n-1}\}$  shown in Figure 4.2.



Figure 4.2: Positions in a line

The particle being at the point  $x_k$  is denoted as  $|x_k\rangle$ , using the Dirac **ket** notation  $|\cdot\rangle$ . To each of these  $n$  **basic states**, we shall associate a column vector:

$$\begin{aligned} |x_0\rangle &\mapsto [1, 0, \dots, 0]^\top \\ |x_1\rangle &\mapsto [0, 1, \dots, 0]^\top \\ &\vdots \\ |x_{n-1}\rangle &\mapsto [0, 0, \dots, 1]^\top \end{aligned} \tag{4.1}$$

The **state** of the particle  $|\psi\rangle$  is a linear combination of  $|x_0\rangle, |x_1\rangle, \dots, |x_{n-1}\rangle$ , by suitable complex weights,  $c_0, c_1, \dots, c_{n-1}$  known as **complex amplitudes**,

$$|\psi\rangle = c_0|x_0\rangle + c_1|x_1\rangle + \dots + c_{n-1}|x_{n-1}\rangle = \sum_{k=0}^{n-1} c_k|x_k\rangle \tag{4.2}$$

We say that the state  $|\psi\rangle$  is a **superposition** of the basic states.  $|\psi\rangle$  represents the particle as being simultaneously in all  $\{x_0, x_1, \dots, x_{n-1}\}$  locations, or a blending of all the  $|x_k\rangle$ .

Thus, every state of our system can be represented by an element of  $\mathbb{C}^n$  as:

$$|\psi\rangle = [c_0, c_1, \dots, c_{n-1}]^\top \tag{4.3}$$

The norm square of the complex number  $c_k$  divided by the norm squared of  $|\psi\rangle$ , called **probability amplitude**  $\bar{c}_k$ , will tell us the probability that, after observing the particle, we will detect it at the point  $x_k$ :

$$p(x_k) = |\bar{c}_k|^2 = \left(\frac{|c_k|}{\|\psi\|}\right)^2 = \frac{|c_k|^2}{\|\psi\|^2} = \frac{|c_k|^2}{\sum_k |c_k|^2} \tag{4.4}$$

Observe that  $p(x_k)$  is always a positive real number and  $0 \leq p(x_k) \leq 1$ , as any genuine probability should be.

When  $|\psi\rangle$  is observed, we will find it in one of the basic states. We might write it as:

$$|\psi\rangle \rightsquigarrow |x_k\rangle \tag{4.5}$$

The probability of obtaining  $|x_k\rangle$  after observing  $|\psi\rangle$  is  $p(x_k)$  where  $k \in \{0, 1, \dots, n-1\}$ .

Two typical operations of ket vectors in the Hilbert space:

- addition:

$$\begin{aligned} |\psi\rangle + |\psi'\rangle &= (c_0 + c'_0) |x_0\rangle + (c_1 + c'_1) |x_1\rangle + \cdots + (c_{n-1} + c'_{n-1}) |x_{n-1}\rangle \\ &= [c_0 + c'_0, c_1 + c'_1, \dots, c_{n-1} + c'_{n-1}]^\top \end{aligned} \quad (4.6)$$

- scalar multiplication:

$$c |\psi\rangle = cc_0 |x_0\rangle + cc_1 |x_1\rangle + \cdots + cc_{n-1} |x_{n-1}\rangle = [cc_0, cc_1, \dots, cc_{n-1}]^\top \quad (4.7)$$

It is worthy noting that a ket's length does not matter as far as physics goes. In other words, the ket  $2|\psi\rangle$  describes *the same physical system* as  $|\psi\rangle$ .

#### 4.1.2 Case 2: single-particle spin system

**Stern-Gerlach experiment** [play the animation video<sup>2</sup> shown in Figure 4.3]: the magnetic field splits the beam of electrons into two streams, found either at the top of the screen or at the bottom, *but none in between!* Conclusion: when the spinning particle is measured in a given direction, it can only be found in two states, *i.e.*, it spins either clockwise or anticlockwise.

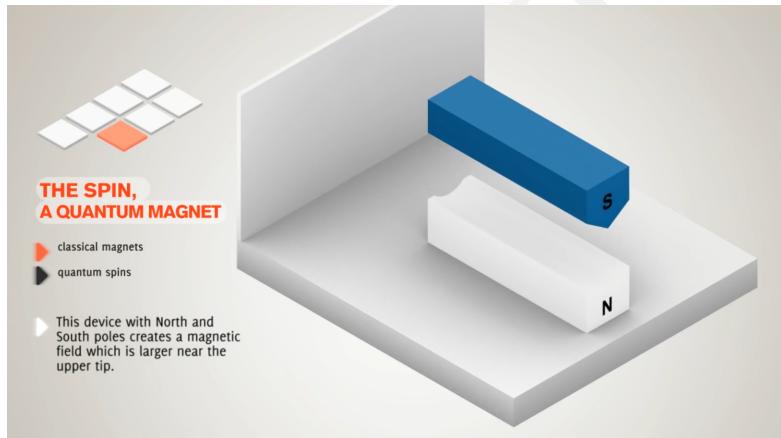


Figure 4.3: Stern-Gerlach experiment

Next, we put the two Stern-Gerlach apparatus (SGA) in different angles, and discuss the observation results.

- **Step 1: 0° SGA experiment.** Figure 4.4 shows the particle's spin direction and SGA's placement orientation before measurement (sub-figure A) and after measurement (sub-figure B). After measurement, the particle is *prepared in state*  $\sigma_z = +1$ . If the particle state is not perturbed, and the SGA's placement orientation is kept the same, the following measurements will always have the same results.

<sup>2</sup> [https://www.bilibili.com/video/BV1ta4y1a7fp?from=search&seid=2882474434643948118&spm\\_id\\_from=333.337.0.0](https://www.bilibili.com/video/BV1ta4y1a7fp?from=search&seid=2882474434643948118&spm_id_from=333.337.0.0)

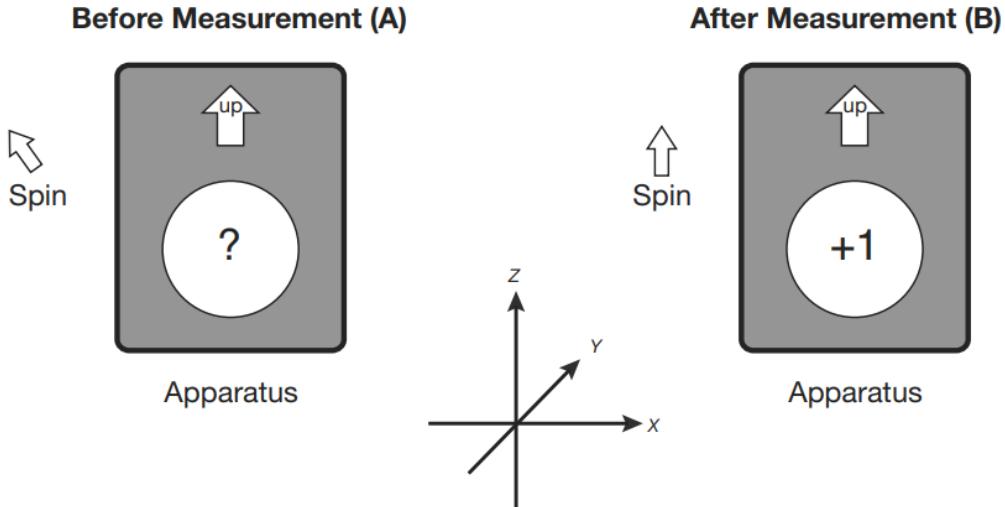


Figure 4.4:  $0^\circ$  Stern-Gerlach apparatus (SGA) experiment.

- **Step 2:  $180^\circ$  SGA experiment.** After preparing the spin by measuring it with SGA, we turn the apparatus upside down and then measure  $\sigma_z$  again (Figure 4.5). What we find is that if we originally prepared  $\sigma_z = +1$ , the upside down apparatus records  $\sigma_z = -1$ .

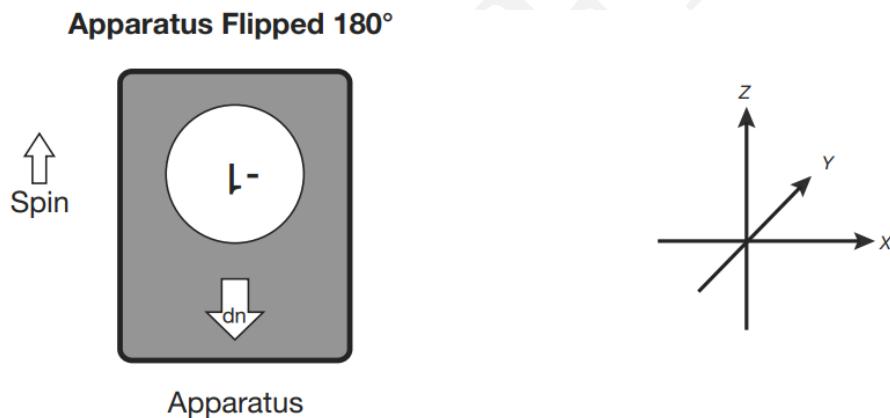


Figure 4.5:  $180^\circ$  Stern-Gerlach apparatus (SGA) experiment.

- **Step 3:  $90^\circ$  SGA experiment.** So far, there is still no difference between classical physics and quantum physics. The difference only becomes apparent when we rotate the apparatus through an arbitrary angle, say  $\pi/2$  radians (90 degrees): (1) The apparatus begins in the upright position (with the up-arrow along the z axis). A spin is prepared with  $\sigma_z = +1$ . (2) rotate the SGA so that the up-arrow points along the x axis (Figure 4.6). (3) make a measurement of what is presumably the x component of the spin,  $\sigma_x$ . The apparatus gives either  $\sigma_x = +1$  or  $\sigma_x = -1$ , and the numbers of  $\sigma_x = +1$  events and  $\sigma_x = -1$  events are statistically equal. In other words, the average value of  $\sigma_x$  is zero.

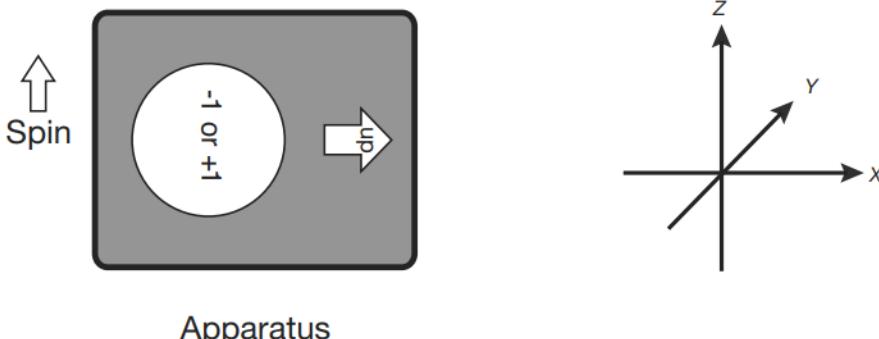
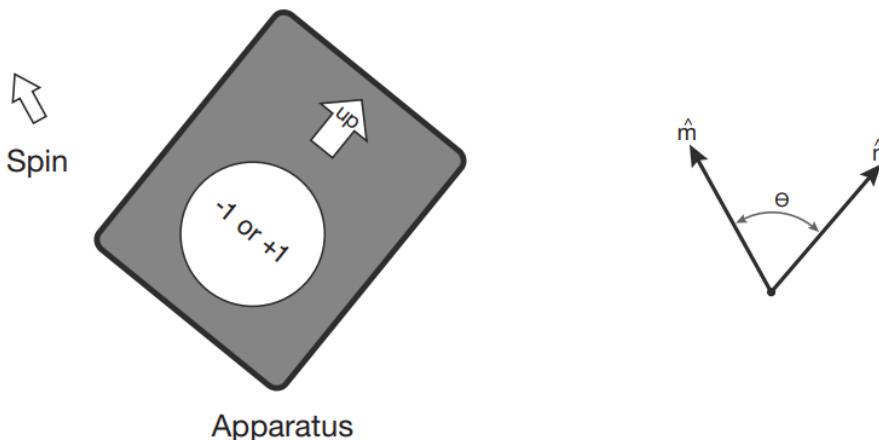
**Apparatus Rotated 90°**

Figure 4.6: 90° Stern-Gerlach apparatus (SGA) experiment.

- **Step 4:  $\theta^\circ$  SGA experiment.** Now lets do the whole thing over again, but instead of rotating the SGA to lie on the x axis, rotate it to an arbitrary direction along the unit vector  $\hat{n}$ . If  $\hat{n}$  lies at an angle  $\theta$  with respect to  $\hat{m}$  (the prepared spin direction of the particle shown in Figure 4.7), each time we do the experiment we get  $\sigma = +1$  or  $\sigma = -1$ , and the average value of the measurement  $\sigma$  is  $\hat{n} \cdot \hat{m} = \cos \theta$ .

**Apparatus Rotated by an Arbitrary Angle**Figure 4.7:  $\theta^\circ$  Stern-Gerlach apparatus (SGA) experiment.

Let the probability of measuring  $\sigma$  is  $p(\sigma)$ , thus we have the following set of equations:

$$\begin{cases} p(\sigma = +1) \cdot (+1) + p(\sigma = -1) \cdot (-1) = \cos \theta \\ p(\sigma = +1) + p(\sigma = -1) = 1 \end{cases} \quad (4.8)$$

From which we can calculate the measuring probabilities of  $\sigma = +1$  and  $\sigma = -1$ .

**State representation.** According to the previous Stern-Gerlach experiment, particle's spin state can be represented as the superposition of basic states in a specific direction.

- **state along the z-axis** can be represented, according to the previous  $0^\circ$  and  $180^\circ$  SGA experiments, as

$$|\psi\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle \quad (4.9)$$

where  $\alpha_u = \langle u|\psi\rangle$  and  $\alpha_d = \langle d|\psi\rangle$  are the probability amplitudes meeting the following equation set:

$$\begin{cases} p(u) = \alpha_u^\dagger \alpha_u = \langle \psi|u\rangle \langle u|\psi\rangle \\ p(d) = \alpha_d^\dagger \alpha_d = \langle \psi|d\rangle \langle d|\psi\rangle \\ p(u) + p(d) = \alpha_u^\dagger \alpha_u + \alpha_d^\dagger \alpha_d = 1 \end{cases} \quad (4.10)$$

- **state along the x-axis.** According to the previous  $90^\circ$  SGA experiments, if SGA initially prepares  $|r\rangle$ , and then is rotated to measure  $\sigma_z$ , there will be equal probabilities for *up* and *down*. Thus,  $\alpha_u^\dagger \alpha_u$  and  $\alpha_d^\dagger \alpha_d$  must both be equal to  $\frac{1}{2}$ . A simple vector that satisfies this rule is:

$$|r\rangle = \frac{1}{\sqrt{2}} |u\rangle + \frac{1}{\sqrt{2}} |d\rangle \quad (4.11)$$

Considering the exclusive constraints between  $r$  and  $l$ , i.e.,  $\langle r|l\rangle = \langle l|r\rangle = 0$ , we have:

$$|l\rangle = \frac{1}{\sqrt{2}} |u\rangle - \frac{1}{\sqrt{2}} |d\rangle \quad (4.12)$$

- **state along the y-axis.** Represent spin states along the y-axis is more complicate because of the following constraints:

$$\begin{cases} \langle i|o\rangle = 0 \\ \langle i|u\rangle \langle u|i\rangle = \frac{1}{2}, \quad \langle i|d\rangle \langle d|i\rangle = \frac{1}{2} \\ \langle o|u\rangle \langle u|o\rangle = \frac{1}{2}, \quad \langle o|d\rangle \langle d|o\rangle = \frac{1}{2} \\ \langle i|l\rangle \langle l|i\rangle = \frac{1}{2}, \quad \langle i|r\rangle \langle r|i\rangle = \frac{1}{2} \\ \langle o|l\rangle \langle l|o\rangle = \frac{1}{2}, \quad \langle o|r\rangle \langle r|o\rangle = \frac{1}{2} \end{cases} \quad (4.13)$$

From which a set of proper representation of state along the y-axis is

$$\begin{cases} |i\rangle = \frac{1}{\sqrt{2}} |u\rangle + \frac{i}{\sqrt{2}} |d\rangle \\ |o\rangle = \frac{1}{\sqrt{2}} |u\rangle - \frac{i}{\sqrt{2}} |d\rangle \end{cases} \quad (4.14)$$

**Transition amplitude.** Suppose the start state is  $|\psi\rangle = [c_0, c_1, \dots, c_{n-1}]^\top$  the end state is

$|\psi'\rangle = [c'_0, c'_1, \dots, c'_{n-1}]^\top$ , the transition amplitude is defined as

$$\langle\psi'|\psi\rangle = [\overline{c'_0}, \overline{c'_1}, \dots, \overline{c'_{n-1}}] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \sum_{k=0}^{n-1} \overline{c'_k} \times c_k \quad (4.15)$$

where  $\langle\psi'| = [\overline{c'_0}, \overline{c'_1}, \dots, \overline{c'_{n-1}}]$  is called the **bra** vector of the corresponding ket vector  $|\psi'\rangle$ .

We can represent the start state, the ending state, and the amplitude of going from the first to the second as the decorated arrow:

$$|\psi\rangle \xrightarrow{\langle\psi'|\psi\rangle} |\psi'\rangle \quad (4.16)$$

#### Note 4.1: Transition amplitude

The transition amplitude between two states may be zero. In fact, that happens precisely when the two states are orthogonal to one another. This simple fact hints at the physical content of orthogonality: orthogonal states are as far apart as they can possibly be. We can think of them as *mutually exclusive alternatives*: for instance, an electron can be in an arbitrary superposition of spin up and down, but after we measure it in the z direction, it will always be *either* up or down, never both up and down.

We can express  $|\psi\rangle$  in the orthonormal basis  $|b_0\rangle, |b_1\rangle, \dots, |b_{n-1}\rangle$  as

$$|\psi\rangle = b_0|b_0\rangle + b_1|b_1\rangle + \dots + b_{n-1}|b_{n-1}\rangle \quad (4.17)$$

where the probability amplitude is also the transition amplitude, i.e.,  $b_j = \langle b_j|\psi\rangle$ , and that  $|b_0|^2 + |b_1|^2 + \dots + |b_{n-1}|^2 = 1$ .

## 4.2 Observables and Measuring

### 4.2.1 Basic concepts

**Specification of a physical system:** On the one hand, its **state space**, i.e., the collection of all the states (discussed in the previous section), and on the other hand, **observable set**, i.e., the physical quantities observed in each state of the state space.

**Observable:** A specific question we pose to the system. For example, if the system is currently in some given state  $|\psi\rangle$ , which values can we possibly observe?

**Measuring:** The process of asking a specific question and receiving a definite answer.

The measurement operations in classic and quantum physics are inherently different. Figure 4.8 shows two key differences.

Classic physics	Quantum physics
<ul style="list-style-type: none"> <li>the act of measuring would <b>not change</b> the system state</li> <li>the result of a measurement on a well-defined state is <b>deterministic</b></li> </ul>	<ul style="list-style-type: none"> <li>the act of measuring would <b>change</b> the system state</li> <li>the result of a measurement on a well-defined state is <b>nondeterministic</b></li> </ul>

Figure 4.8: Comparisons of measurement operations in classic and quantum physics.

#### 4.2.2 The principles

In this part, we present the five principles about the observing and measurement. The first four principles do not involve the evolution of state-vectors with time.

- Principle 1:** The observable or measurable quantities of quantum mechanics are represented by linear operators  $\Omega$ , which must also be Hermitian<sup>3</sup>.
- Principle 2:** The possible results of a measurement are the *eigenvalues* of the operator that represents the observable. The collapsed state is the related *eigenvector* of the operator that represents the observable. If the system is in the eigenstate  $|\lambda_i\rangle$ , the result of a measurement is *guaranteed* to be  $\lambda_i$ .

##### Example 4.1: Positions on a line

In this example, the most obvious observable is **position**. As we have stated already, each observable represents a specific question we pose to the quantum system. Position asks: Where can the particle be found? Which hermitian operator corresponds to position? We are going to tell first how it acts on the basic states:

$$P(|\psi\rangle) = P(|x_i\rangle) = x_i |x_i\rangle \quad (4.18)$$

Considering  $|x_i\rangle = [0, \dots, 1, \dots, 0]^\top$  is the one-hot vector with only the  $i$ -th element equals to 1, we have

$$P = \begin{bmatrix} x_0 & 0 & \cdots & 0 \\ 0 & x_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n-1} \end{bmatrix} \quad (4.19)$$

<sup>3</sup> Two reasons why observable operators must be Hermitian: First, the eigenvalues of an operator are real, which is a necessary condition of realistic experiment. Second, the eigenvectors that represent unambiguously distinguishable results must have different eigenvalues, and must also be orthogonal (see principle 2 and 3).

### Example 4.2: Single-particle spin system

Let's recall the previous Stern-Gerlach experiment, in which a particle with vertical upward spin passing through a z-axis-directional upward SGA generates “+1” observation value and its spin state is kept upward. Meanwhile, given a particle with vertical downward spin, its SGA observation reads “-1” value and its spin state is kept downward. Hence we have the following formulations:

$$\begin{cases} \sigma_z |u\rangle = |u\rangle \\ \sigma_z |d\rangle = -1 |d\rangle \\ \langle u|d\rangle = 0 \end{cases} \quad (4.20)$$

Let  $|u\rangle = [1, 0]^\top$ ,  $|d\rangle = [0, 1]^\top$ , and  $\sigma_z \in \mathbb{R}^{2 \times 2}$ , the above set of equations can be re-formulated and the matrix of x-axis spin operator  $\sigma_x$  can be calculated

$$\begin{cases} \begin{bmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases} \Rightarrow \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (4.21)$$

Similarly, the matrices of x-axis spin operator  $\sigma_x$  and y-axis spin operator  $\sigma_y$  can be written as

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (4.22)$$

### Note 4.2: Measurement operator

There are some truths and misconception about the measurement operator:

**Truths:** (1) Operators are the things we use to calculate eigenvalues and eigenvectors; (2) Operators act on state-vectors (which are abstract mathematical objects), not on actual physical systems; (3) When an operator acts on a state-vector, it produces a new state-vector.

**Misconception:** When a measurement operator  $\Omega$  acts on a state-vector, it produces a new state-vector, but that operation is in no way the same as acting on the state with the operator  $\Phi$ . The former,  $\Omega |\psi\rangle$ , means a state collapse and the formulation only valid when  $|\psi\rangle$  is the eigenvector of  $\Omega$ . The latter,  $\Phi |\psi\rangle$ , is always valid and means that a state transition from the original state  $|\psi\rangle$  to a new state  $\Phi |\psi\rangle$ .

- **Principle 3:** Unambiguously distinguishable<sup>4</sup> states are represented by orthogonal vectors. Inner product of two states is a measure of the inability to distinguish them with certainty.
- **Principle 4:** If  $|\psi\rangle$  is the state-vector of a system, and the observable  $\Omega$  is measured, the probability to observe value  $\lambda_i$  is:

$$p(\lambda_i) = |\langle \lambda_i | \psi \rangle|^2 = \langle \psi | \lambda_i \rangle \langle \lambda_i | \psi \rangle \quad (4.23)$$

But, in general, there is no way to tell for certain which of these values will be observed. There is only a probability  $p(\lambda_i)$ , expressed in terms of the overlap of  $|\psi\rangle$  and  $|\lambda_i\rangle$ , describing that the outcome will be  $\lambda_i$ .

- **Principle 5:** The evolution of a quantum system (that is not a measurement) is given by a unitary operator or transformation, *i.e.*,  $|\psi(t+1)\rangle = U|\psi(t)\rangle$ .

### 4.2.3 The expected value of observing

Suppose that  $\lambda_0, \lambda_1, \dots, \lambda_{n-1}$  is the list of eigenvalues of a measurement operator  $\Omega$ . Let us prepare our quantum system so that it is in state  $|\psi\rangle$  and let us observe the value of  $\Omega$ . We are going to obtain one or another of the aforementioned eigenvalues. Now, let us start all over again many times, say,  $n$  times, and let us keep track of what was observed each time. At the end of our experiment, the eigenvalue  $\lambda_i$  has been seen  $p_i$  times, where  $0 \leq p_i \leq n$  (in statistical jargon, its frequency is  $p_i/n$ ). Now perform the calculation

$$\lambda_0 \times \frac{p_0}{n} + \lambda_1 \times \frac{p_1}{n} + \dots + \lambda_{n-1} \times \frac{p_{n-1}}{n} \quad (4.24)$$

If  $n$  is sufficiently large, this number (known in statistics as the estimated expected value of  $\Omega$ ) will be very close to  $\langle \Omega \rangle_\psi = \langle \Omega \psi, \psi \rangle$ .

### 4.2.4 Multiple-step observing

Before we investigate the multiple-step observing, let's first consider what happens after single-step observing. Suppose the quantum system state is  $|\psi\rangle$ , and the observing operator is  $\Omega$  (with eigenvalues  $\{\lambda_i\}$  and corresponding eigenvectors  $\{|\lambda_i\rangle\}$  as the previous section). After one-step observing, we first get an answer  $\lambda_i$  with probability  $p_i = \langle \psi | \lambda_i \rangle \langle \lambda_i | \psi \rangle$ . Then, the system's state collapse from  $|\psi\rangle$  to the corresponding eigenstate  $|\lambda_i\rangle$  as shown in Figure 4.9.

According to the above discussion, observing in the quantum world will necessarily lead to state collapse. Hence, the result of multiple-step observing depends on the observing order. Take Figure 4.10 as an example, given the quantum state  $|\psi\rangle$ , the result of two-step observing of “ $\Omega \rightarrow \Omega'$ ” is zero, but if we insert an intermediate observing, *i.e.*,  $\Omega''$ , then the result of three-step observing of “ $\Omega \rightarrow \Omega'' \rightarrow \Omega'$ ” is not zero.

---

<sup>4</sup> Two states are physically distinct if there is a measurement that can tell them apart without ambiguity. For example,  $|u\rangle$  and  $|d\rangle$  can be distinguished by measuring  $\sigma_z$ . If you are handed a spin and told that it is either in the state  $|u\rangle$  or the state  $|d\rangle$ , to find out which of the two states is the right one, all you have to do is align SGA with the z axis and measure  $\sigma_z$ .

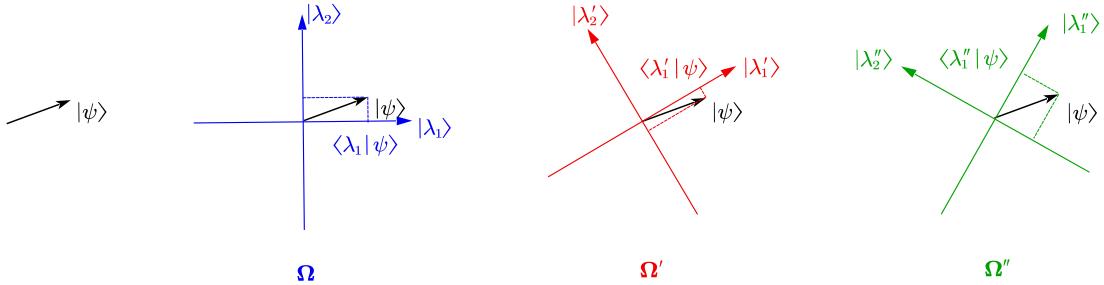


Figure 4.9: Illustration of single-step observing with probabilistic collapse under various measurement operators  $\Omega$ ,  $\Omega'$ , and  $\Omega''$ .

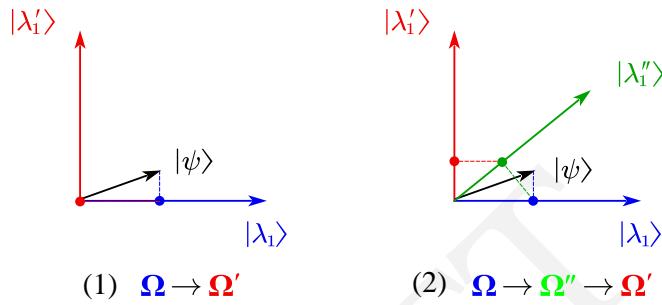


Figure 4.10: Order effects in the multiple-step observing.

### 4.3 Dynamics

The process of quantum computing can be generally divided into three steps:

- (1) Prepare an initial state  $|\psi\rangle$ ;
- (2) Apply a sequence of unitary operators to the state (see Figure 4.11);
- (3) Measure the output and get a final state.

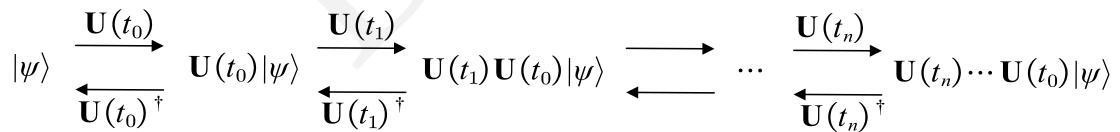


Figure 4.11: Apply a sequence of unitary operators to the state.

Instructor: Chao Liang

# 5 Quantum Gates

This section studies quantum gates. As the basis, we begin with bit and qubit. We then discuss classical gates, reversible gates and quantum gates in turn. Last, we introduce the Bell circuit and its two applications, *i.e.*, superdense coding and quantum teleportation.

## 5.1 Bits and Qubits

**Definition 5.1** (Bit). *A bit is a unit of information describing a two-dimensional classical system.*

Since a two-dimensional classical system has two orthogonal states, hence the bit 0 and bit 1 can be represented as two  $2 \times 1$  binary vectors, *i.e.*,

$$\text{bit-0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \text{bit-1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (5.1)$$

where bit-0 and bit-1 equal to  $|0\rangle$  and  $|1\rangle$  for the convenience of the following discussion.

**Definition 5.2** (Qubit). *A quantum bit or a qubit is a unit of information describing a two-dimensional quantum system.*

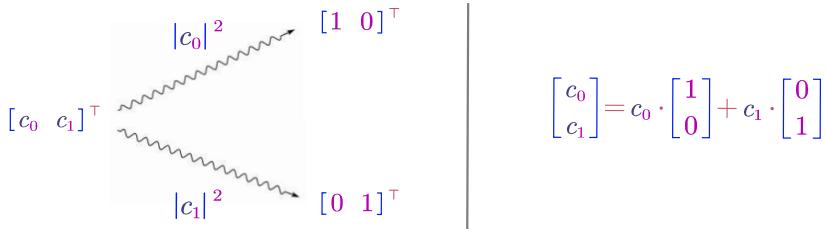


Figure 5.1: Relation between qubit and bit.

The only difference between the above two definitions is the property of the two-dimensional system. The former is a classic system, while the latter is a quantum system. This means that the state of a qubit lies in the complex vector space satisfying the normalization constraint, *i.e.*,

$$|\varphi\rangle = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \quad (5.2)$$

where  $|c_0|^2 + |c_1|^2 = 1$ . Whenever we measure a qubit, it automatically becomes a bit with corresponding collapse probability of  $|c_0|^2$  for bit 0 and  $|c_1|^2$  for bit 1 as shown in Figure 5.1. So we shall never “see” a general qubit.

**Definition 5.3** (Byte). *The byte is a unit of digital information that most commonly consists of eight bits.*

For example, given a byte including eight bits 01101011, the vector representation of these eight bits is:  $[10]^\top, [01]^\top, [01]^\top, [10]^\top, [01]^\top, [10]^\top, [01]^\top, [01]^\top$ . Recall the tensor product in composite system, the state of a byte equals to  $|0\rangle \otimes |1\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle \otimes |1\rangle$ , which is a discrete element of the complex vector space  $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ .

**Definition 5.4** (Qubyte). *The qubyte is a unit of quantum information that consists of eight qubits.*

Similar, the state of a qubyte is the tensor product of eight qubits, i.e.,  $|\varphi_0\rangle \otimes |\varphi_1\rangle \otimes |\varphi_2\rangle \otimes |\varphi_3\rangle \otimes |\varphi_4\rangle \otimes |\varphi_5\rangle \otimes |\varphi_6\rangle \otimes |\varphi_7\rangle$ , which is a continuous element of the complex vector space  $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ .

	Single (qu)bit	Double (qu)bits	Eight (qu)bits / (qu)byte
Classic bit(s)	$ 0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $ 1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$ 01\rangle$ $= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$ 01101011\rangle$ $= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 00000000 \\ 00000001 \\ \vdots \\ 01101010 \\ 01101011 \\ 01101100 \\ \vdots \\ 11111110 \\ 11111111 \end{bmatrix}$
Quantum bit(s)	$ \varphi\rangle$ $= \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = c_0 0\rangle + c_1 1\rangle$ $\text{where } \sum_{i=0}^1  c_i ^2 = 1$	$ \varphi_0\varphi_1\rangle$ $= \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_0 00\rangle + c_1 01\rangle + c_2 10\rangle + c_3 11\rangle$ $\text{where } \sum_{i=0}^3  c_i ^2 = 1$	$ \varphi_0\varphi_1\varphi_2\varphi_3\varphi_4\varphi_5\varphi_6\varphi_7\rangle$ $= \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{106} \\ c_{107} \\ c_{108} \\ \vdots \\ c_{254} \\ c_{255} \end{bmatrix} = \begin{bmatrix} 00000000 \\ 00000001 \\ \vdots \\ 01101010 \\ 01101011 \\ 01101100 \\ \vdots \\ 11111110 \\ 11111111 \end{bmatrix}$ $= c_0 00000000\rangle + \dots + c_{255} 11111111\rangle$ $\text{where } \sum_{i=0}^{255}  c_i ^2 = 1$

Figure 5.2: State vectors of single (qu)bit, double (qu)bits and (qu)byte.

The state vectors of single (qu)bit, double (qu)bits and (qu)byte are shown in Figure 5.2. Let's compare the byte and qubyte, both of them are represented as a 256-dimensional vector. But the former, classic byte, contains only 8 binary numbers, while the latter, quantum byte, contains  $2^8 = 256$  complex numbers. This difference indicates qubyte is much more informative than the classic byte.

## 5.2 Classic Gates

**Matrix representation.** Classical logical gates are ways of manipulating bits. This section studies classical gates from the point of view of matrices. As stated in Section 5.1, we represent  $n$  input bits as a  $2^n \times 1$  vector and  $m$  output bits as a  $2^m \times 1$  vector. How should we represent our logical gates? When one multiplies a  $2^m \times 2^n$  matrix with a  $2^n \times 1$  vector, the result is a  $2^m \times 1$  vector. In symbols:

$$\underbrace{(2^m \times 2^n)}_{\text{gate}} \cdot \underbrace{(2^n \times 1)}_{\text{input}} = \underbrace{(2^m \times 1)}_{\text{output}} \quad (5.3)$$

### Example 5.1: NOT gate

Consider the NOT gate:  $A \xrightarrow{\text{NOT}} Y$

NOT gate takes as input one bit, or a  $2 \times 1$  vector, and outputs one bit, or a  $2 \times 1$  vector. NOT of  $|0\rangle$  equals  $|1\rangle$  and NOT of  $|1\rangle$  equals  $|0\rangle$ . Consider the matrix

$$\text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (5.4)$$

This matrix satisfies

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (5.5)$$

### Example 5.2: AND gate

Consider the AND gate:  $A \xrightarrow{\text{AND}} Y$

The AND gate accepts two bits and outputs one bit, hence we need a  $2^1 \times 2^2$  matrix. Consider the matrix

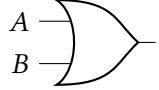
$$\text{AND} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.6)$$

This matrix satisfies  $\text{AND } |11\rangle = |1\rangle$  and  $\text{AND } |01\rangle = |0\rangle$

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (5.7)$$

### Example 5.3: OR gate

Consider the OR gate:



The OR gate similarly accepts two bits and outputs one bit, hence can be represented by a  $2^1 \times 2^2$  matrix.

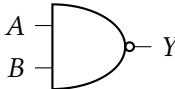
$$\text{AND} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (5.8)$$

This matrix satisfies  $\text{OR}|00\rangle = |0\rangle$  and  $\text{AND}|01\rangle = |1\rangle$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (5.9)$$

### Example 5.4: NAND gate

Consider the NAND gate:



The NAND gate similarly accepts two bits and outputs one bit, hence can be represented by a  $2^1 \times 2^2$  matrix.

$$\text{AND} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad (5.10)$$

This matrix satisfies  $\text{NOT} \times \text{AND} = \text{NAND}$

$$\text{NOT} \cdot \text{AND} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = \text{NAND} \quad (5.11)$$

**Sequential operation.** The way of thinking of NAND brings to light a general situation. When we perform a computation, we often have to carry out one operation followed by another. We call this procedure performing **sequential** operations. Take Figure 5.3a as an example, if A an operation with  $m$  input bits and  $n$  output bits, its matrix will be of size  $2^n \times 2^m$ . Say, B takes the  $n$  outputs of A as input and outputs  $p$  bits, then B is represented by a  $2^p \times 2^n$  matrix, and performing one operation sequentially followed by another operation corresponds to  $B \times A$ , which is a  $(2^p \times 2^n) \times (2^n \times 2^m) = (2^p \times 2^m)$  matrix.

**Parallel operation.** Besides sequential operations, there are parallel operations as shown in Figure 5.3b. Here we have A acting on some bits and B on others. This will be represented by  $A \otimes B$ . Let us be exact with the number of inputs and the number of outputs. A will be of size  $2^n \times 2^m$ . B will be of size  $2^{n'} \times 2^{m'}$ ,  $A \otimes B$  is of size  $2^n 2^{n'} = 2^{n+n'} \times 2^m 2^{m'} = 2^{m+m'}$ .

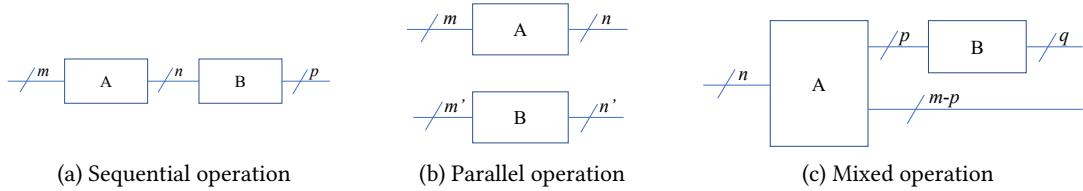


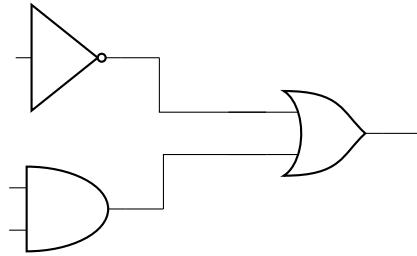
Figure 5.3: Sequential, parallel and mixed operations.

**Mixed operation.** Take Figure 5.3c as an example, let A be an operation that takes  $n$  inputs and gives  $m$  outputs. Let B take  $p < m$  of these outputs and leave the other  $m - p$  outputs alone. B outputs  $q$  bits. A is a  $2^m \times 2^n$  matrix. B is a  $2^q \times 2^p$  matrix. As nothing should be done to the  $m - p$  bits, we might represent this as the  $2^{m-p} \times 2^{m-p}$  identity matrix  $I_{m-p}$ . We do not draw any gate for the identity matrix. The entire circuit can be represented by the following matrix:

$$(B \otimes I_{m-p}) \times A \quad (5.12)$$

#### Example 5.5: Example 1 for mixed operation

Consider the circuit:



This is represented by

$$\text{OR} \times (\text{NOT} \times \text{AND}) \quad (5.13)$$

Let us see how the operations look like as matrices. We first calculate the parallel part:

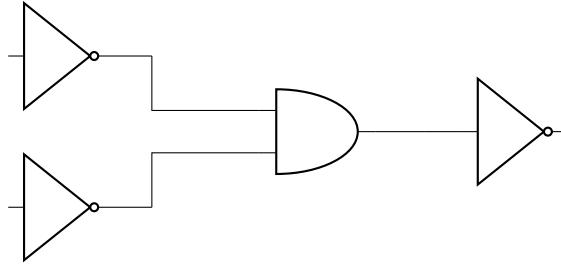
$$\text{NOT} \otimes \text{AND} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (5.14)$$

And then we calculate the whole circuit:

$$\text{OR} \times (\text{NOT} \otimes \text{AND}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.15)$$

**Example 5.6: Example 2 for mixed operation**

Consider the circuit:



This is represented by

$$\text{NOT} \times \text{AND} \times (\text{NOT} \otimes \text{NOT}) \quad (5.16)$$

Let us see how the operations look like as matrices. We first calculate the parallel part:

$$\text{NOT} \otimes \text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (5.17)$$

And then we calculate the whole circuit  $\text{NOT} \times \text{AND} \times (\text{NOT} \otimes \text{NOT})$ :

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (5.18)$$

### 5.3 Reversible Gates

In the quantum world, all operations that are not measurements are reversible and are represented by unitary matrices. The AND operation is not reversible. Given an output of  $|0\rangle$  from AND, one cannot determine if the input was  $|00\rangle$ ,  $|01\rangle$ , or  $|10\rangle$ . So from an output of the AND gate, one cannot determine the input and hence AND is not reversible. In contrast, the NOT gate and the identity gates are reversible.

Reversible gates have a history that predates quantum computing. In the 1960s, Rolf Landauer analyzed computational processes and showed that erasing information, as opposed to writing information, is what causes energy loss and heat. This notion has come to be known as the **Landauers principle**.

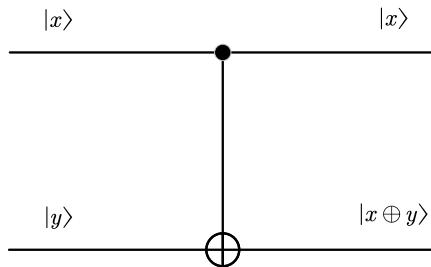
We have found that erasing information is an irreversible, energy-dissipating operation. In the 1970s, Charles H. Bennett continued along these lines of thought. If erasing information is the only operation that uses energy, then a computer that is reversible and does not erase

would not use any energy. Bennett started working on reversible circuits and programs.

A reversible circuit has exactly as many outputs as inputs. Each input can be reconstructed from the output; no bits are lost, so reversible circuits will not give off heat from bit loss.

### 5.3.1 CNOT gate

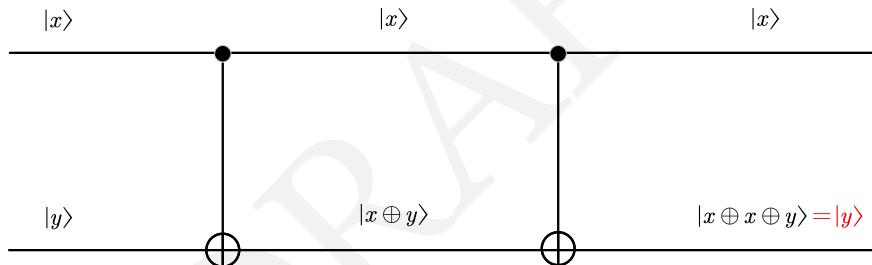
What examples of reversible gates are there? We have already seen that the identity gate and NOT gates are reversible. What else is there? Consider the following controlled-NOT gate shown in Figure 5.4 (a):



(a) Circuit of CNOT gate

00	01	10	11
00	1	0	0
01	0	1	0
10	0	0	1
11	0	0	1

(b) Matrix of CNOT gate



(c) Series circuit of CNOT gates

Figure 5.4: The circuit (a), matrix (b), and reversion (c) of CNOT gate.

This gate has two inputs and two outputs. The top input is the control bit. It controls what the output will be. If  $|x\rangle = |0\rangle$ , then the bottom output of  $|y\rangle$  will be the same as the input. If  $|x\rangle = |1\rangle$ , then the bottom output will be the opposite. If we write the top qubit first and then the bottom qubit, then the controlled-NOT gate takes  $|x, y\rangle$  to  $|x, x \oplus y\rangle$ , where  $\oplus$  is the binary “exclusive or” operation. The matrix that corresponds to this reversible gate is shown in Figure 5.4 (b).

CNOT gate can be reversed by itself as shown in Figure 5.4 (c). State  $|x, y\rangle$  goes to  $|x, x \oplus y\rangle$ , which further goes to  $|x, x \oplus (x \oplus y)\rangle$ . This last state is equal to  $|x, (x \oplus x) \oplus y\rangle$  because  $\oplus$  is associative. Because  $x \oplus x$  is always equal to 0, this state reduces to the original  $|x, y\rangle$ .

### 5.3.2 Toffoli gate

Toffoli gate extends CNOT gate's function by using two controlling bits. The bottom bit flips only when *both* of the top two bits are in state  $|1\rangle$ . We can write this operation as taking state  $|x, y, z\rangle$  to  $|x, y, z \oplus (x \wedge y)\rangle$ . The circuit and matrix representations of Toffoli gate is shown in Figure 5.5 (a) and (b).

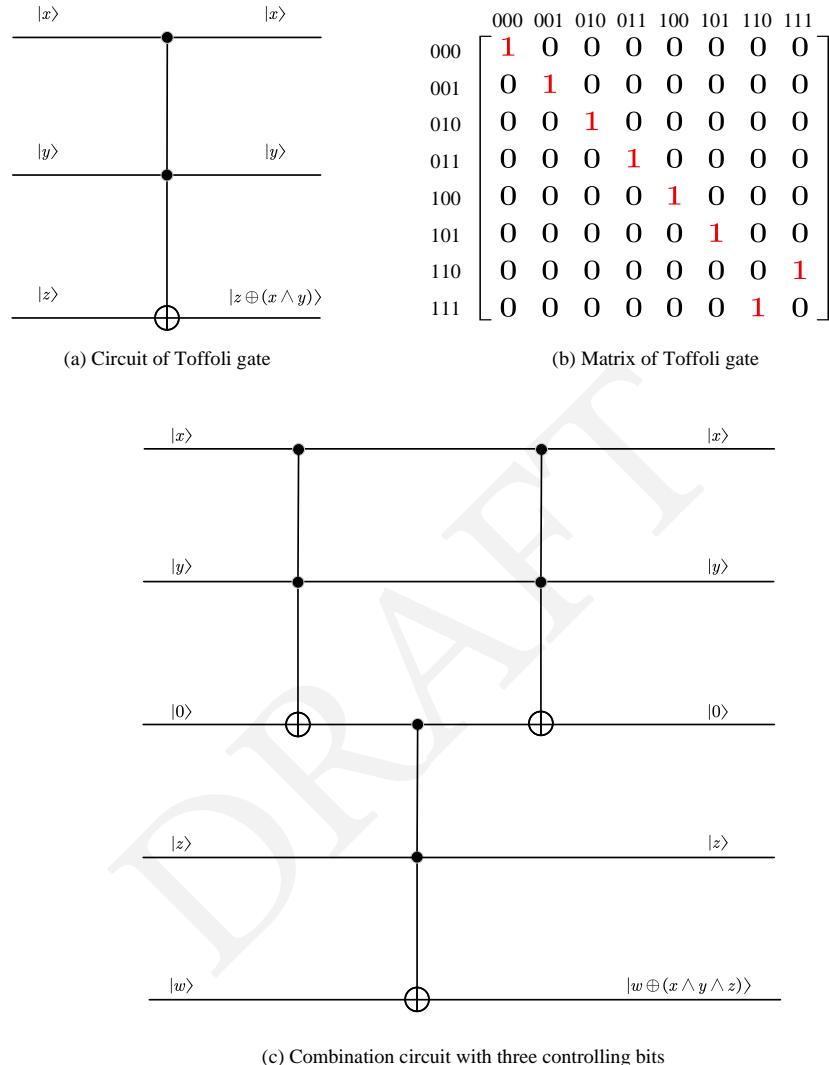


Figure 5.5: The circuit (a), matrix (b), and combination circuit (c) of Toffoli gate.

The NOT gate has no controlling bit, the CNOT gate has one controlling bit, and the Toffoli gate has two controlling bits. We can go on with this by with the combination circuit shown in Figure 5.5 (c).

One reason why the Toffoli gate is interesting is that it is universal. In other words, with

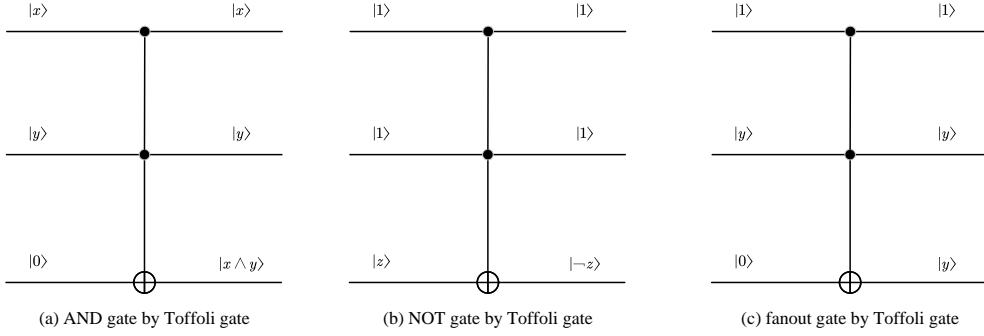


Figure 5.6: AND gate (a), NOT gate (b), and fanout gate (c) by Toffoli gate.

copies of the Toffoli gate, we can make any logical gate. In order to see that the Toffoli gate is universal, we will show that it can be used to make both the AND and NOT gates as shown in Figure 5.6 (a) and (b). Specifically, the AND gate is obtained by setting the bottom  $|z\rangle$  input to  $|0\rangle$ , and the bottom output will then be  $|x \wedge y\rangle$ . The NOT gate is obtained by setting the top two inputs to  $|1\rangle$ , and bottom output will be  $|(1 \wedge 1) \oplus z\rangle = |1 \oplus z\rangle = |\neg z\rangle$ .

Moreover, in order to construct all gates, we must also have a way of producing a fanout of values. In other words, a gate is needed that inputs a value and outputs two of the same values. This can be obtained by setting  $|x\rangle$  to  $|1\rangle$  and  $|z\rangle$  to  $|0\rangle$ . This makes the output  $|1, y, y\rangle$ .

### 5.3.3 Fredkin gate

Another interesting reversible gate is the Fredkin gate. The Fredkin gate also has three inputs and three outputs as shown in Figure 5.7 (a). The top  $|x\rangle$  input is the control input. The output is always the same  $|x\rangle$ . If  $|x\rangle$  is set to  $|0\rangle$ , then  $|y'\rangle = |y\rangle$  and  $|z'\rangle = |z\rangle$ , i.e., the values stay the same. If, on the other hand, the control  $|x\rangle$  is set to  $|1\rangle$ , then the outputs are reversed:  $|y'\rangle = |z\rangle$  and  $|z'\rangle = |y\rangle$ . In short,  $|0, y, z\rangle \mapsto |0, y, z\rangle$  and  $|1, y, z\rangle \mapsto |1, z, y\rangle$ .

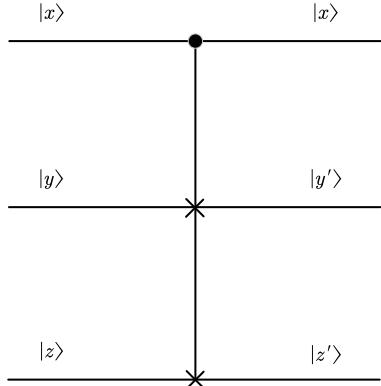
The matrix that corresponds to the Fredkin gate is shown in Figure 5.7 (b), from which we can see that the Fredkin gate is its own inverse. The Fredkin gate is also universal. By setting  $|y\rangle$  to  $|0\rangle$  as shown in Figure 5.7 (c). The NOT gate and the fanout gate can be obtained by setting  $|y\rangle$  to  $|1\rangle$  and  $|z\rangle$  to  $|0\rangle$  as shown in Figure 5.7 (d).

So both the Toffoli and the Fredkin gates are universal. Not only are both reversible gates; a glance at their matrices indicates that they are also unitary.

## 5.4 Quantum Gates

**quantum gate** is simply an operator that acts on qubits. Such operators will be represented by unitary matrices.

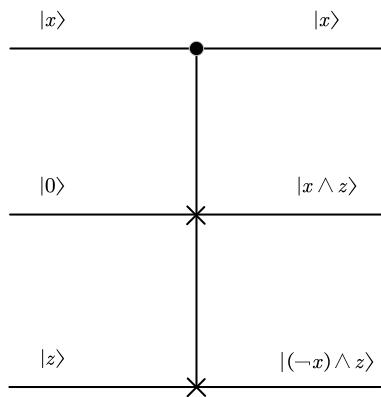
We have already worked with some quantum gates such as the identity operator I, the



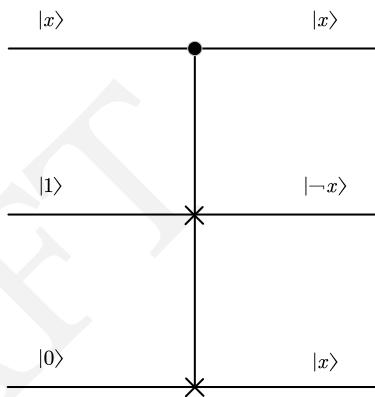
(a) Circuit of Fredkin gate

	000	001	010	011	100	101	110	111
000	1	0	0	0	0	0	0	0
001	0	1	0	0	0	0	0	0
010	0	0	1	0	0	0	0	0
011	0	0	0	1	0	0	0	0
100	0	0	0	0	1	0	0	0
101	0	0	0	0	0	0	1	0
110	0	0	0	0	0	1	0	0
111	0	0	0	0	0	0	0	1

(b) Matrix of Fredkin gate



(c) AND gate by Fredkin gate



(d) NOT gate by Fredkin gate

Figure 5.7: The circuit (a), matrix (b) of Fredkin gate and its functional equivalence to AND gate (c), and NOT and fanout gate (d).

Hadramard gate H, the NOT gate, the CNOT gate, the Toffoli gate, and the Fredkin gate. What else is there? Here we discuss some important quantum gates:

- **Pauli matrices.** They occur everywhere in quantum mechanics and quantum computing. Note that the X matrix is nothing more than our NOT matrix.

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (5.19)$$

- **Square root of NOT.** It is a one-qubit quantum gate and is denoted as  $\sqrt{\text{NOT}}$ . The matrix representation of this gate is

$$\sqrt{\text{NOT}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad (5.20)$$

- **Hardmard gates.** It is defined as

$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad (5.21)$$

The Hardmard gate has two important properties in quantum algorithm (see Example ??). First, we can transition into superposition from classic state through Hardmard gate. Second, we can transition out of superposition without measurement with the help of Hardmard gate.

- **Phase shift gate.** It is defined as

$$R(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} \quad (5.22)$$

This gate performs the following operation on an arbitrary qubit:

$$\cos(\theta') |0\rangle + e^{i\phi} \sin \theta' |1\rangle = \begin{bmatrix} \cos(\theta') \\ e^{i\phi} \sin \theta' \end{bmatrix} \mapsto \begin{bmatrix} \cos(\theta') \\ e^{i(\theta+\phi)} \sin \theta' \end{bmatrix} \quad (5.23)$$

This corresponds to a rotation that leaves the latitude alone and just changes the longitude. The new state of the qubit will remain unchanged. Only the phase will change.

- **Controlled-U gate.** It is equivalent to an IFTHEN statement. If a certain (qu)bit is true, then a particular operation should be performed, otherwise the operation is not performed. For every  $n$ -qubit unitary operation  $U$ , we can create a unitary  $(n+1)$ -qubit operation **controlled**-U or  $^C U$ :

$$^C U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix} \quad (5.24)$$

- **Deutsch gate.** It is very similar to the Toffoli gate. If the inputs  $|x\rangle$  and  $|y\rangle$  are both  $|1\rangle$ , then the phase shift operation  $R(\theta)$  will act on the  $|z\rangle$  input. Otherwise, the  $|z\rangle$  will simply be the same as the  $|z\rangle$ . When  $\theta$  is not a rational multiple of  $\pi$ ,  $D(\theta)$  by itself is a universal three-qubit quantum gate. In other words,  $D(\theta)$  will be able to mimic every other quantum gate.

Up to now, we have discussed classic gates, reversible gates and quantum gates. Their relationship can be illustrated by the following Figure 5.8.

Instructor: Chao Liang

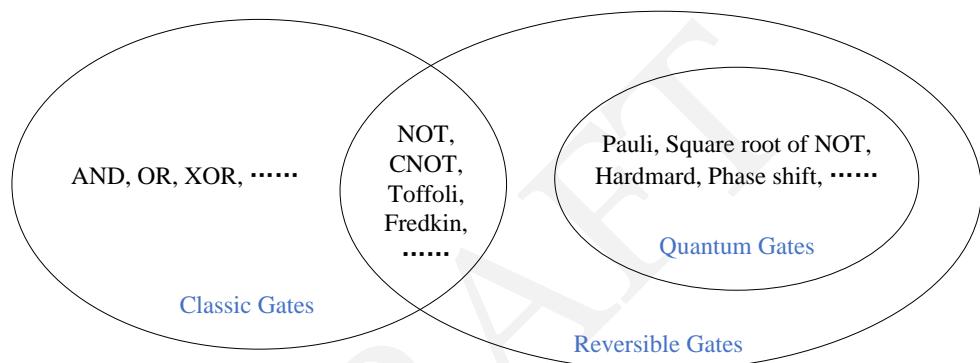


Figure 5.8: The relation of various gates.

# 6 Quantum Cryptography

**Cryptography** is the art of concealing message. The standard cryptography model is shown in Figure 6.1.

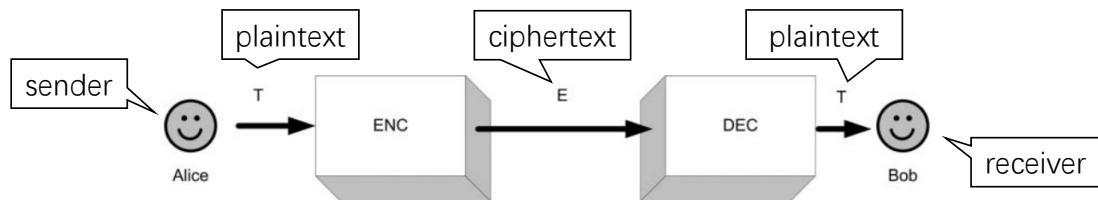


Figure 6.1: The standard cryptography model.

The whole procedure can be divided into two parts:

- Encoder (ENC) is responsible for encoding the input plaintext into ciphertext, which can be formulated as  $\text{ENC}(T, K_E) = E$  where  $T$  is the plaintext,  $K_E$  is encryption key, and  $E$  is the ciphertext.
- Decoder (DEC) is responsible for decoding the received ciphertext into plaintext, which can be formulated as  $\text{DEC}(E, K_D) = T$  where  $K_D$  is the decryption key.

$\text{DEC}(\text{ENC}(T, K_E), K_D) = T$  means that as long as we use the right keys, we can always retrieve the original message intact without any loss of information.

## 6.1 Classic Cryptography

### 6.1.1 Caesar cipher

In cryptography, a Caesar cipher<sup>1</sup> (Figure 6.2a), also known as Caesar's cipher, the shift cipher, Caesar's code or Caesar shift, is one of the simplest and most widely known encryption techniques. It is a type of substitution cipher in which each letter in the plaintext is replaced by a letter some fixed number of positions down the alphabet. As shown in Figure 6.2b, with a left shift of 3, D would be replaced by A, E would become B, and so on. The method is named after Julius Caesar, who used it in his private correspondence.

The essence of Caesar cipher is a simple linear mapping, which has high statistical correlation between the letter in plaintext and that in the ciphertext. This means that by graphing

---

<sup>1</sup>[https://en.wikipedia.org/wiki/Caesar\\_cipher](https://en.wikipedia.org/wiki/Caesar_cipher)

the frequencies of letters in the ciphertext, and by knowing the expected distribution of those letters in the original language of the plaintext, a human can easily spot the value of the shift by looking at the displacement of particular features of the graph. This is known as frequency analysis. As shown in Figure 6.2c, in the English language the plaintext frequencies of the letters E, T, (usually most frequent), and Q, Z (typically least frequent) are particularly distinctive.

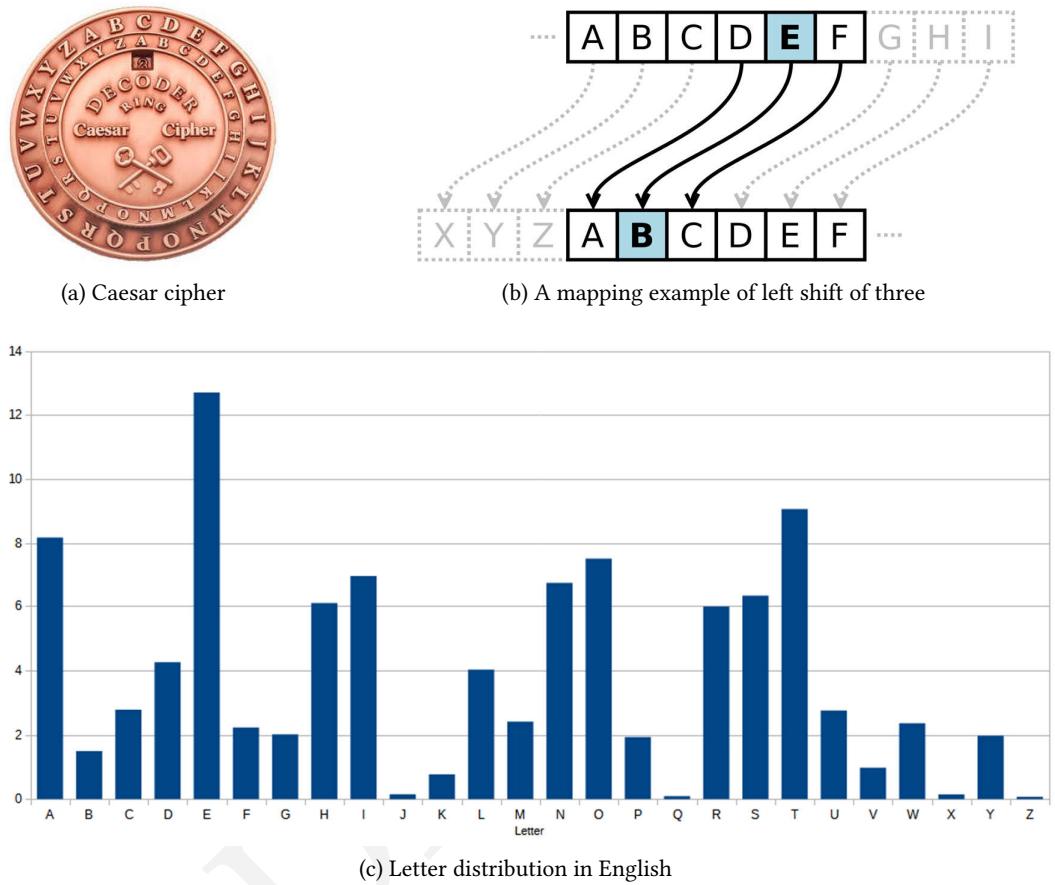


Figure 6.2: Caesar cipher's device, scheme, and defect.

### 6.1.2 One-Time-Pad protocol

In cryptography, the one-time-pad (OTP) protocol<sup>2</sup> is an encryption technique that cannot be cracked, but requires the use of a single-use pre-shared key that is not smaller than the message being sent. In this technique, a plaintext is paired with a random secret key (also referred to as a one-time pad). Then, each bit or character of the plaintext is encrypted by combining it with the corresponding bit or character from the pad using modular addition.

<sup>2</sup>[https://en.wikipedia.org/wiki/One-time\\_pad](https://en.wikipedia.org/wiki/One-time_pad)

<b>One-Time-Pad Protocol</b>						
Original message T	0	1	1	0	1	1
Encryption key K	$\oplus$	1	1	1	0	1
Encrypted message E		1	0	0	0	0
Public channel		↓	↓	↓	↓	↓
Received message E	1	0	0	0	0	1
Decryption key K	$\oplus$	1	1	1	0	1
Decrypted message T		0	1	1	0	1

Figure 6.3: The one-time-pad protocol.

In OTP protocol, both encryption and decryption end share the same key  $K$  in the communication process, which means  $K_E = K_D = K$ . As shown in Figure 6.3, assume that both encoder and decoder share the same working function  $\text{ENC}(T, K) = \text{DEC}(T, K) = T \oplus K$ , than the receiver can decode the ciphertext and get the original message as follows

$$\begin{aligned}
 \text{DEC}(\text{ENC}(T, K_E), K_D) &= \text{DEC}(T \oplus K, K) \\
 &= (T \oplus K) \oplus K \\
 &= T \oplus (K \oplus K) \\
 &= T
 \end{aligned} \tag{6.1}$$

The merit of OTP protocol is that it cannot be cracked, but meanwhile, it has two obvious drawbacks. First, the key in OTP must be longer than the message being sent, which is extremely inconvenient in the transmission and storage process; Second, the key cannot be re-used because of the risk of information leakage<sup>3</sup>.

### 6.1.3 Diffie-Hellman key exchange

To eliminate risks in the key distribution process, Whitfield Diffie and Martin Hellman devised the Diffie-Hellman key exchange method to securely exchange cryptographic keys over a public channel.

---

<sup>3</sup>The eavesdropper may infer part of the original message from the multiple intercepted ciphertext if the key is reused because  $E_1 \oplus E_2 = (T_1 \oplus K) \oplus (T_2 \oplus K) = T_1 \oplus K \oplus K \oplus T_2 = T_1 \oplus T_2$

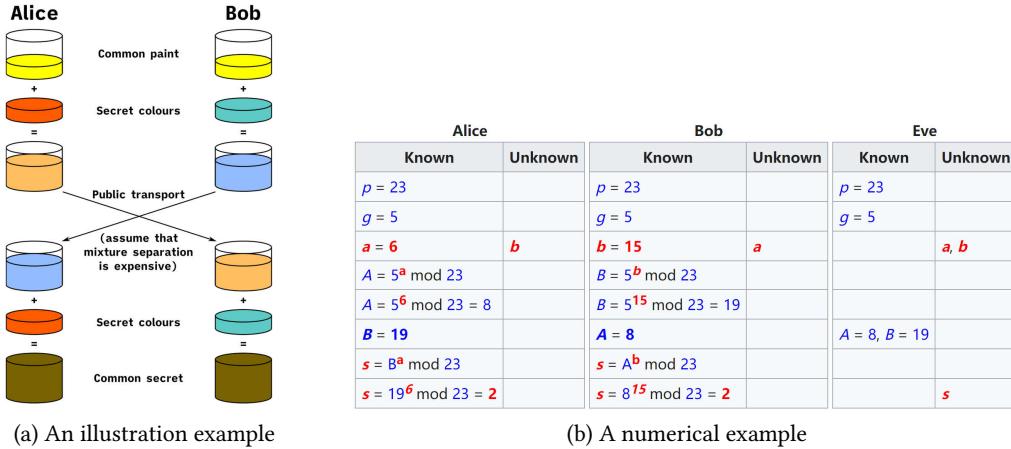


Figure 6.4: Illustrative and numerical examples of Diffie-Hellman key exchange.

An analogy illustrates the concept of public key exchange by using colors instead of very large numbers. As shown in Figure 6.4a, the Diffie-Hellman key exchange process begins by having the two parties, Alice and Bob, publicly agree on an arbitrary starting color that does not need to be kept secret. In this example, the color is yellow. Each person also selects a secret color that they keep to themselves in this case, red and cyan. The crucial part of the process is that Alice and Bob each mix their own secret color together with their mutually shared color, resulting in orange-tan and light-blue mixtures respectively, and then publicly exchange the two mixed colors. Finally, each of them mixes the color they received from the partner with their own private color. The result is a final color mixture (yellow-brown in this case) that is identical to their partner's final color mixture.

The numerical example of the above colorization process is shown in Figure 6.4b. The pigment mixing process is formulated as a classic trapdoor function, also dubbed as one-way function, *i.e.*, modular exponentiation  $f(x) = g^x \text{ mod } p$ . The forward computation, from  $x$  to  $f(x)$ , is easy and fast, while the backward computation, from  $f(x)$  to  $x$ , is extremely hard and computationally prohibitive. In the table,  $a$  and  $b$  represent the secret color that Alice and Bob keep to themselves at the beginning,  $A$  and  $B$  represent the publicly exchanged color, and  $s$  is the final shared color mixture. Thanks to the one-way property of the modular exponentiation, even Eve got  $A$  and  $B$ , he cannot recover  $a$  and  $b$ , not to mention the exchanged key  $s$ .

#### 6.1.4 Public- and private-key cryptography

According to the availability of the encryption key, cryptography algorithm can be divided into private-key cryptography and public-key cryptography (Figure 6.5).

**private-key cryptography**, or symmetric cryptography, uses the same cryptographic keys for both the encryption of plaintext and the decryption of ciphertext. The keys, in practice, represent a shared secret between two or more parties that can be used to maintain a private information link (Figure 6.5a). The requirement that both parties have access to the secret

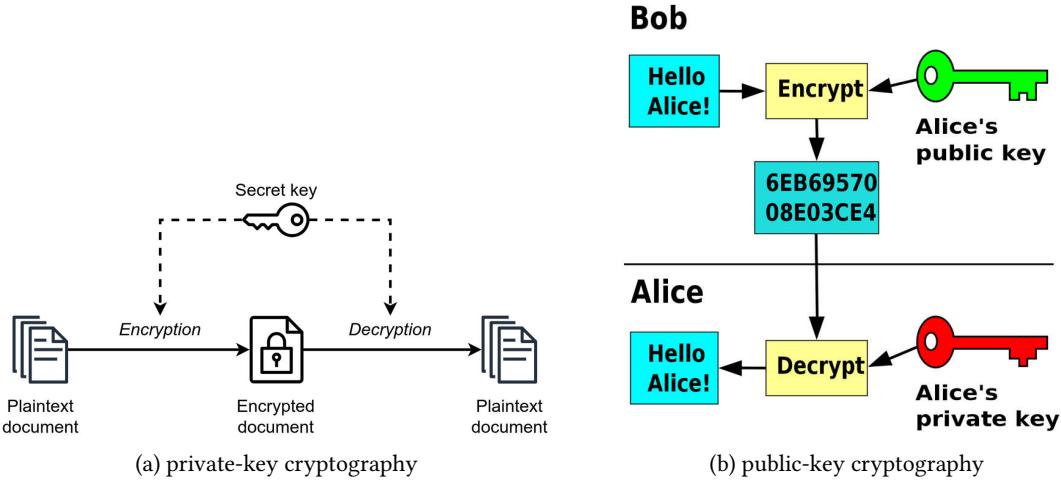


Figure 6.5: Diagrams of private-key cryptography and public-key cryptography.

key is one of the main drawbacks of symmetric-key encryption, in comparison to public-key encryption. OTP protocol is a typical private-key algorithm.

**Public-key cryptography**, or asymmetric cryptography, is the field of cryptographic systems that use pairs of related keys. Each key pair consists of a public key and a corresponding private key. Key pairs are generated with cryptographic algorithms based on mathematical problems termed one-way functions. In a public-key encryption system, anyone with a public key can encrypt a message, yielding a ciphertext, but only those who know the corresponding private key can decrypt the ciphertext to obtain the original message (Figure 6.5b). Diffie-Hellman key exchange belongs to the public-key cryptography.

The plus side of public-key cryptography is that it does have key distribution problem. In contrast, the minus sides include: (1) the one-way property of trapdoor function is a temporary fact that may disappear in the future; (2) public-key cryptography is usually slower than private-key cryptography.

Typical issues in classic cryptography include: (1) success communication; (2) intrusion detection, *i.e.*, Alice and Bob would like to determine whether Eve is, in fact, eavesdropping; (3) authentication, *i.e.*, we would like to ensure that nobody is impersonating Alice and sending false messages.

## 6.2 Quantum Key Exchange

Due to the peculiar effect of quantum observation and measurement, eavesdropping in the classical world has a very different manifestation from that in the quantum world. Specifically, in the classic world, Eve can make copies of arbitrary portions of the encrypted bit stream, and he can listen without affecting the bit stream. While in the quantum world, Eve cannot

make perfect copies of the qubit stream because of the no-cloning theorem, and the very act of measuring the qubit stream alters it.

### 6.2.1 The BB84 protocol

The first quantum key exchange protocol was introduced by Charles Bennett and Gilles Brassard in 1984, and hence the name BB84.

**Problem setting.** Alices goal is to send Bob a key via a quantum channel. Just as in the One-TimePad protocol, her key is a sequence of random (classical) bits obtained, perhaps, by tossing a coin. Alice will send a qubit each time she generates a new bit of her key. But which qubit should she send?

**Plus and times bases.** In this protocol, Alice will employ two different orthogonal bases shown in Figure 6.6. The left part shows the “plus” basis whose formulation is

$$+ = \{|\rightarrow\rangle, |\uparrow\rangle\} = \{[1, 0]^\top, [0, 1]^\top\} \quad (6.2)$$

and the right part shows the “times” basis whose formulation is

$$\times = \{|\nwarrow\rangle, |\nearrow\rangle\} = \left\{ \frac{1}{\sqrt{2}}[-1, 1]^\top, \frac{1}{\sqrt{2}}[1, 1]^\top \right\} \quad (6.3)$$

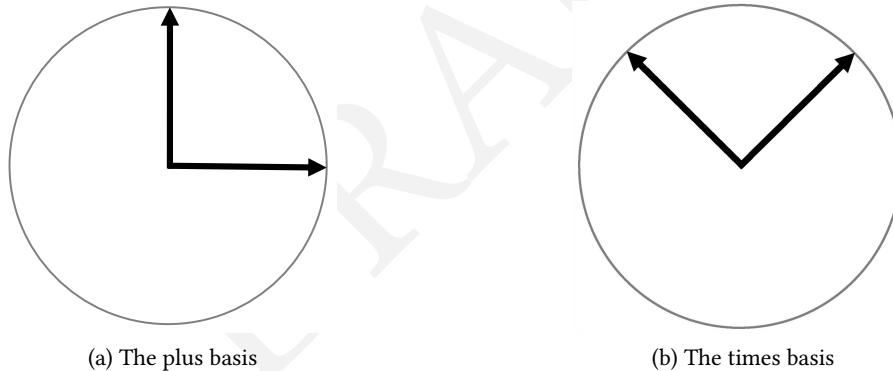


Figure 6.6: Two basis used in BB84.

**Cross representation.** The base vector in one basis can be linearly represented by base vectors in another basis:

- $|\nwarrow\rangle$  with respect to  $+$  will be  $\frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}|\rightarrow\rangle$ .
- $|\nearrow\rangle$  with respect to  $+$  will be  $\frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\rightarrow\rangle$ .
- $|\uparrow\rangle$  with respect to  $\times$  will be  $\frac{1}{\sqrt{2}}|\nearrow\rangle + \frac{1}{\sqrt{2}}|\nwarrow\rangle$ .
- $|\rightarrow\rangle$  with respect to  $\times$  will be  $\frac{1}{\sqrt{2}}|\nearrow\rangle - \frac{1}{\sqrt{2}}|\nwarrow\rangle$ .

**Mapping table.** Alice translates her classic key bits into qubits according to a mapping vocabulary shown in the following table. For example, in the + basis, a  $| \rightarrow \rangle$  will correspond

State/Basis	+	$\times$
$ 0\rangle$	$  \rightarrow \rangle$	$  \nearrow \rangle$
$ 1\rangle$	$  \uparrow \rangle$	$  \nwarrow \rangle$

to a  $|0\rangle$ . If Alice wants to work in the  $\times$  basis and wants to convey a  $|1\rangle$ , she will send a  $| \nwarrow \rangle$ . Similarly, if Alice sends a  $| \uparrow \rangle$  and Bob measures a  $| \uparrow \rangle$  in the + basis, he should record a  $|1\rangle$ .

**BB84 protocol.** There are totally 4 steps in the BB84 protocol:

**step 1** (Alice)

- randomly determine classical bits to send
- randomly determine the bases to send bits
- send the bits in their appropriate basis

<b>Step 1: Alice sends <math>n</math> random bits in random bases</b>												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bits	0	1	1	0	1	1	1	0	1	0	1	0
Alice's random bases	+	+	X	+	+	+	X	+	X	X	X	+
Alice sends	$\rightarrow$	$\uparrow$	$\nwarrow$	$\rightarrow$	$\uparrow$	$\uparrow$	$\nwarrow$	$\rightarrow$	$\nwarrow$	$\nearrow$	$\nwarrow$	$\rightarrow$
Quantum channel	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$

**step 2** (Bob)

- randomly determine the bases to receive bits
- measure the qubit in those random bases

<b>Step 2: Bob receives <math>n</math> random bits in random measurements</b>												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Bob's random bases	X	+	X	X	+	X	+	+	X	X	X	+
Bob observes	$\nearrow$	$\uparrow$	$\nwarrow$	$\nwarrow$	$\uparrow$	$\nearrow$	$\uparrow$	$\rightarrow$	$\nwarrow$	$\nearrow$	$\nwarrow$	$\rightarrow$
Bob's bits	0	1	1	1	1	0	1	0	1	0	1	0

When there is no eavesdropping, Bob has 100% probability to get the correct bit with consistent bases, and 50% probability to get the correct bit with inconsistent bases. Hence the expected correct rate (ECR) for Bob getting the correct bit is  $\frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2} = 75\%$ .

When there is Eve bugging in, he reads the information that Alice transmits, and meanwhile, sends that information onward to Bob. In this case, ECR changes into  $\frac{3}{4} \times \frac{3}{4} + \frac{1}{4}(\frac{1}{2} \times 0 + \frac{1}{2} \times \frac{1}{2}) = 62.5\%$ . The first term means both Eve and Bob get the correct bit, while the second term represents that Eve gets the wrong bit but Bob gets the right one.

The following Table gives another solution to calculate the ECR with eavesdropping, where the first stage considers the operations of receiving basis and sending qubit of Eve, and the second stage considers the receiving basis and bit of Bob.

Eve		Bob		Probability
Receiving basis (consistent to Alice)	Sending qubit (consistent to Alice)	Receiving basis (consistent to Eve)	Receiving bit (consistent to Alice)	
$P(\checkmark) = 1/2$	$P(\checkmark) = 1$	$P(\checkmark) = 1/2$	$P(\checkmark) = 1$	$\frac{1}{2} \cdot 1 \cdot \left[ \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} \right] = \frac{6}{16}$
		$P(\times) = 1/2$	$P(\checkmark) = 1/2$	
		$P(\checkmark) = 1/2$	$P(\checkmark) = 1$	$\frac{1}{2} \cdot \frac{1}{2} \cdot \left[ \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} \right] = \frac{3}{16}$
		$P(\times) = 1/2$	$P(\checkmark) = 1/2$	
		$P(\times) = 1/2$	$P(\checkmark) = 0$	$\frac{1}{2} \cdot \frac{1}{2} \cdot \left[ \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} \right] = \frac{1}{16}$
			$P(\checkmark) = 1/2$	

### step 3 (Alice and Bob)

- publicly compare which basis they used at each step
- scratch out corresponding bits under different bases

Step 3: Alice and Bob publicly compare their bases												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bases	X	X	+	+	X	+	X	+	+	X	+	X
Public channel	◻	◻	◻	◻	◻	◻	◻	◻	◻	◻	◻	◻
Bob's random bases	X	+	+	X	X	+	+	+	+	X	X	+
Which agree?	✓		✓		✓	✓		✓	✓	✓		

### step 4 (Bob)

- randomly chooses half of the  $n/2$  bits
- publicly compares them with Alice

If  $\text{ECR} \leq 1 - \epsilon$ , Eve is listening, Alice and Bob scratch the whole sequence. Otherwise, Alice and Bob scratch out the revealed test subsequence and keep the remains as unrevealed secret private key.

<b>Step 4: Alice and Bob publicly compare half of the remaining bits</b>												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Shared secret keys	1	1		1			0	1	0	1	0	
Randomly chosen to compare			✓					✓	✓			✓
Public channel			⇓					⇓	⇓	⇓		⇓
Shared secret keys	1	1		1			0	1	0	1	0	
Which agree?			✓					✓	✓			✓
Unrevealed secret keys:	1			1			0			1		

### 6.2.2 The B92 protocol

In the BB84 protocol, Alice had two distinct orthogonal bases at her disposal. It turns out that the use of two different bases is redundant, provided one employs a slightly slicker means of measuring. This simplification results in another quantum key distribution protocol, known as B92, invented by Charles Bennett in 1992.

**The non-orthogonal basis.** Alice uses only one non-orthogonal basis

$$\{|\rightarrow\rangle, |\nearrow\rangle\} = \left\{[1, 0]^\top, \frac{1}{\sqrt{2}}[1, 1]^\top\right\} \quad (6.4)$$

**B92 protocol.** There are totally 4 steps in the B92 protocol:

#### step 1 (Alice)

- randomly determine classical bits to send
- send the bits in the appropriate polarization

<b>Step 1: Alice sends <math>n</math> random bits in the <math>\angle</math> basis</b>												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bits	0	0	1	0	1	0	1	0	1	1	1	0
Alice's qubits	→	→	↗	→	↗	→	↗	→	↗	↗	↗	→
Quantum channel	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓

#### step 2 (Bob)

- randomly determine the bases to receive bits
- measure the qubit in those random bases

Step 2: Bob receives $n$ random bits in a random basis												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bits	→	→	↗	→	↗	→	↗	→	↗	↗	↗	→
Quantum channel	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Bob's random bases	X	+	X	X	+	X	+	+	X	+	X	+
Bob's observations	↖	→	↗	↖	↑	↖	→	→	↗	↑	↗	→
Bob's bits	0	?	?	0	1	0	?	?	?	1	?	?

- If Bob uses the + basis and observes a  $|↑\rangle$ , then he knows that Alice must have sent a  $|↗\rangle = |1\rangle$  because if Alice had sent a  $|→\rangle$ , Bob would have received a  $|→\rangle$ .
- If Bob uses the + basis and observes a  $|→\rangle$ , then it is not clear to him which qubit Alice sent. She could have sent a  $|→\rangle$  but she could also have sent a  $|↗\rangle$  that collapsed to a  $|→\rangle$ . Because Bob is in doubt, he will omit this bit.
- If Bob uses the × basis and observes a  $|↖\rangle$ , then he knows that Alice must have sent a  $|→\rangle = |0\rangle$  because if Alice had sent a  $|↗\rangle$ , Bob would have received a  $|↗\rangle$ .
- If Bob uses the × basis and observes a  $|↗\rangle$ , then it is not clear to him which qubit Alice sent. She could have sent a  $|↗\rangle$  but she could also have sent a  $|→\rangle$  that collapsed to a  $|↗\rangle$ . Because Bob is in doubt, he will omit this bit.

#### step 3 (Alice and Bob)

- Bob publicly tells Alice which bits were uncertain
- they both omit uncertain bits

#### step 4 (optional for intrusion detection)

- Bob randomly chooses half of the  $n/2$  bits
- Bob publicly compares them with Alice

### 6.2.3 The EPR protocol

In 1991, Artur K. Ekert proposed a completely different type of quantum key distribution protocol based on entanglement. In the chapter of composite system, we learned that we can prepare a sequence of entangled pairs of qubits like  $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$  or  $\frac{|01\rangle+|10\rangle}{\sqrt{2}}$ . For the discussion convenience, we assume the pair of entangled qubits in the state of  $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$ .

#### step 1 (Alice and Bob)

- Both sides are each assigned one of each of the pairs of entangled qubits

**step 2** (Alice and Bob)

- separately choose a random sequence of bases
- measure their qubits in their chosen basis

Step 2: Alice and Bob measure in each of their random bases												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bases	X	X	+	+	X	+	X	+	+	X	+	X
Alice's observations	↗	↖	→	↑	↗	→	↖	→	→	↗	→	↗
Bob's random bases	X	+	+	X	X	+	+	+	+	X	X	+
Bob's observations	↗	→	→	↗	↗	→	↑	→	→	↗	↖	→

**step 3** (Alice and Bob)

- publicly compare what bases were used
- keep only those bits measured in the same basis

Step 3: Alice and Bob publicly compare their bases												
Bit number	1	2	3	4	5	6	7	8	9	10	11	12
Alice's random bases	X	X	+	+	X	+	X	+	+	X	+	X
Public channel	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓
Bob's random bases	X	+	+	X	X	+	+	+	+	X	X	+
Which agree?	✓		✓		✓	✓		✓	✓		✓	

**step 4** (optional for intrusion or disentangled detection)

- Bob randomly chooses half of the n/2 bits
- Bob publicly compares them with Alice

### 6.3 Quantum Teleportation

**Quantum teleportation** is the process by which the state of an arbitrary qubit is transferred from one location to another.

**Canonical and non-canonical bases for a single qubit.** When working with a single qubit, we worked with the canonical basis,  $\{|0\rangle, |1\rangle\}$  and non-canonical basis,  $\{\frac{|0\rangle+|1\rangle}{\sqrt{2}}, \frac{|0\rangle-|1\rangle}{\sqrt{2}}\}$ , as shown in Figure 6.7.

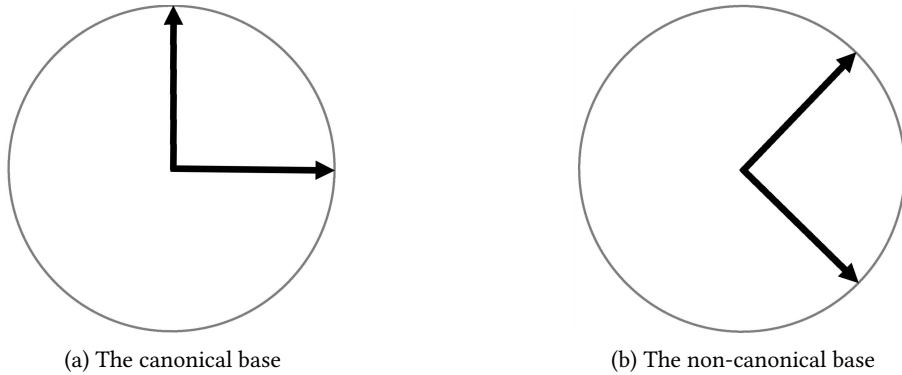


Figure 6.7: Canonical and non-canonical basis used in the quantum teleportation.

**Canonical and non-canonical bases for two qubits.** The teleportation algorithm will work with two entangled qubits, one held by Alice and one held by Bob. The obvious canonical basis for this four-dimensional space is:

$$\{|0_A0_B\rangle, |0_A1_B\rangle, |1_A0_B\rangle, |1_A1_B\rangle\} \quad (6.5)$$

A non-canonical basis, called the Bell basis in honor of John Bell, consists of the following four vectors:

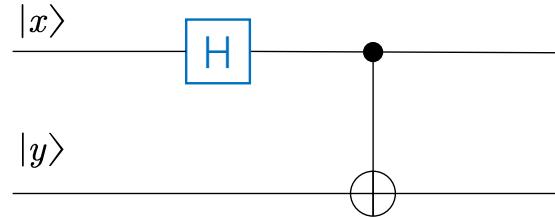
$$\begin{aligned} |\Psi^+\rangle &= \frac{|0_A1_B\rangle + |1_A0_B\rangle}{\sqrt{2}}, & |\Psi^-\rangle &= \frac{|0_A1_B\rangle - |1_A0_B\rangle}{\sqrt{2}}, \\ |\Phi^+\rangle &= \frac{|0_A0_B\rangle + |1_A1_B\rangle}{\sqrt{2}}, & |\Phi^-\rangle &= \frac{|0_A0_B\rangle - |1_A1_B\rangle}{\sqrt{2}} \end{aligned} \quad (6.6)$$

Every vector in this basis is entangled.

**Bell circuit.** How to derive the Bell basis? In the single qubit (two-dimensional) case, the elements of the noncanonical basis can be formed using the Hadamard matrix:

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \text{and} \quad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \quad (6.7)$$

In the two qubits (four-dimensional) case, the elements of the non-canonical basis are derived from Bell circuit:



From which, we have  $|00\rangle \mapsto |\Phi^+\rangle$ ,  $|10\rangle \mapsto |\Phi^-\rangle$ ,  $|01\rangle \mapsto |\Psi^+\rangle$ ,  $|11\rangle \mapsto |\Psi^-\rangle$ .

**Example 6.1:**  $|00\rangle \mapsto |\Phi^+\rangle$

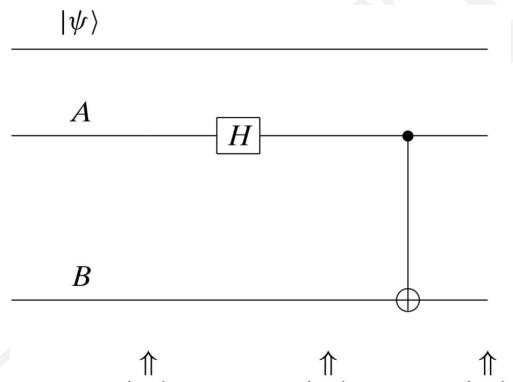
$$\begin{aligned}
 \text{CNOT} \cdot (\mathbf{H}|0\rangle \otimes |0\rangle) &= \text{CNOT} \cdot \left( \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |0\rangle \right) \\
 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\Phi^+\rangle
 \end{aligned} \tag{6.8}$$

**Quantum teleportation protocol.** The whole protocol contains 5 steps:

**step 1 :** Alice has a qubit  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ .

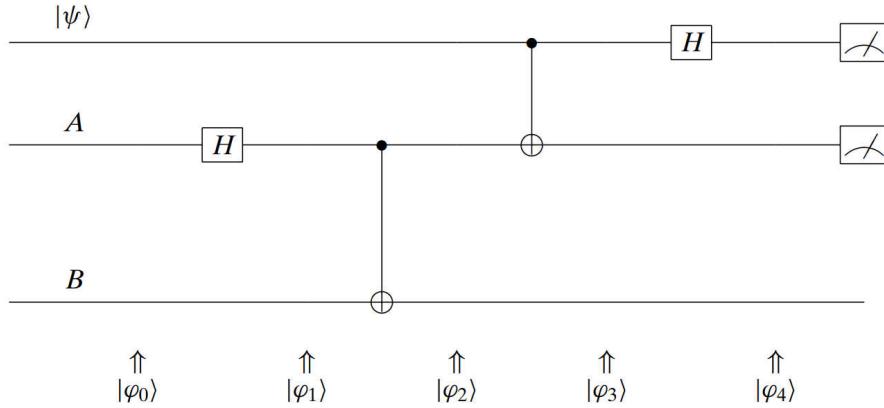
**step 2 :** prepare two entangled quibit A and B

- two entangled quibits are formed as  $|\Phi^+\rangle$
- one is given to Alice and one is given to Bob



$$\begin{aligned}
 |\varphi_0\rangle &= |\psi\rangle \otimes |0_A\rangle \otimes |0_B\rangle = |\psi\rangle |0_A 0_B\rangle \\
 |\varphi_1\rangle &= |\psi\rangle \otimes \frac{|0_A\rangle + |1_A\rangle}{\sqrt{2}} \otimes |0_B\rangle \\
 |\varphi_2\rangle &= |\psi\rangle \otimes |\Phi^+\rangle = |\psi\rangle \otimes \frac{|0_A 0_B\rangle + |1_A 1_B\rangle}{\sqrt{2}} \\
 &= (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{|0_A 0_B\rangle + |1_A 1_B\rangle}{\sqrt{2}} \\
 &= \frac{\alpha|0\rangle(|0_A 0_B\rangle + |1_A 1_B\rangle) + \beta|1\rangle(|0_A 0_B\rangle + |1_A 1_B\rangle)}{\sqrt{2}}
 \end{aligned} \tag{6.9}$$

**step 3 :** Alice lets her  $|\psi\rangle$  interact with her entangled qubit



$$\begin{aligned}
 |\varphi_2\rangle &= \frac{\alpha|0\rangle(|0_A0_B\rangle + |1_A1_B\rangle) + \beta|1\rangle(|0_A0_B\rangle + |1_A1_B\rangle)}{\sqrt{2}} \\
 |\varphi_3\rangle &= \frac{\alpha|0\rangle(|0_A0_B\rangle + |1_A1_B\rangle) + \beta|1\rangle(|1_A0_B\rangle + |0_A1_B\rangle)}{\sqrt{2}} \\
 |\varphi_4\rangle &= \frac{1}{2}(\alpha(|0\rangle + |1\rangle)(|0_A0_B\rangle + |1_A1_B\rangle) + \beta(|0\rangle - |1\rangle)(|1_A0_B\rangle + |0_A1_B\rangle)) \quad (6.10) \\
 &= \frac{1}{2}(\alpha(|000\rangle + |011\rangle + |100\rangle + |111\rangle) + \beta(|010\rangle + |001\rangle - |110\rangle - |101\rangle)) \\
 &= \frac{1}{2}[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\beta|0\rangle + \alpha|1\rangle) \\
 &\quad + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(-\beta|0\rangle + \alpha|1\rangle)]
 \end{aligned}$$

**step 4** : Alice makes a measurement

- Alice measures her two qubits
- Alice determines to which of the four possible states the system collapses

**step 5** : Bob performs the corresponding transformation based on Alice's observations

- Alice sends copies of her two bits (not qubits) to Bob
- Bob uses that information to achieve the desired state

Bob's reconstruction matrices				
Bits received	$ 00\rangle$	$ 01\rangle$	$ 10\rangle$	$ 11\rangle$
Matrix to apply	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

**Example 6.2: Bob's transformation when Alice observes  $|10\rangle$** 

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ -\beta \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha |0\rangle + \beta |1\rangle = |\psi\rangle \quad (6.11)$$

The whole framework of the quantum teleportation protocol is shown in Figure 6.8. Several points should be made about this protocol:

- Alice is no longer in possession of  $|\psi\rangle$ . She has only two classical bits.
- As we have seen, to teleport a single quantum particle, Alice has to send two classical bits. Without receiving them, there is no way that Bob can know what he has. These classical bits travel along a classical channel and thus they propagate at finite speed (less than the speed of light). Entanglement, in spite of its undisputable magic, does not allow you to communicate faster than the speed of light.
- Information teleported from Alice to Bob via qubit is infinite, but it is useless to Bob once he make the measurement (qubit will collapse to a classic bit).
- no particle has been moved at all, only the state.

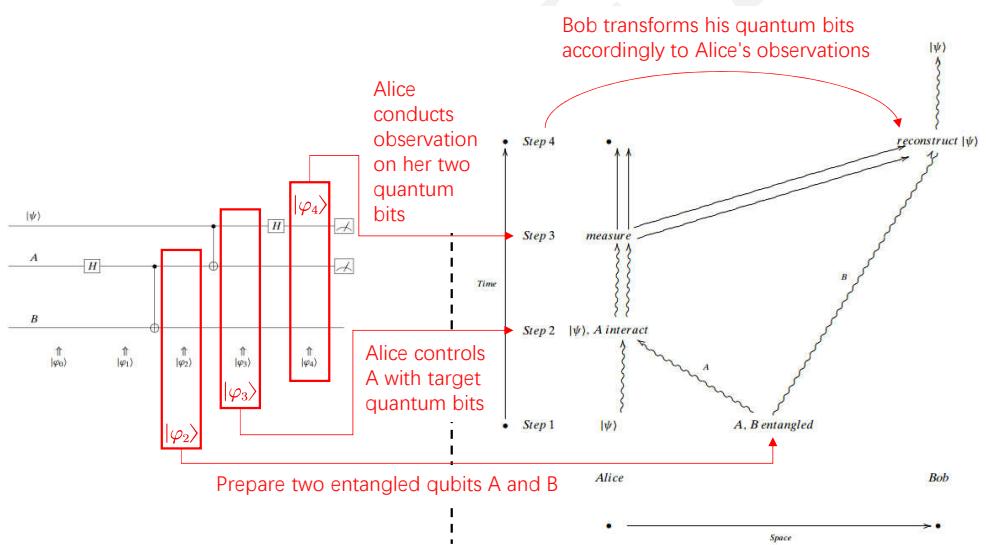


Figure 6.8: The framework of quantum teleportation protocol.

Instructor: Chao Liang