Hiu Yung Wong

Quantum Computing Architecture and Hardware for Engineers

Step by Step

Springer

Quantum Computing Architecture and Hardware for Engineers

Hiu Yung Wong

Quantum Computing Architecture and Hardware for Engineers

Step by Step

Springer

Hiu Yung Wong
Department of Electrical Engineering
San Jose State University
San Jose, CA, USA

ISBN 978-3-031-78218-3 ISBN 978-3-031-78219-0 (eBook)
https://doi.org/10.1007/978-3-031-78219-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my wife, Michelle, my daughter, Jennifer, and my son, James

Preface

The purpose of this book is to teach quantum computing hardware from an engineer's perspective. Engineers play an important role in quantum computers. However, college and graduate engineering students usually do not have the required physics and mathematics training to understand how quantum computer hardware works. This book provides step-by-step guidance to connect engineers to the quantum world. It is based on the teaching materials I created at San José State University for EE274 Quantum Computing Architectures in Spring 2023.

In this book, quantum computers built on silicon spin qubits and superconducting qubits are discussed in detail. The physics, mathematics, and their connection to microwave electronics are presented based on how they fulfill the five DiVincenzo's criteria. Readers will be able to understand the nuts and bolts of qubit initialization, readout, and manipulation. At the end, I also present a superconducting qubit integrated circuit design example, which is not commonly found elsewhere.

The characteristics of this book are three. Firstly, I try to write the equations and explain the process step-by-step to avoid ambiguity. As an engineer, I believe that this is very useful for serious engineers who want to connect their knowledge to quantum computing hardware. Such a connection is difficult to find in the literature. Secondly, I try to connect and contrast the spin qubits with superconducting qubits so that the readers can understand them in a unified theory. Lastly, simulation programs and design examples along with teaching videos are provided.

I learned quantum computing hardware by myself and I believe I have written the book in a way that the confusions and ambiguities an engineer usually faces are highlighted and answered. I also need to thank Dr. Yaniv Jacob Rosen and Dr. Kristin Beck from the Lawrence Livermore National Laboratory for the collaborations since 2021, through which I learned a lot from both of them and also had the opportunity to practice my knowledge in superconducting qubit hardware.

This book is organized into four parts. Part I gives an overview of quantum computers and the basics of linear algebra, i.e., the Schrödinger equation, the density matrix, and the Bloch sphere. These represent the minimal knowledge required to understand quantum computing hardware. Readers are also encouraged to read

my other book, "Introduction to Quantum Computing: From a Layperson to a Programmer in 30 Steps" if one is interested in learning the basics of quantum algorithms.

In Part II, we discuss the physics of spin magnetic moment and magnetic field interaction which is the basics of spin qubits. Besides Larmor precession, we go through the physics and mathematics of Rabi oscillation under a linearly oscillating magnetic field and rotating magnetic field. In the process, we build the mathematical tools and concepts, such as rotating frame and rotating wave approximation which are required to understand qubit operations in many technologies. With this knowledge, we show how a silicon electron qubit can be implemented on a Metal-Oxide-Semiconductor (MOS) technology.

In Part III, we first introduce Lagrangian and Hamiltonian mechanics, which are usually not covered in engineering classes. Then we show how to connect classical mechanics to quantum mechanics through operator promotion. We will discuss the details of a mechanical simple harmonic oscillator and reuse their equations in an LC tank. Then we show that an LC tank is not a suitable qubit due to the lack of anharmonicity and introduce the physics and fabrication of the Josephson junction due to its ability to provide nonlinear inductance for anharmonicity in a superconducting qubit. We go through circuit quantization and discuss the characteristics of the Cooper pair box in its charge qubit and transmon qubit regimes. We highlight the similarities between the superconducting qubit and spin qubit Hamiltonian and show that we can reuse the solutions from the spin qubit after careful treatments. In this process, we share scripts for solving the superconducting qubits.

In Part IV, we discuss the roles of microwave electronics in quantum computers. Then we show the design parameters and methodologies of superconducting integrated quantum chip. Finally, we touch upon errors and decoherence time measurements. After reading this book, readers will be more ready to study more advanced topics such as open quantum systems.

This book is accompanied by many teaching videos and GitHub codes. Readers can find them in Appendix A.

For any comments, please send an email to intro.qc.wong@gmail.com.

San Jose, California, USA Hiu Yung Wong
August 2024

Contents

Part I
Quantum Computer Overview and Essential Linear Algebra and Quantum Mechanics

Chapter 1
Quantum Computer Hardware and Architecture—An Overview

1.1 Introduction

In this chapter, we will have an overview of a quantum computer from an engineer's perspective. We will first look at the basic operations of a quantum computer, namely, qubit initialization, qubit manipulation, and qubit readout. We then introduce and discuss DiVincenzo's criteria which set the minimal requirement for a technology to build a quantum computer. Then we will use a superconducting qubit quantum computer as an example to trace the propagation, attenuation, modification, and amplification of the microwave signals to gain a deeper understanding of the operating principles of a quantum computer and the critical role of classical and microwave electronics in a quantum computer. We will also discuss the scaling of a quantum computer.

1.1.1 Learning Outcomes

Understand the basic operations of a quantum computer; be able to describe the role of classical electronics in a quantum computer; appreciate the challenges in achieving a highly scaled quantum computer.

1.1.2 Teaching Videos

- Search for Ch1 in this playlist
  - https://tinyurl.com/3yhze3jn

H. Y. Wong, *Quantum Computing Architecture and Hardware for Engineers*,
https://doi.org/10.1007/978-3-031-78219-0_1

- Other videos
  - https://youtu.be/M93Qtu3J-5M

1.2 Quantum Computer as a Classical Electronics-Controlled Wavefunction

A quantum computer is nothing but a **quantum bit (qubit)** wavefunction controlled by classical electronics (Fig. 1.1). The qubit wavefunction, such as the spin of an electron, the electronic state of an ion, and the excess number of Cooper pairs in a superconducting qubit, is passive. Without interacting with the external environment, it is a "boring" *passive wavefunction*. A quantum computing process is also nothing but the modification of the wavefunction through an appropriate **Hamiltonian**. We will discuss Hamiltonian in future chapters (e.g., Chap. 4). Hamiltonian is the total energy of the system which *governs the evolution (change) of the wavefunction*. An appropriate Hamiltonian is created by applying a suitable laser pulse or microwave pulse with the right *frequency*, right *shape*, right *amplitude*, and right *phase* at the right *time* for a right *duration*. The pulses are generated in classical electronics and it is not surprising that microwave engineering, laser optics, and signal processing play a key role in a quantum computer. Since the pulses need to be short and fast, high-speed digital circuits (such as application-specific integrated circuits (**ASICs**) and field programmable gate arrays (**FPGAs**)) and efficient classical programming and data processing are crucial in a quantum computer.

At the time of writing, *a quantum computer is just a well-organized physics laboratory*.

Figure 1.1 shows that there are three critical interactions between the classical electronics and the qubit wavefunctions ($QB_{n-1}, \cdots, QB_0$), namely, *qubit initialization*, *qubit manipulation*, and *qubit readout*. These are all performed by sending an appropriate pulse to the qubit from the classical electronics.

Qubit initialization is a process to initialize the qubit to a particular state which is usually the ground state, $|0\rangle$. Most quantum computing algorithm starts with all qubits initialized to $|0\rangle$ [1, 2]. Moreover, in algorithms with error corrections, a steady supply of initialized qubits is required. There are two ways to initialize a qubit to $|0\rangle$. One is through **thermalization**. In this process, we let the qubit wait for a long enough time so that it will decay to the ground state by losing energy to its environment. This is typically a long process. The second approach is to measure the qubit and a pulse is applied to reset the qubit to the ground state if it is not at the ground state (i.e., $|1\rangle \rightarrow |0\rangle$). This is called the **active reset**. We will discuss qubit initialization with more details in Chaps. 11 and 22.

Qubit manipulation is equivalent to applying a quantum gate to the qubit. As mentioned, pulses with the appropriate frequency, shape, amplitude, phase, and

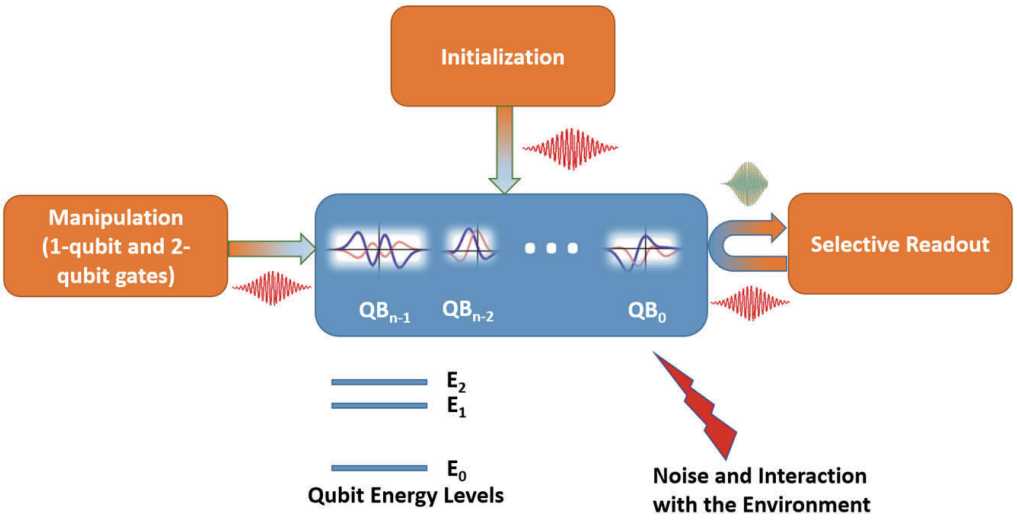

Fig. 1.1 Block diagram of a typical quantum computer

duration are sent to interact with the qubit so that the qubit wavefunction will change. If the pulse only interacts with a single qubit, it is a **1-qubit gate**. If the pulse causes two qubits to interact with each other, it is a **2-qubit gate**. Any quantum gate can be decomposed or approximated by a set of 1-qubit and 2-qubit gates (which is called **a set of universal quantum gates**). This is just like the fact that any classical logic may be implemented by NOT and AND gates. It is possible to have gates for 3 or more qubits. However, they are more difficult to implement and have lower fidelity.

Qubit readout refers to measuring all or some of the qubits. The readout is essential in quantum computing as it is used to measure the final result. In some algorithms, intermediate results need to be measured to determine the next step (e.g. in quantum teleportation, see Chapter 19 in [1]). It is also used to measure ancillary qubits for error correction in the middle of the algorithm. Due to the nature of quantum mechanics, upon measurement, a qubit wavefunction will collapse to the measurement basis states ($|0\rangle$ or $|1\rangle$). This is achieved by applying a pulse to the qubit and observing the change of the reflected pulse which is modulated by the state of the qubit ($|0\rangle$ or $|1\rangle$) in some architectures (such as superconducting qubit). In the silicon electron spin qubit case, the spin state is mapped to a charge state, and the charge state is detected by electronics.

1.3 DiVincenzo's Criteria

In his 2000 paper [3], DiVincenzo suggested that any architecture needed to have the following five criteria to make a useful quantum computer:

1. The ability to initialize the state of the qubits to a simple fiducial state

2. A "universal" set of quantum gates
3. A qubit-specific measurement capability
4. A scalable physical system with well-characterized qubit
5. Long relevant decoherence times

Items 1–3 have been discussed in Sect. 1.2. Item 1 refers to qubit initialization. Initialization needs to have a high fidelity (which is not trivial) and also needs to be fast as mentioned earlier. Item 2 requests the architecture to have a set of universal 1-qubit and 2-qubit gates that can be used to build up (or at least approximate) any quantum circuits. Item 3 requires that one can selectively measure any particular qubits without affecting others. This is required in many circuits such as quantum teleportation and Shor's algorithm [1].

Item 4 refers to the fact that the qubit must be a well-distinguished two-level system. It should only have states $|0\rangle$ and $|1\rangle$ and they need to be well-separated. This is not trivial for some systems. For example, a superconducting qubit can have states $|0\rangle$, $|1\rangle$, $|2\rangle$, $|3\rangle$, $\cdots$ (Fig. 1.1). If the energy separation between $|0\rangle$ and $|1\rangle$, $E_1 - E_0$, is too close to the energy separation between $|1\rangle$ and $|2\rangle$, $E_2 - E_1$, i.e., $E_2 - E_1 \approx E_1 - E_0$, the qubit state can leave the $|0\rangle / |1\rangle$ Hilbert space easily, resulting in a failure in the computation. We will discuss this in detail in Sect. 19.2.

$|0\rangle$ and $|1\rangle$ need to be well-separated because of two reasons. Firstly, there are thermal and **quantum noises**. When there is noise, they cannot be distinguished well unless their energy separation is well above the noise level. When the noise level is large, the qubit can transit from one state to another due to the noise. The noise will also distort the measurement signal making the state indistinguishable [4]. The second reason is due to the so-called **line-width broadening**. The energy of each level is not sharp. It has a certain spread due to **the uncertainty principle** and its interaction with the environment. Therefore, even if there is no noise, the two levels need to be well separated.

Thermal noise is proportional to kT/q, where k is the Boltzmann constant, T is the temperature, and q is the elementary charge. Figure 1.2 shows the thermal energy as a function of temperature in kelvin (K). The qubit bit energy separation ($E_1 - E_0$) of a typical superconducting qubit is about 0.02 meV which is equivalent to the thermal noise at 0.23 K $= 230$ mK. Therefore, to operate a superconducting quantum computer, a **dilution refrigerator** is needed in which the temperature is maintained at about 10 to 20 mK, corresponding to a thermal energy of about 10–20 times lower than the qubit energy separation.

Therefore, besides those mentioned in Sect. 1.2, *cryogenic technologies and electronics* are critical players in quantum computers.

Item 5 is about the **decoherence time** which is a measure of duration a qubit can maintain its state without intentional interaction (such as the initialization, manipulation, and readout in Fig. 1.1). A qubit may lose its information by interacting with the environment. There are two types of decoherence time. The first type is called the T_1*-decoherence time* or the *longitudinal decoherence time*. It is a time constant of how fast an excited state, $|1\rangle$, will decay to a ground state, $|0\rangle$. The second type is called the T_2*-decoherence time* or the *transverse decoherence time*.

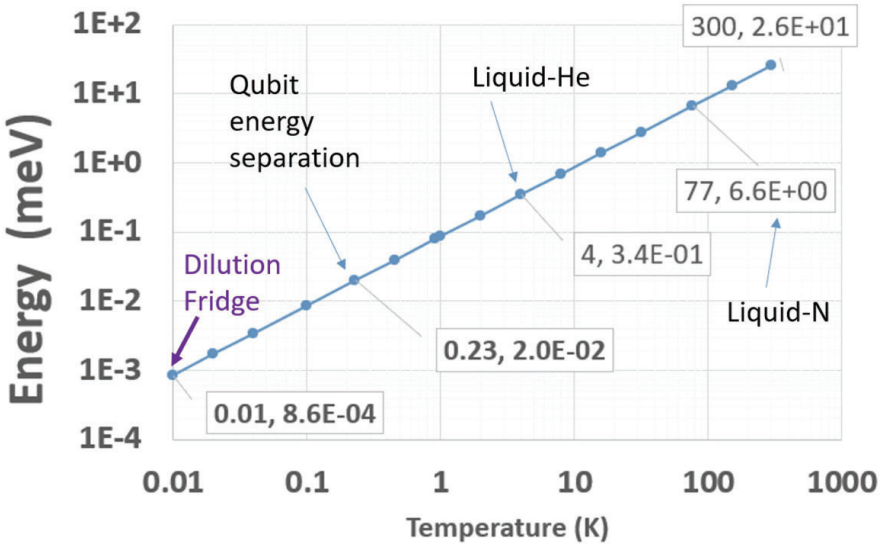

Fig. 1.2 Thermal energy as a function of temperature

It is a time constant of how fast a quantum superposition of $|0\rangle$ and $|1\rangle$ decays to a classical mixture of $|0\rangle$ and $|1\rangle$. In this process, the state loses its phase information too and thus it is also called the **dephasing time**. This will be discussed in detail in Chap. 25.

Item 5 tells us that the qubit needs to have a long decoherence time so that it can perform enough computation (many applications of quantum gates) before it loses its information. Since decoherence is related to noise, we again want to put the qubit at a cryogenic temperature to avoid the thermal noise.

There is a dilemma in DiVincenzo's Criteria. To have a long decoherence time, qubits should be well-isolated from the environment. A qubit that does not have a strong interaction mechanism with the environment usually is well-isolated. However, qubit control requires that the qubit can be coupled to the environment (the control) easily. Usually, these two cannot be achieved at the same time. Therefore, a quantum computer architecture that has a long decoherence time usually has a long gate time (as coupling to the control is weaker). Moreover, we also require dissipative coupling to the environment for initialization if active reset is not preferred.

1.4 Tracing the Signal: A Superconducting Qubit Quantum Computer Example

It would be instructive to look deeper into a quantum computer by tracing how the signals propagate. Here I will choose a superconducting qubit quantum computer which I have worked on. Figure 1.3 shows the schematic of a typical

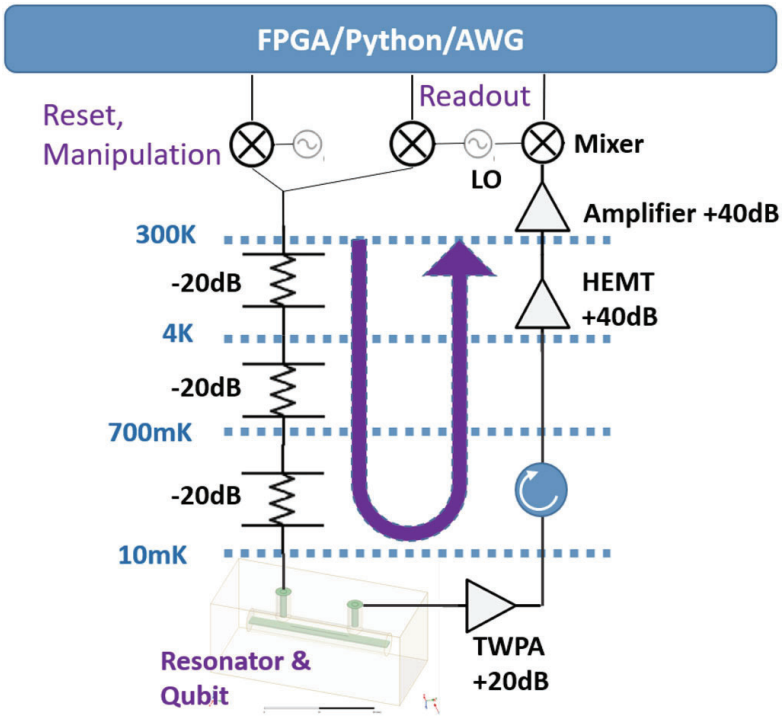

Fig. 1.3 Schematic of a typical superconducting qubit quantum computer

superconducting qubit quantum computer. Every part is classical except the qubit and the **traveling wave parametric amplifier (TWPA)** at 10 mK are quantum.

On the top at room temperature, a high-performance server in conjunction with high-speed electronics such as FPGAs is used to generate signals, analyze data, and make decisions. It controls equipment such as an arbitrary waveform generator (AWG) to generate appropriate pulses (with the right phase, shape, duration, amplitude, and frequency) for qubit initialization, manipulation, and readout. The signal will be up-converted to microwave frequency in the GHz regime through a mixer and a local oscillator (LO) which is required to interact with the qubit as the superconducting qubit operates in the GHz domain (e.g., a qubit energy of 0.02 meV corresponds to about a 5 GHz microwave photon).

The pulses will be generated with a large enough amplitude and go through a chain of attenuators from room temperature to about 10 mK. Note that the manipulation/reset pulses usually have a different frequency (so a different LO is used) than that of the readout pulse. The attenuators are required to attenuate the thermal noise from room temperature so that the thermal noise becomes negligible when it reaches the qubit. At that stage, the thermal noise is smaller than the *quantum noise*. Quantum noise is a result of the *Heisenberg's Uncertainty Principle* and cannot be avoided which poses a fundamental limit.

The pulse does not interact with the qubit directly. It often interacts with the qubit through capacitive or inductive coupling. If it is a manipulation pulse (i.e., a quantum gate) or initialization pulse for active reset, the process is completed.

However, if it is a readout pulse (i.e., a measurement pulse), the reflected or transmitted signal will be detected. In some quantum computers, the readout and manipulation/reset are coupled to the qubit through different paths. Usually, the readout pulse is passed through a resonator capacitively coupled to the qubit. The resonant frequency of the resonator will change depending on the state of the qubit ($|0\rangle$ or $|1\rangle$), and thus the reflected/transmitted pulse is changed accordingly.

When the readout pulse is reflected or transmitted, the signal is very weak and cannot be detected by classical electronics. Therefore, amplification is required. However, every amplifier has a certain noise factor which will be discussed in Chap. 23. When an amplifier amplifies the signal, it also amplifies the noise and thus the **signal-to-noise ratio**, *S/N ratio*, will not be improved. Indeed, it also adds additional noise to the amplified signal due to the noise sources in the amplifier, and thus further degrades the S/N ratio depending on its noise factor. Based on microwave theory [5], to minimize the overall noise factor, it is desirable to have an amplifier with the lowest noise factor at the beginning of the amplification chain. Therefore, a TWPA is added at the beginning. This is a type of **quantum parametric amplifier** [6] and has the minimal possible noise factor. The gain of a TWPA is usually low. Therefore, further amplifications are required. Finally, the signal is downconverted to a lower frequency using a mixer and an LO for signal processing to distinguish $|0\rangle$ and $|1\rangle$ states in classical computers and electronics.

I would also highlight two important aspects related to electrical engineering in this system. Firstly, a *high-electron-mobility transistor* (HEMT) amplifier is commonly used in quantum computers to amplify readout signals. This is because it has a low noise factor. Secondly, superconducting qubits are usually fabricated through *integrated circuit (IC)* technologies. The knowledge in IC can be applied readily to superconducting qubit design except that microwave analysis is required and it is built on materials usually not used in a traditional silicon IC chip. It is also worth noting that a superconducting qubit chip has most of the area occupied by classical microwave components such as capacitors, inductors, transmission lines, and resonators. Readers can see a design example in Chap. 24.

1.4.1 Scaling of Quantum Computers

After understanding what a quantum computer looks like, it is natural to see that the scaling of a quantum computer is not trivial. How can we increase the number of qubits to realize a powerful quantum computer? There are three issues.

Firstly, too many transmission lines need to be wired from room temperature to cryogenic temperature to control each qubit. While multiplexing is possible, this will eventually post a bottleneck to the **throughput** and speed. Throughput refers to the amount of signal that can be processed per unit amount of time. This requires us to miniaturize the classical control electronics and place them in the proximity of the qubits and, thus, need to be cooled down to cryogenic temperature.

Secondly, even if we can miniaturize the control electronics, a typical dilution refrigerator can only have a cooling power of less than 1 mW [7]. A typical classical CPU has a power of 100 W. It is thus almost impossible to put all control electronics in mK-regime. But it is still desirable to bring the electronics as close to the qubit as possible (to increase the throughput) and as cold as possible (to reduce thermal noise). Therefore, cryogenic electronics at 4.2 K (boiling point of liquid He) has been widely studied [8].

Finally, due to decoherence, error-protected qubits are required. In theory, it requires about 100–1000 physical qubits to create an error-protected qubit. This increases the scaling requirement by 100–1000 times to have a useful quantum computer.

1.5 Summary

In this chapter, we give an overview of the operations of a typical quantum computer. A quantum computer has three basic operations, namely, initialization, manipulation, and readout. Any architecture needs to fulfill DiVincenzo's criteria to be a useful quantum computer. Besides being able to perform basic operations, the qubits need to have well-separated levels and long decoherence time. We then use the superconducting qubit as an example to demonstrate the basic operations and constraints of a quantum computer. At the moment of writing, a quantum computer is still just a well-organized sophisticated physics laboratory. To build a quantum computer with many qubits and with error correction is not trivial. This requires a lot of engineering work and engineers are expected to play an important role. In the following chapters, we will review the basics of quantum computing and physics. We will then study the details of how spin qubits and superconducting qubits fulfill DiVincenzo's criteria.

Problems

1.1 Thermal Noise

If we want to operate a superconducting qubit at room temperature, what should be the energy separation between $|0\rangle$ and $|1\rangle$ (i.e., qubit energy) if we want it to be at least 10 times the thermal noise. What is the corresponding frequency? Is this feasible? If you do not know the equations, you can refer to Fig. 1.2 and the discussion in Sect. 1.3. Note that noise energy is proportional to temperature and qubit energy is proportional to the qubit frequency.

References

1. Hiu-Yung Wong. *Introduction to Quantum Computing*. Springer, 2024.
2. M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2011.
3. David P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik*, 48(9–11):771–783, 2000.
4. Hiu Yung Wong, Prabjot Dhillon, Kristin M. Beck, and Yaniv J. Rosen. A simulation methodology for superconducting qubit readout fidelity. *Solid-State Electronics*, 201:108582, 2023.
5. Behzad Razavi. *RF Microelectronics*. Pearson, 2011.
6. C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi. A near–quantum-limited josephson traveling-wave parametric amplifier. *Science*, 350(6258):307–310, 2015.
7. H. Zu, W. Dai, and A.T.A.M. de Waele. Development of dilution refrigerators–a review. *Cryogenics*, 121:103390, 2022.
8. Tom Jiao, Edwin Antunez, and Hiu Yung Wong. Study of cryogenic mosfet sub-threshold swing using ab initio calculation. *IEEE Electron Device Letters*, 44(10):1604–1607, 2023.

Chapter 2
Linear Algebra—Vectors, States, and Measurement

2.1 Introduction

In this and the next few chapters, we will review linear algebra basics. In this chapter, we will review the definitions of vector space, inner product space, and Hilbert space. Although it is not necessary to understand them thoroughly if you feel comfortable with using the equations and mathematical tools given, it is still very instructive if you want to have a basic idea of the foundation of quantum computing and quantum physics. We will also practice *bra-ket* notation and their representations in row and column vectors. We will study how to represent a vector in different bases. Then, we will discuss the effect of measurement in quantum mechanics and the importance of orthonormal basis and normalized vectors. Finally, we will review how to combine two spaces (of two physical systems) into a larger one using the tensor product.

2.1.1 Learning Outcomes

Know how to represent vectors in different bases; be more familiar with *bra-ket* operations; able to perform tensor products.

2.1.2 Teaching Videos

- Search for Ch2 in this playlist
  - https://tinyurl.com/3yhze3jn

H. Y. Wong, *Quantum Computing Architecture and Hardware for Engineers*,
https://doi.org/10.1007/978-3-031-78219-0_2

- Other videos
- https://youtu.be/Xireluiir9Y
- https://youtu.be/cEpjouPIuQY
- https://youtu.be/Wrmigi645J4

2.2 Vector Space, Inner Product, and Hilbert Space

A **vector space** (or **linear space**), in general, has a set of **vectors**, $\{\vec{u}, \vec{v}, \vec{w}, \cdots\}$, and a set of **scalars**, $\{a, b, c, \cdots\}$. There is an **addition operation**, $+$, for the vectors, resulting in another vector in the space (e.g., $\vec{w} = \vec{u} + \vec{v}$). The vector can be scaled by the scalars (e.g., $a\vec{v}$). They need to obey the following **vector axioms**:

$$\begin{aligned}
\vec{u} + (\vec{v} + \vec{w}) &= (\vec{u} + \vec{v}) + \vec{w}, \\
\vec{u} + \vec{v} &= \vec{v} + \vec{u}, \\
\vec{u} + \vec{0} &= \vec{u}, \\
\vec{v} + \vec{-v} &= \vec{0}, \\
a(b\vec{u}) &= (ab)\vec{u}, \\
1\vec{u} &= \vec{u}, \\
a(\vec{u} + \vec{v}) &= a\vec{u} + a\vec{v}, \\
(a + b)\vec{u} &= a\vec{u} + b\vec{u}.
\end{aligned} \tag{2.1}$$

We will not discuss the details and interested readers can refer to standard textbooks or Wikipedia [2]. However, we will emphasize a few things. $\vec{-v}$ refers to the requirement that an *inverse* vector must exist for $\vec{v}$. Therefore $\vec{-v}$ is just $-\vec{v}$. Moreover, there exists a **zero** vector $\vec{0}$ and there should be a **multiplicative identity**, 1, in the scalar set.

If the scalars are only real numbers, then the space is called a **real vector space**. If they include complex numbers, the corresponding space is called a **complex vector space**.

We can further require that a certain vector space *has an operation called the **inner product** defined for its vectors*. Then it is called an **inner product space**. The inner product between two vectors $\vec{u}$ and $\vec{v}$ is written as $\langle \vec{u} | \vec{v} \rangle$ *which is a scalar*. Before we continue to study the properties of inner products, let us discuss how our 3D space behaves as an inner product space.

Example 2.1 **Euclidean vector space** is an inner product space. The 3D space we live in is a Euclidean vector space. Discuss how it fulfills the definition of an inner product space.

For any position vectors in the space, $\vec{v}$ and $\vec{w}$, we already know from our daily experience that they obey Eq. (2.1) with real number scalars. So it is a vector space.

We can set $\hat{x}$, $\hat{y}$, and $\hat{z}$ as its basis vectors (see Chapter 3 in [1]) and they are **orthonormal** (to be discussed in Sect. 2.3.4). Then for any position vectors in the space, $\vec{v}$ and $\vec{w}$, they can be written as a linear combination of the basis vectors,

$$\begin{aligned} \vec{v} &= v_1\hat{x} + v_2\hat{y} + v_3\hat{z}, \\ \vec{w} &= w_1\hat{x} + w_2\hat{y} + w_3\hat{z}, \end{aligned} \tag{2.2}$$

where v_1, v_2, v_3, w_1, w_2, and w_3 are real numbers.

We define the inner product as

$$\begin{aligned} \langle v|w\rangle &= \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}, \\ &= v_1w_1 + v_2w_2 + v_3w_3. \end{aligned} \tag{2.3}$$

It is also helpful to state that the inner product in a real 3D space is equivalent to

$$\langle v|w\rangle = \vec{v}\cdot\vec{w} = |\vec{v}||\vec{w}|\cos\theta, \tag{2.4}$$

where θ is the angle between the two vectors. And it should be noted that $\langle v|w\rangle$ is just the $bra-ket$ notation for the **dot product**, $\vec{v}\cdot\vec{w}$, that we usually use.

Based on our daily experience, it is easy to understand that the inner product in a real 3D space satisfies the following equations:

$$\begin{aligned} \vec{v}\cdot\vec{w} &= \vec{w}\cdot\vec{v}, \\ \vec{v}\cdot\vec{v} &\geq 0, \\ \vec{u}\cdot(a\vec{v} + b\vec{w}) &= a\vec{u}\cdot\vec{v} + b\vec{u}\cdot\vec{w}, \end{aligned} \tag{2.5}$$

where the second equation just tells us that any vector has a zero or positive length. ■

Now, we formally define a general **Euclidean vector space** as a *finite inner product space* with *real* scalars. A *finite* inner product space means that the number of basis vectors (e.g., 3 in our 3D real space) is finite instead of infinite.

Based on our experience with the real 3D Euclidean vector space, we can generalize the criteria of the inner product in Eq. (2.5) to any inner product space with complex scalars and with an arbitrary number of dimensions. For simplicity, we will remove the arrow in the vector notation now. Since $\langle\vec{u}|\vec{v}\rangle = \vec{u}\cdot\vec{v}$ can be a complex number, the definition of an inner product is an operation that obeys the

following rules:

$$\begin{aligned} \langle v|w\rangle &= \langle w|v\rangle^* , \\ \langle v|v\rangle &\geq 0, \\ \langle u|\left(a\,|v\rangle + b\,|w\rangle\right) &= a\,\langle u|v\rangle + b\,\langle u|w\rangle . \end{aligned} \tag{2.6}$$

We can use the inner product to define another quantity called **norm**, which is

$$\|v\| = \sqrt{\langle v|v\rangle}, \tag{2.7}$$

and is just the length of the vector and sometimes we write it as $|\vec{v}|$ in the real 3D space (e.g., Eq. (2.4)).

We then define the **distance** between two vectors, $\vec{u}$ and $\vec{v}$, as

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v|u - v\rangle}, \tag{2.8}$$

which is just the *norm* or the length of their difference.

I introduce *norm* and *distance* because I would like to give a more formal definition of **Hilbert Space**. A Hilbert space is a *real* or *complex* inner product space that is *complete* with respect to its distance. The definition of "completeness" here is *different* from the completeness of basis vector (e.g., Section 10.4 in [1]). It is pretty involving mathematically. We can approximately explain it as the following. There is something called the **Cauchy sequence**. This sequence has its elements eventually become arbitrarily close to the next one (converged). For an inner product space, we can create many Cauchy sequences of its *vectors*. Each Cauchy sequence will converge to a vector (in other words, its distance to other elements will converge to 0). If the converged vector *is also a vector in the inner product space*, i.e., the inner product space contains that converged vector, then it is said to be complete. Since we measure the convergence using distance, it is said to be complete with respect to its distance.

Note that any finite-dimensional inner product space is complete. Therefore, a Euclidean vector space is a Hilbert space because it is complete.

Figure 2.1 summarizes the properties of the spaces discussed.

	Has vectors and scalars which obey (2.1)	Has inner product defined that obeys (2.6)	Is complete?	Has infinite dimensions?	Has complex scalars?
Vector Space	✓	Maybe	Maybe	Maybe	Maybe
Inner Product Space	✓	✓	Maybe	Maybe	Maybe
Hilbert Space	✓	✓	✓	Maybe	Maybe
Euclidean Vector Space	✓	✓	✓	×	×

Fig. 2.1 Summary of the properties of various spaces discussed in this section

2.3 Review of Vector Basics

Here we will succinctly list the basic concepts in linear algebra that are relevant to quantum computing. For a more detailed discussion, please see [1].

2.3.1 Basis Vectors/States, Bra-ket Notation, and Representation of Vectors/States

A state in a quantum system is represented by a vector. In other words, every vector in a Hilbert space is also a state. We may choose some vectors to be the **basis vectors** and every vector in the space can be represented as a linear combination of the basis vectors. For example, in real 3D space, a vector $\vec{v}$ can be expressed as the combination of the unit vectors (as basis vectors), $\hat{x}$, $\hat{y}$, and $\hat{z}$, as

$$\vec{v} = v_1\hat{x} + v_2\hat{y} + v_3\hat{z} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \tag{2.9}$$

where v_1, v_2, and v_3 are scalars.

To write a vector, besides putting an arrow on top of the variable (e.g., $\vec{v}$),

$$\vec{v} = \alpha\hat{v_x} + \beta\hat{v_y}, \tag{2.10}$$

we may also use the *bra-ket* notation in which the variable is written in a *ket*,

$$|v\rangle = \alpha\,|v_x\rangle + \beta\,\left|v_y\right\rangle. \tag{2.11}$$

It should be noted that anything inside the *ket* is just a name. As long as it is not ambiguous, we may write anything that is convenient. For example, if $|v\rangle$ is the vector for a photon polarization state, and $|v_x\rangle$ and $\left|v_y\right\rangle$ are the horizontal and vertical polarization states of a photon, respectively, we may write Eq. (2.11) as,

$$\begin{aligned} |\text{Final Polarization}\rangle = \; & \alpha\,|\text{Horizontal Polarization}\rangle + \\ & \beta\,|\text{Vertical Polarization}\rangle\,. \end{aligned} \tag{2.12}$$

If a different basis vector set is chosen, the representation will be different. For example, in Fig. 2.2, $|V_1\rangle$ can be expressed as a linear combination of the basis vectors in one of the two bases, $|0\rangle$ / $|1\rangle$ or $|+\rangle$ / $|-\rangle$.

$$|V_1\rangle = 1\,|0\rangle + 0\,|1\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{matrix} \rightarrow |0\rangle \\ \rightarrow |1\rangle \end{matrix} \tag{2.13}$$

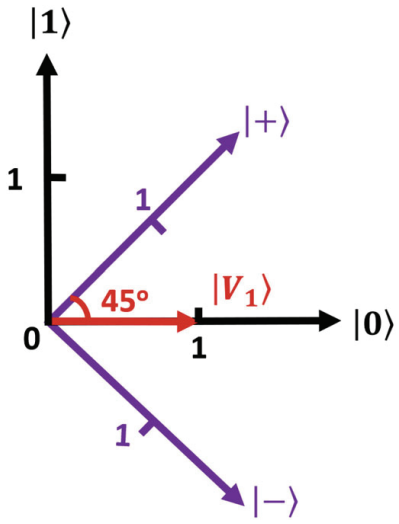

Fig. 2.2 Vector $|V_1\rangle$ can be represented in two different bases, $|0\rangle$ / $|1\rangle$ or $|+\rangle$ / $|-\rangle$

$$= \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\begin{matrix}\cdot \to |+\rangle\\ \to |-\rangle\end{matrix} \tag{2.14}$$

It should be noted that *it has different column vectors in different bases*. This is because each row represents the amount of basis vector, indicated after the right arrows, the vector contains. Also, we have used the $bra - ket$ notation. It should be emphasized that while we label the basis vectors as $|0\rangle$, $|1\rangle$, $|+\rangle$, and $|-\rangle$, they are just some given notations. They are NOT the $|0\rangle$, $|1\rangle$, $|+\rangle$, and $|-\rangle$ in a qubit space, even though they are related through the same equations.

Example 2.2 Change of Basis: Derive Eq. (2.14) from Eq. (2.13).

Here we assume that we already know that $|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ and $|1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$. We could have derived them using trigonometry in Fig. 2.2 but we assume the readers know how to do it. Then,

$$\begin{aligned}|V_1\rangle &= 1\,|0\rangle + 0\,|1\rangle\,,\\ &= 1\left(\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)\right) + 0\left(\frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)\right),\\ &= \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle).\end{aligned} \tag{2.15}$$

■

In Sect. 3.3.5, we will give a transformation matrix to facilitate basis transformation and discuss its properties.

2.3.2 High-Dimensional Vectors

For an n-dimensional space, it has a basis of n basis vectors. Every vector, $\vec{a} = |a\rangle$, in that space, can be expressed as a linear combination of the basis vectors as

$$\vec{a} = a_0\hat{x_0} + a_1\hat{x_1} + \cdots + a_{n-1}\hat{x_{n-1}}, \tag{2.16}$$

or in *bra-ket* notation, it is,

$$\begin{aligned} |a\rangle &= a_0\,|x_0\rangle + a_1\,|x_1\rangle + \cdots + a_{n-1}\,|x_{n-1}\rangle\,, \\ &= a_0\,|0\rangle + a_1\,|1\rangle + \cdots + a_{n-1}\,|n-1\rangle\,, \\ &= \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}, \end{aligned} \tag{2.17}$$

where in line 2, we changed the names of the basis states to emphasize that whatever is put inside the *ket* is just a name. As long as it is not confusing, it does not matter how we name it. When we write the vector in a column form, we have assumed that the basis vectors are orthonormal, which will be discussed soon in Sect. 2.3.4.

Each vector has a corresponding vector in the *bra*-space (**dual correspondence**). This is similar to the fact that every object has an image in the mirror. The *bra* of $|b\rangle$ is written as $\langle b|$. And to construct the *bra* version of $|b\rangle$ in matrix form, we need to perform **conjugate transpose**. That is to swap the rows and columns and then apply complex conjugate to each element. Therefore, the column vector has a row vector in its *bra* version. For example, vector $|b\rangle$, which is expressed as

$$|b\rangle = b_0\,|x_0\rangle + b_1\,|x_1\rangle + \cdots + b_{n-1}\,|x_{n-1}\rangle = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{pmatrix}, \tag{2.18}$$

has its *bra* version expressed as

$$\begin{aligned} \langle b| &= b_0^*\,\langle x_0| + b_1^*\,\langle x_1| + \cdots + b_{n-1}^*\,\langle x_{n-1}|\,, \\ &= \begin{pmatrix} b_0^* & b_1^* \cdots & b_{n-1}^* \end{pmatrix}. \end{aligned} \tag{2.19}$$

The *bra-ket* notation is very useful in linear algebra. For example, the inner product of two vectors, $|b\rangle$ and $|a\rangle$, is just the multiplication between the *bra* of

$|b\rangle$ and the *ket* of $|a\rangle$,

$$\langle b|a\rangle = a_0 b_0^* + a_1 b_1^* + \cdots + a_{n-1} b_{n-1}^*, \tag{2.20}$$

which is the same as how we wrote it in Eq. (2.3).

Example 2.3 For $|a\rangle = \begin{pmatrix} 3i+2 \\ 0 \\ 5 \\ 4-2i \end{pmatrix}$ and $|b\rangle = \begin{pmatrix} 2 \\ i \\ 0 \\ 2i \end{pmatrix}$, find $\langle a|b\rangle$.

$$\langle a|b\rangle = \begin{pmatrix} -3i+2 & 0 & 5 & 4+2i \end{pmatrix} \begin{pmatrix} 2 \\ i \\ 0 \\ 2i \end{pmatrix},$$

$$= (-6i+4) + 0 + 0 + (8i-4) = 2i.$$

■

2.3.3 *Measurement of a Quantum State*

Measurement is not a part of linear algebra. However, I would like to interject this topic so that we can understand the following sections better. The measurement of a quantum state results in the **collapse of the state** to one of its basis states. That means that the measurement outcome is one of the basis states and the original quantum state no longer exists. The process is completely random except that the probability it will collapse to a certain basis vector is the square of the magnitude of the corresponding coefficient (Fig. 2.3). For example, for $|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, the probability it will collapse to $|0\rangle$ is

$$\text{Prob}(|0\rangle) = \alpha\alpha^* = |\alpha|^2, \tag{2.21}$$

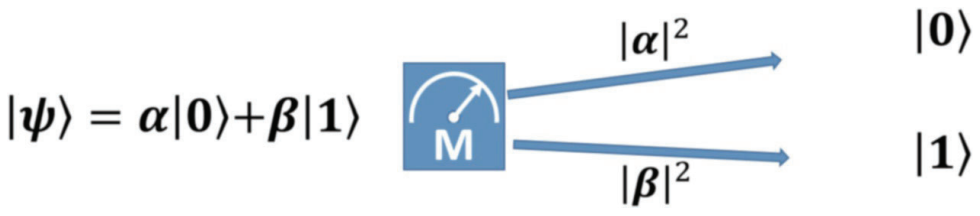

Fig. 2.3 Upon measurement, a state will collapse to one of the basis states with a probability equal to the square of the magnitude of the corresponding coefficient

and the probability it will collapse to $|1\rangle$ is

$$\text{Prob}(|1\rangle) = \beta\beta^* = |\beta|^2. \tag{2.22}$$

2.3.4 Orthonormal Basis and Vector Normalization

An **orthonormal basis** is a basis with orthonormal basis vectors. Let it be an n-dimensional space and thus it has n basis vectors, $|x_0\rangle, |x_1\rangle, \cdots, |x_{n-1}\rangle$. If each basis vector is normalized (with a length of 1) and orthogonal to the others (with 0 overlap or inner product with other basis vectors), it is called an orthonormal basis. This can be written as

$$\begin{aligned}\langle x_i | x_j \rangle &= \begin{cases} 0 & if\ i \neq j, \text{ (orthogonal)} \\ 1 & if\ i = j, \text{ (normalized)} \end{cases} \\ &= \delta_{ij}, \end{aligned} \tag{2.23}$$

where we use the **Kronecker delta** in the last line. Note that when $i = j$, $\langle x_i | x_j \rangle = \langle x_i | x_i \rangle$ and this is just the square of the norm of the basis vector, $|x_i\rangle$ (Eq. (2.7)). If it is one, then it means that the length is also one. Working on an orthonormal basis provides a lot of convenience in calculations due to the fact that $\langle x_i | x_j \rangle$ results in either 0 or 1. We can also thus write the vector in a column or row form as in Eqs. (2.18) and (2.19).

For example, if a vector $|V\rangle$ represented in an orthonormal basis, $|x_i\rangle$, is given by

$$|V\rangle = a_0 |x_0\rangle + \cdots + a_{n-1} |x_{n-1}\rangle, \tag{2.24}$$

to find its norm squared, $\|v\|^2$ (Eq. (2.7)), we have

$$\begin{aligned}\langle v|v\rangle &= (a_0^* a_1^* \cdots a_{n-1}^*) \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}, \\ &= a_0^* a_0 + a_1^* a_1 + \cdots + a_{n-1}^* a_{n-1}, \\ &= |a_0|^2 + |a_1|^2 + \cdots + |a_{n-1}|^2. \end{aligned} \tag{2.25}$$

If $\langle v|v\rangle = 1$, then vector $|v\rangle$ is a **normalized vector**. As shown in Eq. (2.25), this means that a normalized vector has the sum of the coefficient modulus squared equals one. Recalling that upon the measurement of a quantum state, the probability is the corresponding coefficient modulus squared (Eqs. (2.21) and (2.22)), then a

quantum state must be normalized so that the probability of collapsing to *any* of the basis states is one. In other words, this ensures the sum of the probabilities of measuring one of the basis states to be one. Therefore, any quantum state must be a normalized vector.

2.4 Tensor Product

We can combine two or more vector spaces through **tensor product**. For a more detailed discussion, please refer to Chapters 11 and 12 in [1]. What is the meaning of *combing two vector spaces* and why do we want to do that? For example, we can describe the spin of an electron using a 2D Hilbert space. It has two basis vectors, $|0\rangle_1$ and $|1\rangle_1$. Here I used subscript 1 to indicate that this is the vector space belonging to the first electron. The state of any possible spin of the electron is a vector, $|\psi_1\rangle$, in this Hilbert space and is a linear combination of the basis states,

$$|\psi_1\rangle = \alpha_1 |0\rangle_1 + \beta_1 |1\rangle_1 . \tag{2.26}$$

Similarly, if there is a second electron, its basis vectors are $|0\rangle_2$ and $|1\rangle_2$. Its state is given by,

$$|\psi_2\rangle = \alpha_2 |0\rangle_2 + \beta_1 |1\rangle_2 . \tag{2.27}$$

If we want to describe the two electrons together or treat the two electrons as a *single* physical system, then the tensor product is the mathematical tool for us to do so.

As aligned with our common sense, the new system must have a larger vector space. Here, *let me emphasize* that a larger space is *NOT* obtained through a simple extension of a lower-dimension one to a higher-dimension one (e.g., adding a time dimension to the 3D space to become a 4D space-time). It is a result of the tensor product of two lower-dimension spaces. The basis states will be expanded as a tensor product, $\otimes$, of the lower space basis states. The number of the new basis states is the number of the permutations of the lower space basis states. For example, the two-electron system as a whole has four basis states, $|0\rangle_1 \otimes |0\rangle_2$, $|0\rangle_1 \otimes |1\rangle_2$, $|1\rangle_1 \otimes |0\rangle_2$, and $|1\rangle_1 \otimes |1\rangle_2$. We may also omit $\otimes$ by writing it as $|0\rangle_1 |0\rangle_2$, $|0\rangle_1 |1\rangle_2$, $|1\rangle_1 |0\rangle_2$, and $|1\rangle_1 |1\rangle_2$. And if we agree with each other that the first (second) number refers to the first (second) electron, we can also succinctly write it as $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$.

With the new space, we also expect that a state in the combined system must be a linear combination of the new basis vectors. This can be seen clearly by considering the tensor product of the two-electron system. The following demonstrates how to perform tensor products without explaining the background. Readers can treat it as a result of the definitions and should appreciate its similarity to a regular algebraic product.

For example, if the first electron is in state $|\psi_1\rangle$ and the second electron is in state $|\psi_2\rangle$, then the state of the whole system, $|\psi\rangle$, is obtained through the tensor product of $|\psi_1\rangle$ and $|\psi_2\rangle$ (Eqs. (2.26) and (2.27)):

$$\begin{aligned}
|\psi\rangle &= |\psi_1\rangle \otimes |\psi_2\rangle\,, \\
&= (\alpha_1 |0\rangle_1 + \beta_1 |1\rangle_1) \otimes (\alpha_2 |0\rangle_2 + \beta_2 |1\rangle_2), \\
&= \alpha_1\alpha_2 |0\rangle_1 |0\rangle_2 + \alpha_1\beta_2 |0\rangle_1 |1\rangle_2 + \beta_1\alpha_2 |1\rangle_1 |0\rangle_2 + \beta_1\beta_2 |1\rangle_1 |1\rangle_2\,, \\
&= \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 |01\rangle + \beta_1\alpha_2 |10\rangle + \beta_1\beta_2 |11\rangle\,. \qquad (2.28)
\end{aligned}$$

We can also do this in matrix form,

$$\begin{aligned}
|\psi\rangle &= |\psi_1\rangle \otimes |\psi_2\rangle\,, \\
&= \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \otimes \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}, \\
&= \begin{pmatrix} \alpha_1 \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} \\ \beta_1 \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} \end{pmatrix}, \\
&= \begin{pmatrix} \alpha_1\alpha_2 \\ \alpha_1\beta_2 \\ \beta_1\alpha_2 \\ \beta_1\beta_2 \end{pmatrix} \cdot \begin{matrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{matrix} \qquad (2.29)
\end{aligned}$$

where in the last line, the corresponding basis states are indicated. The same methodology is used for higher-dimensional spaces.

If more than 2 spaces need to be combined, we can do this one after another.

2.5 Summary

We review the basic properties of vectors in various vector spaces. In quantum computing, we will work in the Hilbert space. Therefore, the inner product which is an important component of the Hilbert space plays an important role in all calculations. We also discuss the measurement of a quantum state. Although measurement is not a part of linear algebra, it requires that all quantum state needs to be normalized. We also practice how to combine two subsystems into a larger one using tensor product. In the next chapter, we will discuss more advanced linear algebra. We will discuss matrices and operators and their applications in quantum computing.

Problems

2.1 Vector Space

a is a scalar and $|W\rangle$ is a vector. Given that,

$$|aW\rangle = a\,|W\rangle\,, \tag{2.30}$$

using also Eq. (2.1), prove the following equations:

$$\langle V|aW\rangle = a\,\langle V|W\rangle\,. \tag{2.31}$$

$$\langle aV|W\rangle = a^*\,\langle V|W\rangle\,. \tag{2.32}$$

2.2 Orthonormal Basis

Prove Eq. (2.25) using *bra-ket* notation (e.g., Eq. (2.24)) instead of using matrix form.

2.3 Tensor Product

Find the tensor product of $|a\rangle$ and $|b\rangle$ in Example 2.3.

References

1. Hiu-Yung Wong. *Introduction to Quantum Computing*. Springer, 2024.
2. Vector space. https://en.wikipedia.org/wiki/Vector_space. Accessed:2024-01-08.

Chapter 3
Linear Algebra—Operators, Matrices, and Quantum Gates

3.1 Introduction

Vectors are the fundamentals in a Hilbert space. They represent the state of a physical system corresponding to that Hilbert space. But a state or a vector is *passive*. What is interesting is the application of an operator to rotate a vector in the space or, in other words, to transform the state. An operator corresponds to a matrix. In this chapter, we will review the concepts of eigenvalue, eigenvector, Hermitian matrix, and unitary matrix. We will also review how to construct a projection operator and a unitary transformation matrix. Then we will revisit the meaning of a measurement in a quantum system using the new knowledge we have learned. Finally, we will discuss how to perform a tensor product for matrices.

3.1.1 Learning Outcomes

Understand the definitions of Hermitian matrix, unitary matrix, and project matrix; able to transform vectors and matrices from one basis to another; appreciate that applying an operator to a vector is to rotate the vector in its Hilbert space.

3.1.2 Teaching Videos

- Search for Ch3 in this playlist
  - https://tinyurl.com/3yhze3jn

H. Y. Wong, *Quantum Computing Architecture and Hardware for Engineers*,
https://doi.org/10.1007/978-3-031-78219-0_3

- Other videos
  - https://youtu.be/Wrmigi645J4
  - https://youtu.be/Z_fXDssH2JA

3.2 Operators

An **operator** in a Hilbert space maps the vectors in one space to another space. However, sometimes, the second space is the same as the first one. In this case, an operator can be considered to be rotating a vector in its space. Here we will consider the operators that map the vectors to the same space. For example, an operator $\boldsymbol{M}$ applied to vector $|\alpha\rangle$ may give another vector $\left|\alpha'\right\rangle$,

$$\left|\alpha'\right\rangle = \boldsymbol{M}\,|\alpha\rangle\,. \tag{3.1}$$

If the vector is represented as a column vector, it is natural that the operator must be a matrix.

The operators are linear and observe the distribution law,

$$\boldsymbol{M}(C_\alpha\,|\alpha\rangle + C_\beta\,|\beta\rangle) = C_\alpha \boldsymbol{M}\,|\alpha\rangle + C_\beta \boldsymbol{M}\,|\beta\rangle\,, \tag{3.2}$$

where C_α and C_β are complex scalars and $|\beta\rangle$ is another vector.

3.2.1 Dual Correspondence

We introduced the concept of **dual correspondence** in Chap. 2 that every vector $|\alpha\rangle$ in the *ket* space has a corresponding vector $\langle\alpha|$ in the *bra* space. Therefore, there is a corresponding *bra* version for $\left|\alpha'\right\rangle$, i.e., $\left\langle\alpha'\right|$, in Eq. (3.1),

$$\left\langle\alpha'\right| \quad \Leftrightarrow \quad \left|\alpha'\right\rangle. \tag{3.3}$$

Is there a dual correspondence for the *ket*-space operator $\boldsymbol{M}$? Yes, it is the **adjoint** of $\boldsymbol{M}$, i.e.. $\boldsymbol{M}^\dagger$. The adjoint of a matrix is just the **conjugate transpose** of that matrix. This is the same as how we find the *bra* version of a vector (e.g., Eq. (2.19)). Therefore, Eq. (3.3) can be rewritten as,

$$\langle\alpha|\,\boldsymbol{M}^\dagger \quad \Leftrightarrow \quad \boldsymbol{M}\,|\alpha\rangle\,. \tag{3.4}$$

Note that the operator is applied from the **right** of the *bra* vector. This can be easily appreciated by recognizing that a *bra* vector is a row vector instead of a column vector.

The dual equation of Eq. (3.2) is thus,

$$(C_\alpha^* \langle\alpha| + C_\beta^* \langle\beta|)\boldsymbol{M}^\dagger = C_\alpha^* \langle\alpha| \boldsymbol{M}^\dagger + C_\beta^* \langle\beta| \boldsymbol{M}^\dagger. \tag{3.5}$$

Note that complex conjugate is taken for scalars C_α and C_β in the *bra* version.

Based on the first line in Eq. (2.6), we can further derive that

$$\begin{aligned} \langle\beta|\alpha'\rangle &= \langle\alpha'|\beta\rangle^* , \\ \langle\beta| \boldsymbol{M} |\alpha\rangle &= \langle\alpha| \boldsymbol{M}^\dagger |\beta\rangle^* . \end{aligned} \tag{3.6}$$

These are some useful equations we will use in quantum mechanics.

3.3 Matrices

We mentioned that operators can be represented as matrices when the vectors are represented as row or column vectors. Therefore, it is important to understand some of the important properties of matrices.

3.3.1 Eigenvalues and Eigenvectors

A matrix maps (transforms) a vector to another vector. For a given matrix, there is a set of vectors to which it only scales by a scalar when it is applied. These vectors are called the **eigenvectors** of the matrix. The corresponding amounts it scales are the **eigenvalues** of the matrix. For example, if $|i\rangle$ is an eigenvector of $\boldsymbol{M}$, then

$$\boldsymbol{M} |i\rangle = \lambda_i |i\rangle , \tag{3.7}$$

where λ_i is a scalar and the eigenvalue of $\boldsymbol{M}$, corresponding to the eigenvector, $|i\rangle$.

For an $n \times n$ matrix, it has n eigenvalues (counting multiplicities) over the complex field (the eigenvalues can be complex or real). For the same operator, it can be represented in a different matrix form if a different basis is chosen. For some matrices (**diagonalizable** matrices), if the eigenvectors are chosen to be the basis states (**eigenbasis**), then the matrix is a diagonal matrix with the eigenvalues along the diagonal,

$$\boldsymbol{M} = \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} \end{pmatrix}_{In\ M's\ eigenbasis} . \tag{3.8}$$

The process of finding the eigenbasis so that the matrix is in a diagonal form is called the **diagonalization**. Diagonalization is a very important tool in solving the Schrödinger equation. Note again that *not all matrices are diagonalizable*. Readers are encouraged to refer to Section 9.2 in [1] to review how to find the eigenvalues and eigenvectors and, thus, the diagonalization of a matrix.

If the eigenvectors and eigenvalues are given, we can also construct the matrix from the eigenvectors and eigenvalues using this equation,

$$\boldsymbol{M} = \sum_{i=0}^{n-1} \lambda_i \, |i\rangle \, \langle i| \, . \tag{3.9}$$

This is trivial if the matrix is in the eigenbasis which has the form of Eq. (3.8). This is still true in general and can be proved by using the basis transformation to be discussed in Sect. 3.3.5.

3.3.2 Hermitian Matrix

We discussed earlier that the adjoint of an operator $\boldsymbol{M}$ is written as $\boldsymbol{M}^\dagger$. When it is written as a matrix, the adjoint of $\boldsymbol{M}$ is its conjugate transpose. If the adjoint of a matrix equals itself, it is also called a **self-adjoint** or **Hermitian** matrix.

Example 3.1 Show that $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is Hermitian.

$$\begin{aligned} \sigma_y^\dagger &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^{T*}, \\ &= \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}^{*}, \\ &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y. \end{aligned} \tag{3.10}$$

Therefore, it is Hermitian. Here σ_y^{T*} refers to applying a transpose operation followed by complex conjugation to σ_y. ■

3.3.3 Projection Operator

A projection operator, $\boldsymbol{P}$, is an operator that satisfies the following equation:

$$\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}, \tag{3.11}$$

which means that applying it twice is the same as applying it once (**idempotent**). For our purpose, we want to be more specific on what it does. Therefore, we will label it as $\boldsymbol{P}_{|v\rangle}$ to indicate that it can be an operator to extract the $|v\rangle$ component from any vectors. $|v\rangle$ needs to be a *normalized* vector. Therefore, $\langle v|v\rangle = 1$ (see Eq. (2.25) and after). For example, $\boldsymbol{P}_{|v\rangle}\,|\alpha\rangle$ should give us the $|v\rangle$ component in $|\alpha\rangle$. To construct $\boldsymbol{P}_{|v\rangle}$, we can use this equation,

$$\boldsymbol{P}_{|v\rangle} = |v\rangle\,\langle v|\,. \tag{3.12}$$

Let us try two examples to understand better.

Example 3.2 Show $\boldsymbol{P}_{|v\rangle} = \boldsymbol{P}_{|v\rangle}\boldsymbol{P}_{|v\rangle}$.

$$\begin{aligned}\boldsymbol{P}_{|v\rangle}\boldsymbol{P}_{|v\rangle} &= (|v\rangle\,\langle v|)(|v\rangle\,\langle v|),\\ &= |v\rangle\,(\langle v|v\rangle)\,\langle v|\,,\\ &= |v\rangle\,\langle v|\,,\\ &= \boldsymbol{P}_{|v\rangle}.\end{aligned} \tag{3.13}$$

Therefore, as long as $|v\rangle$ is a *normalized* vector, $\boldsymbol{P}_{|v\rangle}$ satisfies the definition of a projection operator. ■

Example 3.3 In a 1-qubit system, a general state can be written as $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Find the $|0\rangle$ component of $|\psi\rangle$ using the corresponding projection operator.

It is trivial that the $|0\rangle$ component of $|\psi\rangle$ is $\alpha\,|0\rangle$ from the given expression. Let us use the projection operator to find it, too. Firstly, we recall that the column and row representations of $|0\rangle$ and $\langle 0|$ are,

$$\begin{aligned}|0\rangle &= \begin{pmatrix}1\\0\end{pmatrix},\\ \langle 0| &= \begin{pmatrix}1\ 0\end{pmatrix}.\end{aligned} \tag{3.14}$$

Therefore, the projection operator for $|0\rangle$ is

$$\begin{aligned}\boldsymbol{P}_{|\boldsymbol{0}\rangle} &= |0\rangle\,\langle 0|\,,\\ &= \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\ 0\end{pmatrix},\\ &= \begin{pmatrix}1\ 0\\0\ 0\end{pmatrix}.\end{aligned} \tag{3.15}$$

We may use two methods to find the answer. Firstly, by using *bra-ket* notation, we have,

$$\begin{aligned} P_{|0\rangle}\,|\psi\rangle &= |0\rangle\,\langle 0|\,(\alpha\,|0\rangle + \beta\,|1\rangle), \\ &= |0\rangle\,(\alpha\,\langle 0|0\rangle + \beta\,\langle 0|1\rangle), \\ &= |0\rangle\,\alpha = \alpha\,|0\rangle\,. \end{aligned} \tag{3.16}$$

We may also use the matrix method and we have,

$$\begin{aligned} P_{|0\rangle}\,|\psi\rangle &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \\ &= \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \\ &= \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha\,|0\rangle\,. \end{aligned} \tag{3.17}$$

Both methods give the same result as expected. ■

3.3.4 Unitary Matrix

A **unitary matrix**, $\boldsymbol{U}$, is a matrix that satisfies the following equations:

$$\begin{aligned} \boldsymbol{U}\boldsymbol{U}^{\dagger} &= \boldsymbol{U}^{\dagger}\boldsymbol{U} = \boldsymbol{I}, \\ \boldsymbol{U}^{\dagger} &= \boldsymbol{U}^{-\mathbf{1}}. \end{aligned} \tag{3.18}$$

Unlike a Hermitian matrix which is equal to its adjoint, a unitary matrix has its **inverse** equal to its adjoint. The most important property of a unitary matrix is that it preserves the inner product of two vectors when both vectors are transformed by the same unitary matrix. For example, after the transformation, vectors $|g\rangle$ and $|f\rangle$ become $\left|g'\right\rangle = \boldsymbol{U}\,|g\rangle$ and $\left|f'\right\rangle = \boldsymbol{U}\,|f\rangle$, respectively. The inner product of the new vectors is

$$\begin{aligned} \left\langle g'\middle|f'\right\rangle &= (\langle g|\,\boldsymbol{U}^{\dagger})(\boldsymbol{U}\,|f\rangle), \\ &= \langle g|\,(\boldsymbol{U}^{\dagger}\boldsymbol{U})\,|f\rangle\,, \\ &= \langle g|\,\boldsymbol{I}\,|f\rangle\,, \\ &= \langle g|f\rangle\,, \end{aligned} \tag{3.19}$$

where Eqs. (3.4) and (3.18) are used in line 1 and line 3, respectively. Since a unitary matrix preserves the inner product of two vectors, it also *preserves the norm* of any vector as the norm of a vector is just the square root of the inner product of the vector to itself (Eq. (2.7)). Therefore, later we will see that a quantum gate must be unitary so that the state vector norm is not changed after each operation and keeps normalized.

It should also be noted that when a unitary matrix is written in matrix form, each of its columns is a normalized vector and is orthogonal to other columns. This is the same for the rows. This means that if the matrix is,

$$\begin{aligned} \boldsymbol{U} &= \begin{pmatrix} b_{0,0} & b_{0,1} & \cdots & b_{0,n-1} \\ b_{1,0} & b_{1,1} & \cdots & b_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n-1,0} & b_{n-1,1} & \cdots & b_{n-1,n-1} \end{pmatrix}, \\ &= \left(|v_0\rangle \ |v_1\rangle \cdots |v_{n-1}\rangle \right), \end{aligned} \tag{3.20}$$

where we have set

$$|v_i\rangle = \begin{pmatrix} b_{0,i} \\ b_{1,i} \\ \vdots \\ b_{n-1,i} \end{pmatrix}, \tag{3.21}$$

then we have

$$\langle v_i | v_j \rangle = \delta_{i,j}. \tag{3.22}$$

3.3.5 *Transformation of Basis*

Sometimes we want to work on a different basis for convenience. Then we need to perform an appropriate transformation of the vectors and matrices. For example, in Fig. 3.1, vector $|V\rangle$ might be originally represented in the old basis $|0\rangle / |1\rangle$ as $|V\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$. We want to find its representation in a new basis $|0'\rangle / |1'\rangle$ and it might be $|V\rangle = \alpha'_0 |0'\rangle + \alpha'_1 |1'\rangle$. We have done something similar in Fig. 2.2. Here, we want to show an equation to help us perform the transformation.

Suppose an n-dimensional vector is represented in a vector form with the basis vectors in the old basis being $|0\rangle$, $|1\rangle$, $\cdots$, $|n-1\rangle$. Now we want to represent it in a new basis with basis vectors $|0'\rangle$, $|1'\rangle$, $\cdots$, $|n-1'\rangle$. The transformation matrix

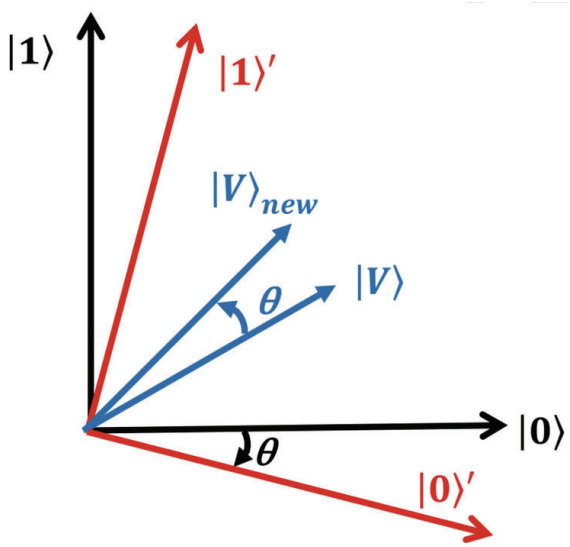

Fig. 3.1 Representation of vector $|V\rangle$ in the new basis $|0'\rangle / |1'\rangle$ is the same as the representation of vector $|V\rangle_{new}$ in the old basis $|0\rangle / |1\rangle$

to represent a vector in the new basis is given by

$$\boldsymbol{U} = \begin{pmatrix} \langle 0'|0\rangle & \langle 0'|1\rangle & \cdots & \langle 0'|n-1\rangle \\ \langle 1'|0\rangle & \langle 1'|1\rangle & \cdots & \langle 1'|n-1\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n-1'|0\rangle & \langle n-1'|1\rangle & \cdots & \langle n-1'|n-1\rangle \end{pmatrix}. \tag{3.23}$$

By using this matrix, we can find the coefficients of the vector in the new basis through matrix multiplication.

$$\begin{pmatrix} \alpha'_0 \\ \alpha'_1 \\ \vdots \\ \alpha_{n-1'} \end{pmatrix} = \begin{pmatrix} \langle 0'|0\rangle & \langle 0'|1\rangle & \cdots & \langle 0'|n-1\rangle \\ \langle 1'|0\rangle & \langle 1'|1\rangle & \cdots & \langle 1'|n-1\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n-1'|0\rangle & \langle n-1'|1\rangle & \cdots & \langle n-1'|n-1\rangle \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{pmatrix}. \tag{3.24}$$

We may better appreciate the meaning of this equation if we realize that the i-th row of the left-hand side (α'_i), which represents the amount of $|i'\rangle$ component in the new basis, is the sum of the amount of each component in the old basis (e.g., α_j) weighted by their overlaps (commons) with $|i'\rangle$, i.e., $\langle i'|j\rangle$.

Equation (3.24) also reveals another important thing. As discussed, a matrix applying to a vector is also a transformation of the vector. Therefore, the left-hand side can also be regarded as a new vector $|V\rangle_{new}$ after a certain operation *in the old basis*. What is this operation? As shown in Fig. 3.1, if the new basis can be obtained by rotating the old basis clockwise by an angle, θ, the operation is equivalent to a counterclockwise rotation of the vector by an angle, θ, in the old basis. In the figure, it can be seen that the representation of vector $|V\rangle$ in the new basis $|0'\rangle / |1'\rangle$ is the same as the representation of vector $|V\rangle_{new}$ in the old basis $|0\rangle / |1\rangle$. In general,

when we represent a vector in a new basis formed by a transformation $\boldsymbol{U^{-1}}$ of the old basis, it is the same as transforming the vector in the old basis by its inverse, i.e., $\boldsymbol{U}$.

Example 3.4 For the problem in Fig. 2.2, construct the transformation matrix to convert the representation of $|V_1\rangle$ in the old $|0\rangle / |1\rangle$ basis to the new $|+\rangle / |-\rangle$ basis.

Firstly, we recognize that the old basis has basis vectors $|0\rangle$ and $|1\rangle$. The new basis has basis vectors $\left|0'\right\rangle = |+\rangle$ and $\left|1'\right\rangle = |-\rangle$. Therefore, the transformation matrix is

$$\begin{aligned}\boldsymbol{U} &= \begin{pmatrix}\langle 0'|0\rangle & \langle 0'|1\rangle \\ \langle 1'|0\rangle & \langle 1'|1\rangle\end{pmatrix} = \begin{pmatrix}\langle +|0\rangle & \langle +|1\rangle \\ \langle -|0\rangle & \langle -|1\rangle\end{pmatrix}, \\ &= \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 \\ 1 & -1\end{pmatrix}. \end{aligned} \tag{3.25}$$

Now, if we apply $\boldsymbol{U}$ to $|V_1\rangle = |0\rangle$ as given in the first line of Eq. (2.15), we get

$$\boldsymbol{U}\,|V_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 \\ 1 & -1\end{pmatrix}\begin{pmatrix}1 \\ 0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 \\ 1\end{pmatrix}, \tag{3.26}$$

which has the same coefficients as those in the last line of Eq. (2.15). As discussed, $\boldsymbol{U}\,|V_1\rangle$ is also the vector formed after rotating $|V_1\rangle$ counterclockwise by 45° in the old basis. ■

If two vectors $|g\rangle$ and $|f\rangle$ are represented in a new basis, we expect that their inner product will not change. Since representing the vectors in a new basis is equivalent to transforming the vectors in the old basis, i.e., $\left|g'\right\rangle = \boldsymbol{U}\,|g\rangle$ and $\left|f'\right\rangle = \boldsymbol{U}\,|f\rangle$, then it means that *the transformation matrix must be unitary* in order to preserve their inner product (See Eq. (3.19)). Therefore, it also obeys Eq. (3.18), i.e., $\boldsymbol{U^\dagger U} = \boldsymbol{I}$.

Similar to vectors, when the basis is changed, matrices also need to be transformed accordingly. A matrix $\boldsymbol{M}$ is transformed to $\boldsymbol{M'}$ through

$$\boldsymbol{M'} = \boldsymbol{UMU^\dagger}. \tag{3.27}$$

This is also called the **similarity transformation**. It is not difficult to see that this makes sense. Assume $|w\rangle = \boldsymbol{M}\,|v\rangle$. This means that vector $|v\rangle$ is transformed by an operator $\boldsymbol{M}$ to another vector $|w\rangle$ and they are all in the same old basis. Now if we want to work on a new basis by applying the basis transformation matrix $\boldsymbol{U}$, we have,

$$\begin{aligned}\boldsymbol{U}\,|w\rangle &= \boldsymbol{U}(\boldsymbol{M}\,|v\rangle), \\ &= \boldsymbol{UMI}\,|v\rangle\,,\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{U}\boldsymbol{M}(\boldsymbol{U}^{\dagger}\boldsymbol{U})\,|v\rangle\,, \\
&= (\boldsymbol{U}\boldsymbol{M}\boldsymbol{U}^{\dagger})(\boldsymbol{U}\,|v\rangle), \qquad (3.28)
\end{aligned}
$$

which clearly shows that $|w\rangle$ in the new basis ($\boldsymbol{U}\,|w\rangle$) equals the operator in the new basis (($\boldsymbol{U}\boldsymbol{M}\boldsymbol{U}^{\dagger}$), Eq. (3.27)) multiplying $|v\rangle$ in the new basis ($\boldsymbol{U}\,|v\rangle$). This preserves the relationship between the vectors in the old basis, i.e., $|w\rangle = \boldsymbol{M}\,|v\rangle$.

3.4 Measurement of a Quantum State—Part 2

We have discussed some of the basics of measurement in Sect. 2.3.3. When a quantum state is measured, the state will collapse to one of the basis states. For example, if we perform a spin measurement on a spin qubit, the quantum state will collapse to either spin-up, $|\uparrow\rangle$, or spin-down, $|\downarrow\rangle$. Experimentally, we will also obtain a real number in the measurement (e.g., $\frac{1}{2}$ or $-\frac{1}{2}$).

In general, depending on what we are measuring, the basis states it will collapse to and the real values measured are the eigenvectors and eigenvalues, respectively, of a Hermitian matrix, $\boldsymbol{A}$. For example, if we are performing a spin measurement of an electron, this measurement corresponds to the Hermitian matrix, $\frac{1}{2}\boldsymbol{\sigma}_z = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$, which has eigenvalues of $\frac{1}{2}$ and $-\frac{1}{2}$. It should be noted that the corresponding Hermitian matrix is **NOT an operator to perform the measurement**. It is only that its eigenvectors are the states it will collapse to and its eigenvalues are the numbers being measured experimentally.

It should also be clear to the readers why the corresponding operators must be Hermitian. This is because the Hermitian matrix has real eigenvalues which are what will be measured experimentally.

We mentioned that the probability of a state $|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ collapsing to one of the basis states (which are the eigenstates of the corresponding Hermitian matrix) is the square of the modulus of the corresponding coefficient. Here, we will give a more versatile definition of the probability, $Prob(|i\rangle)$, it will collapse to basis state $|i\rangle$. That is,

$$
Prob(|i\rangle) = \langle\Psi|\,\boldsymbol{P}_{|i\rangle}\,|\Psi\rangle\,, \qquad (3.29)
$$

where the projection operator to $|i\rangle$ is used.

Example 3.5 Derive Eq. (2.21) using Eq. (3.29).

$$
\begin{aligned}
Prob(|0\rangle) &= \langle\Psi|\,P_{|0\rangle}\,|\Psi\rangle\,, \\
&= (\alpha^*\,\langle 0| + \beta^*\,\langle 1|)(|0\rangle\,\langle 0|)(\alpha\,|0\rangle + \beta\,|1\rangle), \\
&= \alpha^*\,\langle 0|0\rangle\,\langle 0|\,\alpha\,|0\rangle\,,
\end{aligned}
$$

$$= \alpha^* \alpha = |\alpha|^2, \tag{3.30}$$

where from line 2 to line 3, we have used the fact that $\langle 0|1\rangle = 0$. ■

3.4.1 Expectation Value in a Measurement

If $\boldsymbol{A}$ is the Hermitian matrix corresponding to a measurement and has eigenvectors $|0\rangle$ and $|1\rangle$, then the **expectation value** or the average value obtained by performing the measurement on many identically prepared state $|\Psi\rangle$ is the sum of the eigenvalues (λ_0, λ_1) of each eigenvector weighted by the probability of the eigenvector to which $|\Psi\rangle$ will collapse. Therefore, the expectation value of $\boldsymbol{A}$ (or the average measured value) for the given state $|\Psi\rangle$ is

$$\begin{aligned}
\langle \boldsymbol{A} \rangle &= Prob(|0\rangle)\lambda_0 + Prob(|1\rangle)\lambda_1, \\
&= \langle \Psi | \, P_{|0\rangle} \, |\Psi\rangle \, \lambda_0 + \langle \Psi | \, P_{|1\rangle} \, |\Psi\rangle \, \lambda_1, \\
&= \langle \Psi | 0 \rangle \, \langle 0 | \Psi \rangle \, \lambda_0 + \langle \Psi | 1 \rangle \, \langle 1 | \Psi \rangle \, \lambda_1, \\
&= \langle \Psi | \, (|0\rangle \, \langle 0| \, \lambda_0 + |1\rangle \, \langle 1| \, \lambda_1) \, |\Psi\rangle \, , \\
&= \langle \Psi | \, \boldsymbol{A} \, |\Psi\rangle \, .
\end{aligned} \tag{3.31}$$

In the last line, we used the fact that working in $\boldsymbol{A}$'s eigenbasis, $\boldsymbol{A}$ is a diagonal matrix with the eigenvalues along the diagonal which is $\boldsymbol{A} = |0\rangle \, \langle 0| \, \lambda_0 + |1\rangle \, \langle 1| \, \lambda_1$ (see Eqs. (3.8) and (3.9)).

3.5 Tensor Product of Matrices

In Sect. 2.4, we discussed how to construct a larger space by combining smaller spaces using the **tensor product**. The state/vector of the combined system can be described by the tensor product of the states/vectors of the smaller systems (Eq. (2.28)). Note that it can also be a linear combination of the tensor products if they are **entangled** which will be discussed later. We also need to create an operator for the combined system so that it is equivalent to the individual operators in the subsystems. For example, if $\boldsymbol{M_1}$ is applied to $|\psi_1\rangle$ and $\boldsymbol{M_2}$ is applied to $|\psi_2\rangle$, what is the equivalent operator $\boldsymbol{M}$ applied to state of the combined system, i.e., $|\psi_1\rangle \otimes |\psi_2\rangle$?

We construct $\boldsymbol{M}$ using a tensor product of $\boldsymbol{M_1}$ and $\boldsymbol{M_2}$,

$$\boldsymbol{M} = \boldsymbol{M_1} \otimes \boldsymbol{M_2}. \tag{3.32}$$

As a result, we have

$$\begin{aligned} \boldsymbol{M}\,|\psi\rangle &= \boldsymbol{M}(|\psi_1\rangle \otimes |\psi_2\rangle), \\ &= (\boldsymbol{M_1} \otimes \boldsymbol{M_2})(|\psi_1\rangle \otimes |\psi_2\rangle), \\ &= (\boldsymbol{M_1}\,|\psi_1\rangle) \otimes (\boldsymbol{M_2}\,|\psi_2\rangle). \end{aligned} \tag{3.33}$$

Note that the operator in each subsystem only applies to the state in that system. For example, a magnetic pulse to rotate the spin state of electron 1 ($\boldsymbol{M_1}$) is only physically applied to electron 1 and should not have an effect on electron 2. If it has, this is already an operator in the combined system.

When the operators are expressed in their matrix form, we follow the approach in Eq. (2.29) to perform the tensor product.

Example 3.6 If $\boldsymbol{M_1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\boldsymbol{M_2} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, find $\boldsymbol{M} = \boldsymbol{M_1} \otimes \boldsymbol{M_2}$.

$$\begin{aligned} \boldsymbol{M} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \\ &= \begin{pmatrix} a\begin{pmatrix} e & f \\ g & h \end{pmatrix} & b\begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ c\begin{pmatrix} e & f \\ g & h \end{pmatrix} & d\begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{pmatrix}, \\ &= \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}. \end{aligned} \tag{3.34}$$

■

3.6 Summary

In this chapter, we have reviewed some fundamental concepts of matrix. A Hermitian matrix has real eigenvalues. Therefore, all measurements must be corresponding to a Hermitian matrix. However, it is emphasized that the Hermitian matrix is not an operator that results in a measurement. It is only that its eigenvectors are the collapsed states after a measurement and its eigenvalues are the experimentally measured values. A unitary matrix preserves the inner products of vectors and preserves the vector norms. Therefore, the transformation matrix for basis changing must be unitary. We also learn how to create the operators of a combined system using a tensor product of the operators in the subsystems. Now, we have reviewed

most of the essential basic linear algebra and we can start studying the physics of quantum computers, namely, the Schrödinger equation in the next chapter.

Problems

3.1 Dual Correspondence

Prove Eq. (3.6) by using Eq. (2.6).

3.2 Adjoint Matrix

Find the adjoint matrix of $\begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}$. Is it Hermitian? Is it unitary?

3.3 Transformation

How is a general vector $|\Psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ transformed in the example in Fig. 2.2?

3.4 Tensor Product

Transform $|0\rangle_1 \otimes |0\rangle_2$ using the matrices in Eq. (3.34). Firstly, transform each qubit individually in its own space and then find the combined vector using a tensor product. Secondly, transform $|0\rangle_1 \otimes |0\rangle_2$ in the combined space using the corresponding matrix in the combined space. Show that both methods give the same result.

3.5 Diagonal Matrix

Show that this is a diagonal matrix by performing appropriate substitutions: $\boldsymbol{A} = |0\rangle \langle 0| \lambda_0 + |1\rangle \langle 1| \lambda_1$. See also Eq. (3.31).

Reference

1. Hiu-Yung Wong. *Introduction to Quantum Computing*. Springer, 2024.

Chapter 4
Schrödinger Equation and Quantum Gates

4.1 Introduction

To realize an operator or a quantum gate, we need to set up the hardware so that it has the appropriate energy landscape, which is called the Hamiltonian. A quantum state will then evolve by following the Schrödinger equation of the given Hamiltonian. In this chapter, we will first study how to solve the Schrödinger equation using matrix mechanics with both diagonal and non-diagonal Hamiltonians. Then we will discuss how a quantum gate can be generated for a given Hamiltonian. We will then review a few important 1-qubit quantum gates. We will also discuss the CNOT gate, which is a 2-qubit entanglement gate, and demonstrate how to use it to create an entanglement state by combining it with other 1-qubit gates.

4.1.1 Learning Outcomes

Understand the meaning of the Schrödinger equation; be able to solve the Schrödinger equation in matrix form for different types of Hamiltonians; be familiar with the basic gates and entanglement creation.

4.1.2 Teaching Videos

- Search for Ch4 in this playlist
  - https://tinyurl.com/3yhze3jn

H. Y. Wong, *Quantum Computing Architecture and Hardware for Engineers*,
https://doi.org/10.1007/978-3-031-78219-0_4

- Other videos
  - https://youtu.be/wyenXYGu51o
  - https://youtu.be/DvJPM3ACkNw
  - https://youtu.be/Wrmigi645J4
  - https://youtu.be/tKx-JZg0qYk

4.2 Schrödinger Equation

The **Schrödinger equation** is the *governing equation* in quantum mechanics. It is difficult to solve. However, it is relatively easy if it is applied to a 1-qubit system, which is the case in most parts of this book. The Schrödinger equation is given as

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \boldsymbol{H} |\psi\rangle , \tag{4.1}$$

where i is the imaginary number, $\sqrt{-1}$, t is time, and $\hbar$ is the **reduced Planck constant**. $\hbar = \frac{h}{2\pi}$ with the **Planck constant**, $h = 6.626 \times 10^{-34} J \cdot s$. $\boldsymbol{H}$ is the **Hamiltonian** *of the system that we are investigating*. The Hamiltonian is the total energy of the system, which is the sum of the potential and kinetic energies. We will discuss it more in depth in Chap. 13. Here, we assume that $\boldsymbol{H}$ is given. Also, we write $\boldsymbol{H}$ in boldface because we treat it as a matrix here. In the following chapters, we will start writing it as an operator after we have learned the necessary knowledge. $|\psi\rangle$ is the state of the system.

Let us first descriptively understand what the Schrödinger equation tries to tell us. It says that the rate of change of the state ($\frac{\partial |\psi\rangle}{\partial t}$) is proportional to (scaled by $i\hbar$) the Hamiltonian multiplied by the state ($\boldsymbol{H} |\psi\rangle$).

Recalling that we represent a state as a vector, the Hamiltonian must be a matrix. Writing and solving the Schrödinger equation in this way is called the **matrix mechanics** as proposed by Heisenberg in contrast to Schrödinger's wave formulation. For finite (discrete) Hilbert spaces such as those of qubit systems, matrix mechanics is often more convenient.

Let us now consider a 1-qubit system. A single qubit is a 2D Hilbert space with complex scalars with two basis states, $|0\rangle$ and $|1\rangle$. Again, $|0\rangle$ and $|1\rangle$ are just the labels of the basis states and it does not matter what the underlying physics is. A general state in the system, $|\psi\rangle$, can be represented as a linear combination of the basis states

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{4.2}$$

where α and β are complex scalars. If α and β are determined, then $|\psi\rangle$ is determined. Therefore, solving the Schrödinger equation for $|\psi\rangle$ in the 1-qubit system is equivalent to finding α and β.

Since it is a 2D space, the matrix must be 2×2 in size. We assume the Hamiltonian to be

$$\boldsymbol{H} = \begin{pmatrix} H_{00} & H_{01} \\ H_{10} & H_{11} \end{pmatrix}, \tag{4.3}$$

where H_{00}, H_{01}, H_{10}, and H_{11} are complex numbers. Then Eq. (4.1) becomes

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} H_{00} & H_{01} \\ H_{10} & H_{11} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{4.4}$$

To find α and β, we perform scalar multiplication on the left-hand side and matrix multiplication on the right-hand side,

$$\begin{aligned} \begin{pmatrix} i\hbar \frac{\partial \alpha}{\partial t} \\ i\hbar \frac{\partial \beta}{\partial t} \end{pmatrix} &= \begin{pmatrix} H_{00} & H_{01} \\ H_{10} & H_{11} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \\ &= \begin{pmatrix} H_{00}\alpha + H_{01}\beta \\ H_{10}\alpha + H_{11}\beta \end{pmatrix}. \end{aligned} \tag{4.5}$$

The vectors on the left and right are the same and so do their coefficients. Therefore, we obtain two simultaneous *differential equations*,

$$i\hbar \frac{\partial \alpha}{\partial t} = H_{00}\alpha + H_{01}\beta, \tag{4.6}$$

$$i\hbar \frac{\partial \beta}{\partial t} = H_{10}\alpha + H_{11}\beta. \tag{4.7}$$

To solve Eqs. (4.6) and (4.7), we need to solve a second-order differential equation by substituting one into another. We will study two cases to understand how to solve them in general.

4.3 Solving Schrödinger Equation

In general, a **matrix differential equation** with the following form,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \boldsymbol{H} |\psi\rangle ; \quad |\psi(t=0)\rangle = |\psi_0\rangle , \tag{4.8}$$

has a general solution of

$$|\psi(t)\rangle = e^{\frac{\boldsymbol{H}}{i\hbar}t} |\psi_0\rangle = e^{-i\frac{\boldsymbol{H}}{\hbar}t} |\psi_0\rangle , \tag{4.9}$$

when $\boldsymbol{H}$ is a *constant matrix* (independent of time). If it is time-dependent, more sophisticated equations are needed and readers can refer to Chapter 2 in [2]. Therefore, if we know how to perform **matrix exponential**, we can obtain the solution, too.

4.3.1 Diagonal Hamiltonian

If the given Hamiltonian is diagonalized, then H_{01} and H_{10} are zero. A Hamiltonian (or, in general, an operator) is diagonal if the basis being used is the eigenbasis of the Hamiltonian (see Sect. 3.3.1). The equations to be solved become

$$i\hbar \frac{\partial \alpha}{\partial t} = H_{00}\alpha, \tag{4.10}$$

$$i\hbar \frac{\partial \beta}{\partial t} = H_{11}\beta. \tag{4.11}$$

We can see that now α and β are **decoupled**, and each equation contains only one variable and can be solved independently. Note that α and β refer to the amount of each basis state ($|0\rangle$ and $|1\rangle$) that $|\psi\rangle$ has. Therefore, **non-zero off-diagonal elements** *enable the coupling between different basis states*. For example, even if $\beta = 0$ at $t = 0$, eventually, β will become non-zero if there are non-zero off-diagonal elements which enable the coupling.

The solutions to the equations are

$$\alpha(t) = \alpha_0 e^{-i\frac{H_{00}}{\hbar}t}, \tag{4.12}$$

$$\beta(t) = \beta_0 e^{-i\frac{H_{11}}{\hbar}t}, \tag{4.13}$$

where α_0 and β_0 are constants and they are the initial values of $\alpha(t)$ and $\beta(t)$ at $t = 0$. One may substitute Eqs. (4.12) and (4.13) into Eqs. (4.10) and (4.11), respectively, to show that they are indeed the solutions.

Therefore, the state (vector) of the 1-qubit system changes as a function of time when it has a diagonal Hamiltonian $\boldsymbol{H} = \begin{pmatrix} H_{00} & 0 \\ 0 & H_{11} \end{pmatrix}$ as

$$|\psi\rangle = \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} \alpha_0 e^{-i\frac{H_{00}}{\hbar}t} \\ \beta_0 e^{-i\frac{H_{11}}{\hbar}t} \end{pmatrix}. \tag{4.14}$$

We may also check this by using Eq. (4.9). When the Hamiltonian is *diagonal*, we can exponentiate it easily by only exponentiating the diagonal elements (the proof

will be given in Example 4.1). That is,

$$\begin{aligned} e^{-i\frac{\boldsymbol{H}}{\hbar}t} &= e^{-i\frac{\begin{pmatrix} H_{00} & 0 \\ 0 & H_{11} \end{pmatrix}}{\hbar}t}, \\ &= \begin{pmatrix} e^{-i\frac{H_{00}}{\hbar}t} & 0 \\ 0 & e^{-i\frac{H_{11}}{\hbar}t} \end{pmatrix}. \end{aligned} \tag{4.15}$$

Therefore, Eq. (4.9) becomes

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\frac{\boldsymbol{H}}{\hbar}t}\,|\psi_0\rangle\,, \\ \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} &= \begin{pmatrix} e^{-i\frac{H_{00}}{\hbar}t} & 0 \\ 0 & e^{-i\frac{H_{11}}{\hbar}t} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}, \\ \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} &= \begin{pmatrix} \alpha_0 e^{-i\frac{H_{00}}{\hbar}t} \\ \beta_0 e^{-i\frac{H_{11}}{\hbar}t} \end{pmatrix}, \end{aligned} \tag{4.16}$$

which is the same as the solution in Eq. (4.14).

Example 4.1 Prove Eq. (4.15).

We prove this by using the Taylor expansion of $e^{-i\frac{\boldsymbol{H}}{\hbar}t}$ and the definition of the zero exponent of a matrix, $\boldsymbol{H}$,

$$\boldsymbol{H}^0 = \boldsymbol{I}. \tag{4.17}$$

The Taylor series of $e^{-i\frac{\boldsymbol{H}}{\hbar}t}$ is

$$\begin{aligned} e^{-i\frac{\boldsymbol{H}}{\hbar}t} &= \sum_{k=0}^{\infty} \frac{(-i\frac{\boldsymbol{H}}{\hbar}t)^k}{k!}, \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} -i\frac{H_{00}}{\hbar}t & 0 \\ 0 & -i\frac{H_{11}}{\hbar}t \end{pmatrix}^k, \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} (-i\frac{H_{00}}{\hbar}t)^k & 0 \\ 0 & (-i\frac{H_{11}}{\hbar}t)^k \end{pmatrix}, \\ &= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}(-i\frac{H_{00}}{\hbar}t)^k & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{1}{k!}(-i\frac{H_{11}}{\hbar}t)^k \end{pmatrix}, \\ &= \begin{pmatrix} e^{-i\frac{H_{00}}{\hbar}t} & 0 \\ 0 & e^{-i\frac{H_{11}}{\hbar}t} \end{pmatrix}. \end{aligned} \tag{4.18}$$

where in line 2, Eq. (4.3) is used to substitute $\boldsymbol{H}$. In line 3, we use the fact that when a diagonal matrix multiplies itself, it is the same as each diagonal element multiplies itself. In line 4, we just use the definition of matrix summation, and in line 5, the Taylor series of number exponential is used. ■

4.3.2 Non-diagonal Hamiltonian

If the Hamiltonian is not diagonal (i.e., at least one of the off-diagonal elements is non-zero), then we cannot use Eq. (4.15). We need to solve the system of linear equations in Eqs. (4.6) and (4.7). This is tedious. However, if we already know the eigenvalues and eigenvectors of $\boldsymbol{H}$, we can work on its eigenbasis to find the solutions and then transform it back to the basis we are interested in. By the way, since $\boldsymbol{H}$ is the total energy of the system, its eigenvalues are also called the **eigenenergies**.

We had discussed how to construct a general transformation matrix in Sect. 3.3.5. Assume we are in an old basis with basis states $|0\rangle$ and $|1\rangle$. In this basis, $\boldsymbol{H}$ is *not* diagonal. We can work in the eigenbasis of $\boldsymbol{H}$ (the new basis with basis vectors $\left|0'\right\rangle$ and $\left|1'\right\rangle$) by creating a transformation matrix, $\boldsymbol{U}$, based on Eq. (3.23)

$$\boldsymbol{U} = \begin{pmatrix} \langle 0'|0\rangle & \langle 0'|1\rangle \\ \langle 1'|0\rangle & \langle 1'|1\rangle \end{pmatrix}. \tag{4.19}$$

Then Eq. (4.1) becomes

$$\begin{aligned} \boldsymbol{U} i\hbar \frac{\partial}{\partial t} |\psi\rangle &= \boldsymbol{U}\boldsymbol{H}\boldsymbol{I} |\psi\rangle , \\ i\hbar \frac{\partial}{\partial t} \boldsymbol{U} |\psi\rangle &= \boldsymbol{U}\boldsymbol{H}\boldsymbol{U}^{\dagger}\boldsymbol{U} |\psi\rangle . \end{aligned} \tag{4.20}$$

This is like how we derived Eq. (3.28) by applying $\boldsymbol{U}$, which is time independent, from the left and using the identity, $\boldsymbol{U}^{\dagger}\boldsymbol{U} = \boldsymbol{I}$. More specifically, we can set $\left|\psi'\right\rangle = \boldsymbol{U} |\psi\rangle$ and $\boldsymbol{H'} = \boldsymbol{U}\boldsymbol{H}\boldsymbol{U}^{\dagger}$ which is *diagonal* and we can solve

$$i\hbar \frac{\partial}{\partial t} \left|\psi'\right\rangle = \boldsymbol{H'} \left|\psi'\right\rangle , \tag{4.21}$$

as how we did in the diagonal Hamiltonian case in Eq. (4.16). After obtaining, $\left|\psi'\right\rangle$, we can get $|\psi\rangle$ by using

$$|\psi\rangle = \boldsymbol{U}^{\dagger} \left|\psi'\right\rangle . \tag{4.22}$$

Of course, the difficulty is to find the eigenvectors of $\boldsymbol{H}$ which is computationally intensive when the matrix is large.