

# CSCI3350 Introduction to Quantum Computing (2026 Spring)

Xiao Liang

 <https://xiao-liang.github.io>

Department of Computer Science and Engineering  
The Chinese University of Hong Kong

(Updated on January 13, 2026)

**Disclaimer:** These lecture notes originate from the Spring 2026 offering of the course. They have been lightly edited for accuracy and completeness, and may still contain typographical or other errors.

## Contents

<b>0 Physics and Computing: from Newton to Shrödinger</b>	<b>1</b>
0.1 Era I: Classical Mechanics and Mechanical Computing . . . . .	1
0.2 Era II: Electromagnetism and Electronic Computing . . . . .	3
0.2.1 Mechanical vs. Electronic Computing . . . . .	4
0.3 Era III: Quantum Mechanics and Quantum Computing . . . . .	4
0.3.1 Schrödinger's Equation . . . . .	4
0.3.2 An Inaccurate-yet-Helpful Analogy . . . . .	5
0.4 Our Focus: Quantum Computing . . . . .	6
<b>1 Preliminaries: Linear Algebra over Complex Vector Spaces</b>	<b>8</b>
1.1 Vector Spaces: A Review . . . . .	8
1.1.1 Formal Definition of Vector Space . . . . .	9
1.1.2 Reviewing Basic Concepts in Vector Spaces . . . . .	10
1.2 Complex Vector Spaces . . . . .	14
1.2.1 Reviewing Complex Number Arithmetic . . . . .	14
1.3 Inner Products: Adding Geometry to Vector Spaces . . . . .	16
1.3.1 Starting Point: the Euclidean Space $\mathbb{R}^3$ . . . . .	17
1.3.2 Complex Inner-Product Space $\mathbb{C}^n$ . . . . .	19
1.4 Hilbert Space—the Playground for Quantum Computing . . . . .	21
1.5 *Bonus Section: Why Complex Numbers for Quantum Mechanics? . . . . .	22
1.5.1 Why Complex Numbers? . . . . .	22
1.5.2 Complex Numbers in Simple Harmonic Oscillation . . . . .	23
1.5.3 Complex Numbers in Electromagnetic Wave Equations . . . . .	24
<b>2 Postulate 1: State Space</b>	<b>27</b>
2.1 Two-Dimensional Quantum Systems—A Qubit . . . . .	27
2.1.1 Example 1: Spin- $\frac{1}{2}$ in a Magnetic Field . . . . .	28
2.1.2 Example 2: Two-Level Atom (or “Artificial Atom”) . . . . .	29
2.2 The Math for a Single Qubit . . . . .	30
2.3 Quantum Systems of Higher Dimensions . . . . .	31

2.4	Dirac Notation . . . . .	32
2.4.1	Worked Examples: Calculations in Dirac Notation . . . . .	34
<b>3</b>	<b>Postulate (2/4): State Evolution</b>	<b>39</b>
3.1	Commonly Used Unitaries . . . . .	39
<b>4</b>	<b>Postulate (3/4): Measurement (Born Rule)</b>	<b>42</b>
4.1	What a Quantum Measurement Means (Physical Intuition) . . . . .	42
4.2	Postulate 3: the Formal Statement . . . . .	44
4.3	Example: Measuring a Single Qubit . . . . .	44
4.3.1	Case 1: measurement in the computational ( $Z$ ) basis . . . . .	45
4.3.2	Case 2: measurement in the Hadamard ( $X$ ) basis . . . . .	46
	<b>Bibliography</b>	<b>48</b>
	<b>A From matrices to measurable properties: why <math>Z</math> represents an observable</b>	<b>49</b>

# Chapter 0

## Physics and Computing: from Newton to Shrödinger

Nothing from [Chapter 0](#) will be tested on quizzes or exams!

Our computational power is bounded—and enabled—by our understanding of physics. Different physical laws suggest different ways to represent and process information. As we move from gears and pulleys to transistors and integrated circuits and, finally, to quantum systems, we get new primitives for encoding and manipulating information. In this chapter, we will move conceptually:

- from classical mechanics (Newton) and mechanical computation,
- to electromagnetism (Maxwell) and electronic computation,
- to quantum mechanics (Schrödinger) and quantum computation.

### 0.1 Era I: Classical Mechanics and Mechanical Computing

Classical mechanics is governed by Newton's laws, which describe how forces influence motion. The foundational equation is Newton's second law:

$$\mathbf{F} = m \mathbf{a} = \frac{d\mathbf{p}}{dt}. \quad (1)$$

Here:

- $\mathbf{F}$  is the net force vector acting on an object (magnitude and direction).
- $m$  is the mass (a scalar), measuring how strongly the object resists changes in motion.
- $\mathbf{a} = \frac{d\mathbf{v}}{dt}$  is the acceleration vector, the time derivative of velocity.
- $\mathbf{p} = m \mathbf{v}$  is the momentum vector (for constant mass  $m$  and velocity  $\mathbf{v}$ ).
- $\frac{d\mathbf{p}}{dt}$  is the time derivative of momentum, equal to the net force.

Given forces, masses, and initial conditions, we can predict trajectories. The exemplary systems (see [Figure 1](#)) that we have solved in secondary school speaks for this idea.

This predictive power underpins early computational devices built on mechanical principles: if we can engineer interacting motions (gears, levers, pulleys), we can compute. The lesson is simple: with precise control over mechanical components, we can implement arithmetic and logic through motion and force.

We now present two examples of mechanical computing devices: abacus and gear calculators.

#### Mechanical Computing Devices: Abacus



Figure 1: Exemplary Systems Solvable by Classical Mechanics

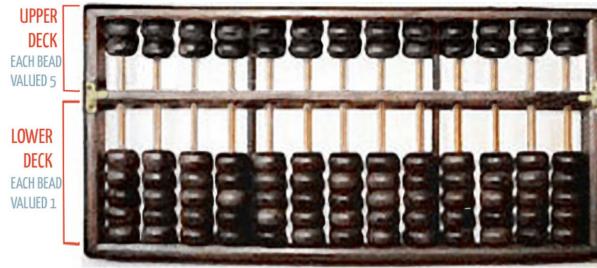


Figure 2: Abacus

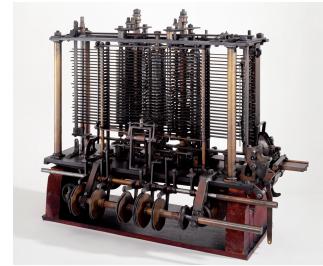
An abacus represents numbers using beads on rods and implements arithmetic by moving beads according to rules. It is simple yet powerful for certain calculations.

### Mechanical Computing Devices: Gear Calculators

Gear-based calculators (e.g., Pascal's Pascaline; Babbage's designs) encode numbers as gear positions; addition and subtraction follow from rotations and carries propagating through gear trains.



(a) Pascaline



(b) Babbage's analytical engine (concept)

Mechanical computation excels when problems can be mapped to a small number of coupled motions, but it struggles with tasks requiring deep memory, branching logic, or billions of fast operations.

## 0.2 Era II: Electromagnetism and Electronic Computing

A second revolution comes with Maxwell's unification of electricity and magnetism. The governing equations are:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} && \text{(Gauss's law)} \\ \nabla \cdot \mathbf{B} &= 0 && \text{(Gauss's law for magnetism)} \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{(Faraday's law)} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} && \text{(Ampère–Maxwell law),}\end{aligned}$$

where  $\mathbf{E}$  and  $\mathbf{B}$  are the electric and magnetic fields,  $\rho$  is charge density,  $\mathbf{J}$  is current density,  $\epsilon_0$  vacuum permittivity, and  $\mu_0$  vacuum permeability. These laws predict electromagnetic waves and enable circuitry, antennas, and information transmission (see [Figure 4](#)).

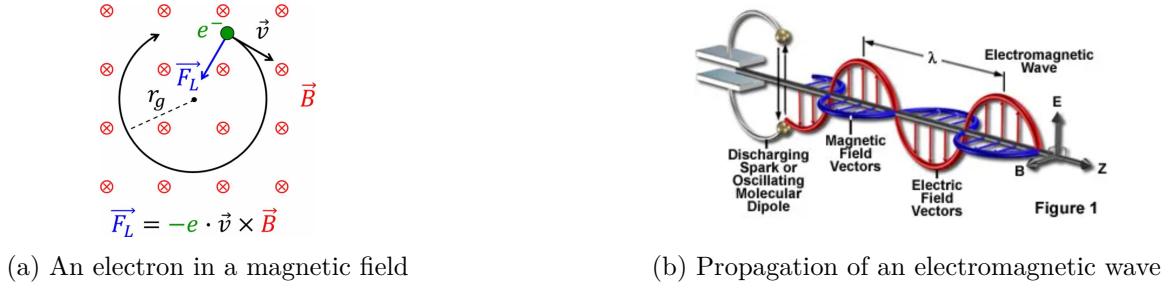


Figure 4: Illustrative Systems in Electromagnetism

If we can harness fields and waves (wires, antennas, circuits), we can compute with electrons and light.

### Electronic Computing Devices

Modern devices—from calculators and desktops to smartphones and cloud servers—encode information electronically.

Two common and crucial tasks for all these devices are *encoding* and *computing*:

- **Encoding**
  - Bits as high/low voltage levels.
  - Wires carry bits; memory stores them as charge states or magnetic domains.
- **Computing**
  - Logic gates (AND, OR, NOT, NAND, etc.) built from transistors implement Boolean functions.
  - Synchronous circuits orchestrated by clocks compose gates into processors.
  - Algorithms are sequences of logic operations on registers of bits.

### 0.2.1 Mechanical vs. Electronic Computing

Mechanical calculators are well-suited for a few coupled motions (gears, cams, analog integrators). They struggle when a task requires:

- exploring enormous combinatorial possibilities,
- processing very large datasets with high precision,
- running millions to billions of steps quickly,
- flexible memory and branching logic.

Electronics handle these via fast arithmetic, hierarchical memory, and parallelism. Examples include:

- high-dimensional PDEs and large-scale simulation,
- large-scale linear algebra and optimization,
- symbolic computation and automated reasoning,
- data-intensive analytics.

## 0.3 Era III: Quantum Mechanics and Quantum Computing

At microscopic scales, classical laws fail to explain experiments. Quantum mechanics provides new rules for evolution and measurement, altering how we encode and process information.

Concrete examples for microscopic-scale systems include:

- Atoms and molecules ( $\sim 0.1\text{--}1 \text{ nm}$ ): explains spectral lines, chemical bonds.
- Tunneling through thin barriers (few nm): exploited in flash memory and scanning tunneling microscopes.
- Quantum dots (1–10 nm): nanocrystals with size-tunable colors used in displays.

### 0.3.1 Schrödinger's Equation

Quantum systems are governed by the Schrödinger equation:<sup>1</sup>

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (2)$$

Here:

- $i := \sqrt{-1}$  is the imaginary unit. (Complex numbers are essential in quantum theory. We will return to this topic in [section 1.5](#).).
- $\hbar$  is Planck's reduced constant,  $\hbar \approx 1.055 \times 10^{-34} \text{ J} \cdot \text{s}$ .<sup>2</sup>

---

<sup>1</sup>What we present in [Equation \(2\)](#) is the simplest form of the (time-dependent) Schrödinger equation. There exist more advanced formulations that capture a broader class of quantum systems.

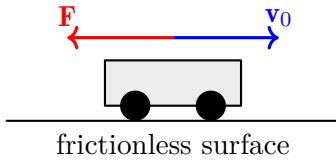
<sup>2</sup>Named after [Max Planck](#), Nobel Prize in Physics (1918).

- $|\psi(t)\rangle$  is the state vector of the system at time  $t$ .
- $\hat{H}$  is the so-called *Hamiltonian operator*. We will discuss it in more details soon.

We do not need a fully rigorous understanding of the Schrödinger equation—this is not a quantum mechanics course. But it helps to have an intuitive feel for it. We will build that intuition by drawing an analogy with classical mechanics.

### 0.3.2 An Inaccurate-yet-Helpful Analogy

Consider the cart on a frictionless surface with initial velocity  $\mathbf{v}_0$  and an applied force  $\mathbf{F}$  (shown in [Figure 5](#)). In classical mechanics, knowing  $\mathbf{F} = m\mathbf{a}$  and initial conditions lets us predict where the cart will be at time  $t$ .



[Figure 5: Cart on a frictionless surface](#)

By analogy, the Schrödinger equation plays the role of the “law of motion” in quantum mechanics:

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

Think of:

- $|\psi(t)\rangle$  as a *vector* describing the system’s state at time  $t$ ;
- $\hat{H}$  (the Hamiltonian) as a *rule-book* or *matrix* encoding how the system behaves (energies, forces, boundaries);

With the above understanding, the equation says: *Change in the system’s state over time is determined by the rule-book acting on the state!*

In more details:

- $|\psi(t)\rangle$ : Think it as “where/what the system is like right now.” It’s like a vector you update over time.
- $\hat{H}$ : Think it as “all the ingredients of the system”—mass, potential energy, constraints, fields, etc. It acts like a matrix that “pushes” the system’s state forward in time.
- $i\hbar \frac{\partial}{\partial t}$ : It reads as “how fast the state changes with time,” scaled by a constant  $i$  times  $\hbar$ .

**In Summary:** The following is all you need to know about Shrödinger’s equation:

- Pretend the quantum state is like a vector that tells us where the system could be. The Hamiltonian is a big matrix that encodes the rules of the world. The Schrödinger equation says the matrix pushes the vector forward in time. If we know the rule and where we started, we can predict what we’ll see later.

## Some caveats

The previous analogy is not accurate. For example,

- The “vector”  $|\psi\rangle$  is not a single, deterministic status of the system; it actually encodes *probabilities of many possible status*.
- $\hat{H}$  is *total energy operator* (kinetic + potential).

Indeed, there is an alternative formulation of classical mechanics, equivalent to Newton’s, called *Hamiltonian mechanics*. The Schrödinger equation often feels more natural to those familiar with this formalism. If you are interested, you can easily find online resources on the topic, e.g., [David Tong’s course](#).

But again, our CSCI3350 is not a quantum mechanics course. We do not need a fully rigorous interpretation of the equation. Instead, we will focus on its major implications for how we perform computation (i.e., *the Four Postulates* that we will discuss later).

We may return to the interpretation of Shrödinger’s equation when needed—for example, when discussing Hamiltonian complexity or applications of quantum computing to quantum simulation and quantum chemistry, where the meaning of the Hamiltonian  $\hat{H}$  becomes more relevant.

## 0.4 Our Focus: Quantum Computing

With the quantum background sketched above, our goal is to understand how Schrödinger evolution enables us to encode information and perform computation.

We will first utilize a few lectures to build up the basic knowledge.

Once we have the basics, we will explore the following aspects of quantum computing:

- **Classical vs. quantum power:** Are there problems quantum computers can solve that classical computers cannot? What features are uniquely quantum?
- **Limits:** Can quantum computers solve all problems we care about? Which problems remain hard even for quantum computers?
- **Applications and interfaces:** Quantum machine learning, quantum networks, quantum cryptography, quantum chemistry, and more.

## Course Style and Scope

**Theory-first, no programming:** The course style is similar to [CSCI3160 Design and Analysis of Algorithms](#), focusing on the algorithmic ideas, pseudocode, and theoretical analysis. It does not involve programming, real-world implementation, or software engineering considerations.

**Beyond algorithms:** We also explore the fundamental principles of quantum information — including concepts like superposition, entanglement, and measurement. The course may also cover several important topics that are not algorithmic in nature but are central to the field of quantum computing, such as quantum error correction, fault-tolerant computation, proofs of quantumness, and non-local games.

**What This Course Is Not:** This course is *not* intended for students who:

- want to learn how to build a quantum computer,
- want hands-on practice in a specific quantum programming language,
- expect a full quantum mechanics/physics course.

Note that this arrangement is not unique to our course of CSCI3350. It is a common practice for quantum computing courses offered by most CS departments. You can explore similar courses at:

- [COMS4281](#) at Columbia University
- [CS498QC](#) at UIUC
- [15-859BB](#) at CMU
- [CS358H](#) at UT Austin

# Chapter 1

## Preliminaries: Linear Algebra over Complex Vector Spaces

Quantum computing is formulated in the language of linear algebra over complex vector spaces. This mathematical framework serves as the foundation for quantum states, measurements, and unitary evolutions. To prepare for these ideas, we first establish the necessary linear-algebraic basics.

We begin by recalling linear algebra on real vector spaces—the material typically covered in an undergraduate course. This review anchors the discussion in concepts you are already familiar with: span, linear independence, bases, dimension, linear transformations, and their matrix representations.

Next, we review the arithmetic of complex numbers. We will cover their algebraic form, geometric interpretation on the complex plane, conjugation, modulus and argument, Euler’s formula, and basic operations and identities that recur throughout quantum computing.

Finally, we extend the familiar setting of real vector spaces to complex vector spaces. We will introduce complex vector spaces and inner products, emphasize conjugate-linearity in inner products, and discuss norms, orthogonality, and unitary operators. With these tools in place, we will be ready to develop the quantum formalism.

### 1.1 Vector Spaces: A Review

Start from what we are familiar with: the real vector space  $\mathbb{R}^3$ . This vector space consists of:

- An underlying number field  $\mathbb{R}$ . Elements in  $\mathbb{R}$  are called *scalars*.
- A set  $\mathbb{R}^3 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\}$ . Elements in  $\mathbb{R}^3$  are called *vectors*.
- An *addition operation* “+” between elements in  $\mathbb{R}^3$ :  $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$ , where  $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$  and  $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$ . Note that the result of the addition is again an element in  $\mathbb{R}^3$ .
- An *multiplication operation* “.” between the field  $\mathbb{R}$  and the set  $\mathbb{R}^3$ :  $c \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c \cdot x_1 \\ c \cdot x_2 \\ c \cdot x_3 \end{bmatrix}$ , where  $c \in \mathbb{R}$  and  $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ . Note that the result of the multiplication is again an element in  $\mathbb{R}^3$ .

### Teaching Suggestions

During the lecture, show an example with  $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ,  $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ , and  $c = 3$ .

#### 1.1.1 Formal Definition of Vector Space

The above example illustrates all the important components of a vector space. But the real definition is more general. We now recall the full definition in [Definition 1.1.1](#).

**Definition 1.1.1** (Vector Space). *A vector space over a field  $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ <sup>1</sup> is a non-empty set  $V$  together with two binary operations*

$$\cdot : \mathbb{F} \times V \rightarrow V \quad \text{and} \quad + : V \times V \rightarrow V$$

*which satisfy the eight axioms listed below. In this context, the elements of  $V$  are commonly called vectors, the elements of  $\mathbb{F}$  are called scalars, the operations “.” is called scalar multiplication, and the operations “+” is called vector addition or simply addition.*

*To have a vector space, the following eight axioms must be satisfied for every  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ , and  $a, b \in \mathbb{F}$ :*

1. **Associativity of vector addition:**  $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ .
2. **Commutativity of vector addition:**  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ .
3. **Identity element of vector addition:** *There exists an element  $\mathbf{0} \in V$ , called the zero vector, such that  $\mathbf{v} + \mathbf{0} = \mathbf{v}$  for all  $\mathbf{v} \in V$ .*
4. **Inverse elements of vector addition:** *For every  $\mathbf{v} \in V$ , there exists an element  $-\mathbf{v} \in V$ , called the additive inverse of  $\mathbf{v}$ , such that  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ .*
5. **Compatibility of scalar multiplication with field multiplication:**

$$a \cdot (b \cdot \mathbf{v}) = (ab) \cdot \mathbf{v},$$

*where the “ $ab$ ” in the right-hand side of the equation means the field multiplication of  $\mathbb{F}$  between its elements  $a$  and  $b$ .*

6. **Identity element of scalar multiplication:**  $1 \cdot \mathbf{v} = \mathbf{v}$ , where  $1$  denotes the multiplicative identity in  $\mathbb{F}$ .
7. **Distributivity of scalar multiplication with respect to vector addition:**  $a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v}$ .
8. **Distributivity of scalar multiplication with respect to field addition:**  $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$ .

◇

---

<sup>1</sup>Actually, the concept of “field” can be more general. However, a formal treatment is out of the scope of this course. For this course, we will only use two fields: the real number field  $\mathbb{R}$  and the complex number field  $\mathbb{C}$ .

**More Examples of Vector Spaces.** Anything satisfies the requirements in [Definition 1.1.1](#) would be a vector space. Besides the familiar  $\mathbb{R}^n$ , let us see some more examples:

- **Space  $C[a, b]$ :** Each point of this space is a continuous real-valued function on (the real interval)  $[a, b]$ . The set of all these functions forms a real vector space with the algebraic operations defined in the usual way:

$$(x + y)(t) = x(t) + y(t),$$

$$(ax)(t) = ax(t) \quad (a \in \mathbb{R}).$$

Namely, this an instantiation of [Definition 1.1.1](#) with  $V = C[a, b]$  and  $\mathbb{F} = \mathbb{R}$ .

- **Space  $P(\mathbb{R})$ :** All polynomials with coefficients from  $\mathbb{R}$  form a vector space, under the standard polynomial arithmetic. Namely, this an instantiation of [Definition 1.1.1](#) with  $V = P(\mathbb{R})$  and  $\mathbb{F} = \mathbb{R}$ .

It is not hard to verify that these two spaces satisfy all the requirements stated in [Definition 1.1.1](#).

Note also that as long as we have a vector space—regardless of what it looks like—it already possesses all the concepts we will review in [Section 1.1.2](#).

### 1.1.2 Reviewing Basic Concepts in Vector Spaces

In a vector space, we study fundamental concepts such as span, linear independence, basis, dimension, and linear operators/transformations (represented by matrices). Let's begin by reviewing these ideas.

There are additional important concepts—such as eigenvalues and eigenvectors—which we will review later as needed.

#### Span

**Definition 1.1.2** (Linear combination). *Given vectors  $v_1, \dots, v_k \in V$  and scalars  $a_1, \dots, a_k \in \mathbb{F}$ , a linear combination of the  $v_i$  is  $a_1v_1 + \dots + a_kv_k$ .* ◇

**Definition 1.1.3** (Subspace). *For a non-empty subset  $S \subseteq V$ , given vector  $v, w \in S$  and scalars  $a, b \in \mathbb{F}$ , if  $av + bw \in S$ , then  $S$  is a subspace of  $V$ .* ◇

**Definition 1.1.4** (Span). *For a subset  $S \subseteq V$ , the span of  $S$  is*

$$\text{Span}(S) = \left\{ \sum_{i=1}^k a_i v_i : k \in \mathbb{N}, v_i \in S, a_i \in \mathbb{F} \right\}.$$

If  $S = \{v_1, \dots, v_n\}$  is finite we write  $\text{Span}\{v_1, \dots, v_n\}$ . ◇

**Proposition 1.1.1.**  *$\text{Span}(S)$  is a subspace of  $V$  and is the smallest subspace of  $V$  containing  $S$  (i.e., contained in every subspace that contains  $S$ ).*

**Example 1.1.1.** In  $\mathbb{R}^3$ ,  $\text{Span}\{(1, 0, 0), (0, 1, 0)\} = \{(x, y, 0) : x, y \in \mathbb{R}\}$ , the  $xy$ -plane.

#### Linear Independence

**Definition 1.1.5** (Linear independence). A list (or set) of vectors  $v_1, \dots, v_n \in V$  is linearly independent if

$$a_1v_1 + \dots + a_nv_n = 0 \Rightarrow a_1 = \dots = a_n = 0.$$

Otherwise they are linearly dependent.  $\diamond$

**Example 1.1.2.** In  $\mathbb{R}^3$ , the vectors  $(1, 0, 0), (0, 1, 0), (1, 1, 0)$  are dependent because  $(1, 1, 0) = (1, 0, 0) + (0, 1, 0)$ .

## Basis

**Definition 1.1.6** (Basis). A basis of  $V$  is a set of vectors  $\mathcal{B} = \{b_1, \dots, b_n\}$  such that

1.  $\mathcal{B}$  spans  $V$ , i.e.,  $\text{Span}(\mathcal{B}) = V$ , and
2.  $\mathcal{B}$  is linearly independent.

$\diamond$

**Theorem 1.1.1** (Uniqueness of coordinates). If  $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  is a basis of  $V$ , every  $\mathbf{v} \in V$  can be written uniquely as  $\mathbf{v} = \sum_{i=1}^n a_i \mathbf{b}_i$  with scalars  $a_i \in \mathbb{F}$ . The vector  $(a_1, \dots, a_n)$  is the coordinate vector of  $\mathbf{v}$  relative to  $\mathcal{B}$ .  $\diamond$

**Example 1.1.3.** The standard basis of  $\mathbb{R}^n$  is  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , where  $\mathbf{e}_i$  has a 1 in the  $i$ -th position and 0 elsewhere.

## Dimension

**Definition 1.1.7** (Dimension). If  $V$  has a finite basis, its dimension, denoted as  $\dim V$ , is the number of vectors in any basis. If no finite basis exists,  $V$  is infinite-dimensional.  $\diamond$

**Theorem 1.1.2** (Well-definedness). If  $V$  is finite-dimensional, all bases of  $V$  have the same cardinality. Hence  $\dim V$  is well-defined.  $\diamond$

**Proposition 1.1.2** (Dimension inequalities). Let  $V$  be finite-dimensional.

1. Any (linearly) independent set has size at most  $\dim V$ ; any spanning set has size at least  $\dim V$ .
2. If  $U$  is a subspace of  $V$ , then  $\dim U \leq \dim V$ , with equality iff  $U = V$ .

## Linear maps and Matrices

**Definition 1.1.8** (Linear map (or operator)). Let  $V$  and  $W$  be vector spaces over the same field  $\mathbb{F}$ . A **linear map** (or operator or transformation) from  $V$  to  $W$  is a function  $T : V \rightarrow W$  with the following properties:

- **Additivity:**

$$T(u + v) = T(u) + T(v) \quad \text{for all } u, v \in V.$$

- **Homogeneity:**

$$T(av) = aT(v) \quad \text{for all } a \in \mathbb{F} \text{ and all } v \in V.$$

We use  $L(V, W)$  to denote the set of all linear maps from  $V$  to  $W$ . ◊

**Definition 1.1.9** (Matrix of a linear map). Fix ordered bases  $\mathcal{B} = (v_1, \dots, v_n)$  of  $V$  and  $\mathcal{C} = (w_1, \dots, w_m)$  of  $W$ . For any linear map  $T : V \rightarrow W$ , there exists a unique matrix  $M_T \in \mathbb{F}^{m \times n}$  such that for all  $x \in V$ ,

$$[T(x)]_{\mathcal{C}} = M_T \cdot [x]_{\mathcal{B}},$$

where  $[x]_{\mathcal{B}}$  denotes the coordinates of vector  $x \in V$  with respect to  $\mathcal{B}$ , and  $[T(x)]_{\mathcal{C}}$  denotes the coordinates of vector  $T(x) \in W$  with respect to  $\mathcal{C}$ .

Moreover, the  $j$ -th column of  $M_T$  is  $[T(v_j)]_{\mathcal{C}}$ , i.e., the coordinates of  $T(v_j)$  with respect to  $\mathcal{C}$ . ◊

**Example 1.1.4** (A  $4 \times 3$  matrix as the matrix of a linear map). Let  $V = \mathbb{R}^3$  and  $W = \mathbb{R}^4$  over  $\mathbb{R}$ . Fix ordered bases

$$\mathcal{B} = (v_1, v_2, v_3), \quad v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

and

$$\mathcal{C} = (w_1, w_2, w_3, w_4), \quad w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad w_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Define a linear map  $T : V \rightarrow W$  by prescribing its values on the basis  $\mathcal{B}$ :

$$T(v_1) = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad T(v_2) = \begin{bmatrix} 0 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \quad T(v_3) = \begin{bmatrix} -1 \\ 0 \\ 4 \\ 1 \end{bmatrix}.$$

Because  $\mathcal{C}$  is the standard basis of  $\mathbb{R}^4$ , the coordinate vectors  $[T(v_j)]_{\mathcal{C}}$  are exactly the columns written above. Therefore, by definition,

$$M_T = [T]_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \\ 0 & 1 & 4 \\ -1 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 3},$$

whose  $j$ -th column is  $[T(v_j)]_{\mathcal{C}}$ .

Now take any  $x \in V$ . Write  $x$  in  $\mathcal{B}$ -coordinates:

$$[x]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \text{meaning} \quad x = a v_1 + b v_2 + c v_3.$$

By linearity,

$$T(x) = a T(v_1) + b T(v_2) + c T(v_3),$$

so in  $\mathcal{C}$ -coordinates,

$$[T(x)]_{\mathcal{C}} = a[T(v_1)]_{\mathcal{C}} + b[T(v_2)]_{\mathcal{C}} + c[T(v_3)]_{\mathcal{C}} = M_T \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

This verifies the defining relation  $[T(x)]_{\mathcal{C}} = M_T [x]_{\mathcal{B}}$ .

Concretely, pick  $[x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ , i.e.,  $x = 1 \cdot v_1 + 2 \cdot v_2 - 1 \cdot v_3$ . Then

$$[T(x)]_{\mathcal{C}} = M_T \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 0 \cdot 2 + (-1) \cdot (-1) \\ 1 \cdot 1 + 3 \cdot 2 + 0 \cdot (-1) \\ 0 \cdot 1 + 1 \cdot 2 + 4 \cdot (-1) \\ -1 \cdot 1 + 2 \cdot 2 + 1 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ -2 \\ 2 \end{bmatrix}.$$

Independently, compute using  $T(v_j)$ :

$$T(x) = 1 \cdot T(v_1) + 2 \cdot T(v_2) - 1 \cdot T(v_3) = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 3 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ -2 \\ 2 \end{bmatrix},$$

which matches  $[T(x)]_{\mathcal{C}}$  above, as expected.

### **Kernel, Image, Support, Rank-Nullity**

**Definition 1.1.10** (Image (aka Range or Column Space) and Rank). *The image of a linear operator  $T \in L(V, W)$ , denoted by  $\text{Im}(T)$ , is the set defined as*

$$\text{Im}(T) = \{T(v) : v \in V\}.$$

*It is also known as the range of  $T$ .*

*Note that  $\text{Im}(T)$  is exactly that vector space spanned by the column vectors of (the matrix representation of)  $T$ . For this reason,  $\text{Im}(T)$  is also known as the column space of  $T$ .*

*The dimension of  $\text{Im}(T)$  is known as the rank of  $T$ , denoted as  $\text{rank}(T)$ .*  $\diamond$

**Definition 1.1.11** (Kernel (aka Null Space) and Nullity). *The kernel of a linear operator  $T \in L(V, W)$ , denoted by  $\text{Ker}(T)$ , is the set defined as*

$$\text{Ker}(T) = \{v \in V : T(v) = \mathbf{0}\}.$$

*It is also known as the null space of  $T$ .*

*The dimension of  $\text{Ker}(T)$  is known as the nullity of  $T$ , denoted as  $\text{nullity}(T)$ .*  $\diamond$

**Theorem 1.1.3** (Rank–Nullity). *For a linear map  $T \in L(V, W)$  between finite-dimensional spaces  $V$  and  $W$ , it holds that*

$$\dim V = \text{rank}(T) + \text{nullity}(T).$$

◊

**Proposition 1.1.3.** *For  $T : V \rightarrow W$  with  $V$  finite-dimensional:*

- (a)  *$T$  is injective  $\iff \text{Ker}(T) = \{0\} \iff \text{nullity}(T) = 0$ .*
- (b)  *$T$  is surjective  $\iff \text{rank}(T) = \dim W$ .*
- (c) *If  $\dim V = \dim W$ , then injective  $\iff$  surjective  $\iff$  bijective.*

## 1.2 Complex Vector Spaces

We emphasize again that the properties listed in [Section 1.1.2](#) are general. They hold for all vector spaces.

We will be particularly interested in complex vector spaces, i.e., the case of [Definition 1.1.1](#) with  $\mathbb{F} = \mathbb{C}$ . Even after fixing  $\mathbb{F} = \mathbb{C}$ , there remain many ways to construct vector spaces satisfying [Definition 1.1.1](#)—by choosing different sets  $V$  and defining addition and scalar multiplication accordingly. All such constructions are called complex vector spaces.

Among these possibilities, the specific case with  $\mathbb{F} = \mathbb{C}$  and  $V = \mathbb{C}^n$  for some  $n \in \mathbb{N}$  (with the natural addition and multiplication) is sufficient for our purposes. This will be our main focus, and we will rarely discuss other vector spaces in this course.

### 1.2.1 Reviewing Complex Number Arithmetic

A *complex number* is a number of the form

$$z = a + bi,$$

where  $a, b \in \mathbb{R}$ , and  $i$  is the imaginary unit defined by

$$i^2 = -1.$$

The set of all complex numbers is denoted by  $\mathbb{C}$ .

For  $z = a + bi$ ,

- $a$  is called the *real part* of  $z = a + bi$ , denoted  $\text{Re}(z) = a$ .
- $b$  is called the *imaginary part* of  $z$ , denoted  $\text{Im}(z) = b$ .

### Euler's Formula and Polar Form

Probably the most important formula involving complex numbers is Euler's formula:

$$e^{i\theta} = \cos(\theta) + i \sin(\theta). \tag{1.1}$$

[Equation \(1.1\)](#) implies that every complex number  $z$  can be written in the so-called *polar form*

$$z = re^{i\theta},$$

with  $r \geq 0$  and  $\theta \in \mathbb{R}$ . Think why.

The following two identities are direct consequences of Euler's formula and are highly useful.

$$\cos \theta = \frac{1}{2} (e^{i\theta} + e^{-i\theta}) \quad \text{and} \quad \sin \theta = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).$$

One way to appreciate Euler's formula is by recognizing how it can be used to derive various trigonometric identities—identities that many students struggle to learn and memorize when first encountering them in secondary school.

As an example, let me show how to derive the identity  $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$  using Euler's formula:

$$\begin{aligned}\cos(\alpha + \beta) &= \operatorname{Re}(e^{i(\alpha+\beta)}) \\ &= \operatorname{Re}(e^{i\alpha}e^{i\beta}) \\ &= \operatorname{Re}((\cos(\alpha) + i\sin(\alpha)) \cdot (\cos(\beta) + i\sin(\beta))) \\ &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta),\end{aligned}$$

where  $\operatorname{Re}(\cdot)$  denotes the real part of a complex number.

### **Basic Arithmetic Operations**

Let  $z_1 = a + bi$  and  $z_2 = c + di$  be two complex numbers. Then:

$$z_1 + z_2 = (a + c) + (b + d)i,$$

$$z_1 - z_2 = (a - c) + (b - d)i.$$

The product of  $z_1 = a + bi$  and  $z_2 = c + di$  is given by:

$$z_1 \cdot z_2 = (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

I assume you are quite familiar with the above basic arithmetic operations on complex numbers. The only operations that might be somewhat tricky to memorize are division and roots. Let us quickly review them.

- **Division of Complex Numbers:** For two complex numbers  $z_1 = a + bi$  and  $z_2 = c + di$ , divide by multiplying the numerator and denominator by the conjugate of  $z_2$ :

$$\frac{z_1}{z_2} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(a + bi)(c - di)}{c^2 + d^2} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} = \frac{(ac + bd)}{c^2 + d^2} + \frac{(bc - ad)}{c^2 + d^2}i.$$

- **Roots of Complex Numbers:** Let  $n$  be a positive integer. For a complex number  $z = r(\cos \theta + i \sin \theta)$ , there are exactly  $n$  distinct the  $n$ -th roots. They are given by:

$$z_k = r^{\frac{1}{n}} \left( \cos \left( \frac{\theta + 2k\pi}{n} \right) + i \sin \left( \frac{\theta + 2k\pi}{n} \right) \right),$$

where  $k = 0, 1, 2, \dots, n - 1$ . These roots are the  $n$  equally spaced solutions on the complex plane.

### **De Moivre's Formula**

De Moivre's formula (also known as de Moivre's theorem or de Moivre's identity) states that for any real number  $\theta$  and integer  $n$ , the following holds:

$$(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta).$$

### Modulus of a Complex Number

The *modulus* (or absolute value) of  $z$ , denoted  $|z|$ , is the distance from  $z$  to the origin in the complex plane:

$$|z| = \sqrt{a^2 + b^2}.$$

**Equivalent characterizations.** Using complex conjugation: if  $\bar{z} = a - bi$ , then

$$|z|^2 = z\bar{z} = (a+bi)(a-bi) = a^2 + b^2, \quad \text{so} \quad |z| = \sqrt{z\bar{z}}.$$

In polar form  $z = re^{i\theta}$  with  $r \geq 0$  and  $\theta \in \mathbb{R}$ ,

$$|z| = r.$$

**Basic Properties.** For all  $z, w \in \mathbb{C}$  and  $\lambda \in \mathbb{C}$ :

$$\begin{aligned} |z| &\geq 0, \quad |z| = 0 \iff z = 0, \\ |\bar{z}| &= |z|, \quad |-z| = |z|, \\ |zw| &= |z||w|, \quad \left|\frac{z}{w}\right| = \frac{|z|}{|w|} \quad (w \neq 0), \\ |z+w| &\leq |z| + |w| \quad (\text{triangle inequality}), \\ ||z| - |w|| &\leq |z-w| \quad (\text{reverse triangle inequality}), \\ \text{If } z = re^{i\theta}, \text{ then } |z^n| &= |z|^n = r^n \text{ for } n \in \mathbb{Z}. \end{aligned}$$

## 1.3 Inner Products: Adding Geometry to Vector Spaces

We have defined vector spaces and reviewed several foundational notions in [Section 1.1.2](#). These concepts capture the algebraic structure of vector spaces and allow us to manipulate and represent vectors symbolically and computationally.

Despite this, the bare notion of a vector space is not yet rich enough for many of our goals. In particular, a general vector space does not come equipped with any “geometric structure:” there is no canonical way to talk about lengths or volumes, nor about angles or distances between vectors.

For example, it is not a priori clear how to define length and angle in the vector space  $C[a, b]$  and  $P(\mathbb{R})$  we discussed on [Page 10](#).

Without such notions, we cannot formalize geometric ideas like orthogonality, projections, or “nearest” vectors, all of which are indispensable in analysis, optimization, numerical linear algebra, signal processing, and our focus of quantum computing!

To address this, we now endow vector spaces with geometry via the concept of an *inner product*.

### 1.3.1 Starting Point: the Euclidean Space $\mathbb{R}^3$

It is widely known that the  $\mathbb{R}^3$  is a *Euclidean space*, which means that there are well-defined geometry concepts. Let's first recall this fact.

In Euclidean 3-space  $\mathbb{R}^3$ , vectors are triples  $\mathbf{v} = (v_1, v_2, v_3)$  with real components.

We have a notion of length of vectors by the so-called *Euclidean norm*:

$$\|\mathbf{v}\|_2 := \sqrt{v_1^2 + v_2^2 + v_3^2},$$

where the “2” in the subscript of  $\|\cdot\|_2$  is because that Euclidean norm is also called the quadratic norm,  $L^2$  norm, 2 norm, or square norm.

We also have a notion of distance between the “end points” of vectors by the so-called *Euclidean distance*:

$$\text{dist}(u, v) := \|u - v\|_2 = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}.$$

Actually, all these geometric notions can be unified by the concept of *inner product*. In the special case of  $\mathbb{R}^3$ , the inner product of  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$  is

$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

This single operation encodes several geometric notions: length (norm), distance, angles, and orthogonality. Let's see how.

#### Norm induced by inner product

The *norm* (length) induced by the inner product is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

**Example 1.3.1.** If  $\mathbf{v} = (3, -4, 12)$ , then  $\|\mathbf{v}\| = \sqrt{3^2 + (-4)^2 + 12^2} = \sqrt{9 + 16 + 144} = \sqrt{169} = 13$ .

#### Distance induced by norm (induced by inner product)

The distance between  $\mathbf{v}$  and  $\mathbf{w}$  is the norm of their difference:

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\| = \sqrt{(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})}.$$

**Example 1.3.2.** For  $\mathbf{v} = (1, 2, 3)$  and  $\mathbf{w} = (4, 0, -1)$ ,

$$\mathbf{v} - \mathbf{w} = (-3, 2, 4), \quad d(\mathbf{v}, \mathbf{w}) = \sqrt{(-3)^2 + 2^2 + 4^2} = \sqrt{9 + 4 + 16} = \sqrt{29}.$$

#### Angle induced by inner product

The angle  $\theta \in [0, \pi]$  between nonzero vectors is defined by the *cosine formula*

$$\cos \theta := \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|} \iff \langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta.$$

**Rationale behind this definition:** Place the tails of vectors  $\mathbf{a}$  and  $\mathbf{b}$  at the same point (say, the origin). The three sides of the triangle are:

- side 1: the segment representing  $\mathbf{a}$ , length  $\|\mathbf{a}\|$ ,
- side 2: the segment representing  $\mathbf{b}$ , length  $\|\mathbf{b}\|$ ,
- side 3: the segment from the head of  $\mathbf{b}$  to the head of  $\mathbf{a}$ , which is  $\mathbf{a} - \mathbf{b}$ , length  $\|\mathbf{a} - \mathbf{b}\|$ .

Let  $\theta$  be the angle between  $\mathbf{a}$  and  $\mathbf{b}$ , i.e., the smaller angle in  $[0, \pi]$  formed when both are drawn from the origin.

Next, apply the law of cosines to that triangle. Recall that

**Lemma 1.3.1** (law of cosines). *For any triangle with side lengths  $x, y, z$  and included angle  $\theta$  between  $x$  and  $y$ , it holds that*

$$z^2 = x^2 + y^2 - 2xy \cos \theta.$$

◊

Identify  $x = \|\mathbf{a}\|$ ,  $y = \|\mathbf{b}\|$ , and  $z = \|\mathbf{a} - \mathbf{b}\|$ . Therefore,

$$\|\mathbf{a} - \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\|\cos \theta. \quad (1.2)$$

Expand  $\|\mathbf{a} - \mathbf{b}\|^2$  using the inner product. By the inner-product definition of the norm,

$$\|\mathbf{a} - \mathbf{b}\|^2 = \langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle = \langle \mathbf{a}, \mathbf{a} \rangle - 2\langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\langle \mathbf{a}, \mathbf{b} \rangle. \quad (1.3)$$

Compare the two expressions. [Equation \(1.2\)](#) and [Equation \(1.3\)](#) and canceling common terms yields

$$\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\|\|\mathbf{b}\|\cos \theta.$$

**Example 1.3.3.** For  $\mathbf{v} = (1, 1, 0)$  and  $\mathbf{w} = (1, -1, 0)$ ,

$$\langle \mathbf{v}, \mathbf{w} \rangle = 1 \cdot 1 + 1 \cdot (-1) + 0 \cdot 0 = 0,$$

so  $\cos \theta = 0$  and  $\theta = \pi/2$  (they are perpendicular).

### Orthogonality and projections

Vectors are *orthogonal* if their inner product is zero.

Given a nonzero vector  $\mathbf{u}$ , the projection of  $\mathbf{v}$  onto  $\mathbf{u}$  is

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} = \|\mathbf{v}\| \cos(\theta) \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

The component orthogonal to  $\mathbf{u}$  is  $\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$ . See [Figure 1.1](#).

**Example 1.3.4.** Let  $\mathbf{u} = (1, 2, 2)$  and  $\mathbf{v} = (3, 0, 1)$ . Then  $\langle \mathbf{v}, \mathbf{u} \rangle = 3 \cdot 1 + 0 \cdot 2 + 1 \cdot 2 = 5$  and  $\langle \mathbf{u}, \mathbf{u} \rangle = 1 + 4 + 4 = 9$ , so

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{5}{9}(1, 2, 2) = \left( \frac{5}{9}, \frac{10}{9}, \frac{10}{9} \right),$$

and  $\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v}) = \left( \frac{22}{9}, -\frac{10}{9}, -\frac{1}{9} \right)$  is orthogonal to  $\mathbf{u}$ .

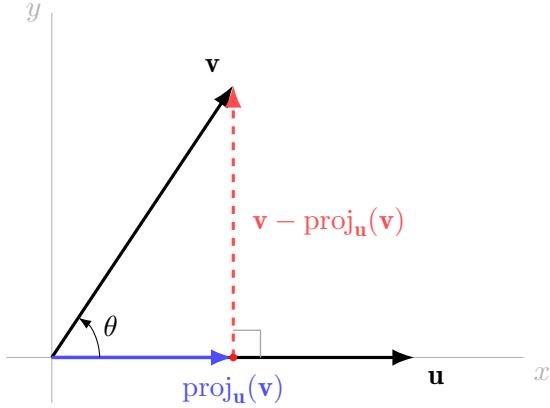


Figure 1.1: Illustrating Projection

### Key inequalities in $\mathbb{R}^3$

For all  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ :

**Cauchy–Schwarz:**  $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|,$

**Triangle inequality:**  $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$

These follow from the inner product structure.

### 1.3.2 Complex Inner-Product Space $\mathbb{C}^n$

Note that the geometric concepts we reviewed in [Section 1.3.1](#) treat the inner product as a “black box”—they do not depend on how the inner product is defined or computed. Consequently, as long as we equip a vector space with a suitable inner product, it will induce all the geometric notions we discussed in [Section 1.3.1](#).

Next, we introduce in [Definition 1.3.1](#) an additional definition concerning linear operators on vector spaces (equipped with an inner product). This definition is important and should be understood in contrast to [Definition 1.1.10](#). However, since it involves the notion of an inner product, we have to postpone its introduction until this point.

**Definition 1.3.1** (Support (aka Co-image or Row Space)). *Let  $V$  be an inner-product space and  $W$  be a vector space. The support of a linear operator  $T \in L(V, W)$ , denoted by  $\text{Supp}(T)$ , is defined to be the orthogonal complement of its kernel:*

$$\text{Supp}(T) := \text{Ker}(T)^\perp := \{v \in V : \langle v, u \rangle = 0 \ \forall u \in \text{Ker}(T)\}.$$

*It is also known as the coimage of  $T$ .*

*Note that  $\text{Supp}(T)$  is exactly that vector space spanned by the row vectors of (the matrix representation of)  $T$ . For this reason,  $\text{Supp}(T)$  is also known as the row space of  $T$ .*

◊

### Inner Product on $\mathbb{C}^n$

As we said, quantum computing uses the particular complex vector spaces  $\mathbb{C}^n$ . We now proceed

to define a inner product for it:

**Definition 1.3.2** (Standard inner product on  $\mathbb{C}^n$ ). *For  $v = (v_1, \dots, v_n), w = (w_1, \dots, w_n) \in \mathbb{C}^n$ , define*

$$\langle v, w \rangle := \sum_{i=1}^n \overline{v_i} \cdot w_i = \mathbf{v}^\dagger \mathbf{w},$$

where  $\overline{v_i}$  is the complex conjugate of  $v_i$ . ◊

**Why  $\overline{v_i}$  in Definition 1.3.2?** The most natural option—reusing the same definition as for real vector spaces (which replaces  $\overline{v_i}$  with  $v_i$ )—does not make much sense for  $\mathbb{C}^n$ . For example, if we reuse the real inner product definition to define inner product for  $\mathbb{C}^n$ , then the inner product induced norm may no longer be a real number (which cannot be a reasonable measure for “length”). On the other hand, it is easy to see that Definition 1.3.2 always ensure that  $\|v\| := \sqrt{\langle v, v \rangle}$  is a real number.

Moreover, it is easy to see that Definition 1.3.2, when restricted to real vector spaces, always becomes the same definition as

$$\langle v, w \rangle := \sum_{i=1}^n v_i \cdot w_i,$$

This is simply because that  $\bar{r} = r$  for all  $r \in \mathbb{R}$ .

Thus, Definition 1.3.2 should be considered as the more general, more “correct” definition for inner product.

**Geometry Structure over  $\mathbb{C}^n$ .** With Definition 1.3.2, we can define (almost) all the geometry concepts as in Section 1.3.1 for  $\mathbb{C}^n$ , including norm, distance, orthogonality, projection, Cauchy–Schwarz inequality and Triangle inequality. Note that the only exception is the concept of angle; its definition will be more tricky as naively copying that same definition will lead to a complex value for  $\cos(\theta)$ . Fortunately, we will not need the notion anyway in this course. Thus, we do not present further discussion here.

**Example 1.3.5** (Worked examples in  $\mathbb{C}^3$ ). *Let  $v = (1+i, 2, i)$  and  $w = (2, 1-i, -i)$  in  $\mathbb{C}^3$ . Using the convention  $\langle v, w \rangle = \sum_{k=1}^3 \overline{v_k} w_k$ , we have*

$$\begin{aligned} \langle v, w \rangle &= \overline{1+i} \cdot 2 + \overline{2} \cdot (1-i) + \overline{i} \cdot (-i) \\ &= (1-i) \cdot 2 + 2(1-i) + (-i) \cdot (-i) \\ &= 2 - 2i + 2 - 2i - 1 \\ &= 3 - 4i, \end{aligned}$$

$$\begin{aligned} \|v\|^2 &= |1+i|^2 + |2|^2 + |i|^2 = 2 + 4 + 1 = 7, \quad \|v\| = \sqrt{7}, \\ \|w\|^2 &= |2|^2 + |1-i|^2 + |-i|^2 = 4 + 2 + 1 = 7, \quad \|w\| = \sqrt{7}. \end{aligned}$$

*Orthogonality holds iff  $\langle v, w \rangle = 0$ .*

## 1.4 Hilbert Space—the Playground for Quantum Computing

Our course on quantum computing will take place entirely in the complex vector space  $\mathbb{C}^n$ , equipped with the (standard) inner product:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n \overline{u_i} v_i = \mathbf{u}^\dagger \mathbf{v}.$$

This vector space is an example of the so-called *Hilbert space*, which is named after the famous German mathematician David Hilbert.

You will often hear people in quantum computing talk about Hilbert space. In most cases, this simply refers to this inner-product space  $\mathbb{C}^n$  we defined so far (or its infinite-dimensional analogues).

There are two caveats worth noting:

- The term “Hilbert space” refers to a class of inner-product spaces that satisfy certain axioms. The space  $\mathbb{C}^n$  with the inner product above is one example, but there are many others (including infinite-dimensional spaces) that also qualify as Hilbert spaces.
- The formal definition of a Hilbert space involves some math-analytic conditions, notably *completeness*. We will not develop the full theory in this undergraduate course. For those who are interested, please refer to [Supplementary Box 1.4.1](#).

Nothing from [Supplementary Box 1.4.1](#) will be tested on quizzes or exams!

### Supplementary Box 1.4.1: What is a Hilbert Space?

A Hilbert space is a *complete* (real or complex) inner-product space. That is, it is a *complete* vector space equipped with an inner product.

We have already explained what a vector space and inner product are. The only thing left is the notion of *completeness*. We now explain this term:

- This term is used to describe a *metric space*. We do not formally define a metric space. Roughly, it is a space that allows us to talk about the concept of “distance”. An inner-product space is always a metric space, because inner product induces a norm, and a norm always induces a metric  $d(\cdot, \cdot)$  as follows:

$$d(\mathbf{u}, \mathbf{v}) := \|\mathbf{u} - \mathbf{v}\| := \sqrt{\langle \mathbf{u}, \mathbf{v} \rangle}.$$

- In mathematical analysis, a metric space  $M$  is called *complete* (or a Cauchy space) if every Cauchy sequence of points in  $M$  has a limit that is also in  $M$ .

Intuitively, a space is complete if there are “no points missing from it” (inside or at the boundary). For instance, the set of rational numbers is not complete—for example, the number  $e$  is “missing” from it, even though one can construct a Cauchy sequence of rational numbers (i.e., the following Maclaurin series) that converges to  $e$ .

- Note that  $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ .
- Let  $s_i = \sum_{k=0}^i \frac{1}{k!}$ . Then,  $\{s_i\}$  is a Cauchy sequence as it satisfies the Cauchy property that:  $\forall \varepsilon > 0, \exists N$  such that  $\forall m, n > N$ , it holds that  $|s_m - s_n| < \varepsilon$ .

- Also,  $\lim_{n \rightarrow \infty} s_n = e$ .
- Essentially, a complete metric space is the minimal place where you can talk about concepts like *limit* and *convergence*. It has nice math-analytical properties.

## 1.5 \*Bonus Section: Why Complex Numbers for Quantum Mechanics?

Nothing from [Section 1.5](#) will be tested on quizzes or exams!

We keep saying that quantum mechanics and quantum computing need to use the mathematical language of complex vector spaces—more precisely, Hilbert spaces. A natural question is why we must use complex numbers, and why we cannot simply treat everything using real vector spaces, which we are more familiar and comfortable with.

In this bonus section, we provide some partial explanations for the necessity of complex numbers in quantum mechanics.

Please note that what we provide here should be viewed as rationale and clues pointing to the indispensable role of complex numbers in quantum mechanics. As for a definitive reason why quantum mechanics must require complex numbers, that remains beyond the current reach of human knowledge in physics.

### 1.5.1 Why Complex Numbers?

#### Short Answer

**Real-valued quantum mechanics fails to capture all phenomena.** Attempts to describe quantum mechanics with only real numbers result in a restricted theory that cannot reproduce all experimental results, such as entanglement and certain interference effects.

#### Longer answer

We provide more details. But what provided below are not really an “answer” yet. They are just some justification/rationale behind the essential role of complex numbers in quantum mechanics.

1. Complex numbers are not unique to quantum mechanics. They are already there in classical physics, specially in the topics involving oscillatory phenomena and **wave functions**. See [Section 1.5.2](#) and [Section 1.5.3](#) as an example—it is more mathematically convenient and straightforward to express these concepts in terms of complex exponentials rather than real trigonometric functions.
2. Quantum mechanics have a **wave-like nature**:
  - As demonstrated by phenomena such as interference and diffraction (check [this short YouTube video](#) on double-split experiment).
  - Schrödinger equation inherently uses complex numbers. For a non-relativistic particle, the time-dependent Schrödinger equation is:

$$i\hbar \frac{\partial}{\partial t} |\psi(x,t)\rangle = \hat{H} |\psi(x,t)\rangle, \quad (1.4)$$

where:

- $\psi(x,t)$  is the wavefunction (a complex-valued function),
- $\hbar$  is the reduced Planck constant,
- $\hat{H}$  is the Hamiltonian operator (energy operator).

**Moreover**, the exemplary systems described in Sections 1.5.2 and 1.5.3 are second-order differential equations. Cosine functions are already sufficient to serve as their solutions; it is fine if you think exponential functions as being overkilling. In contrast, the Schrödinger's equation is 1st order differential equation; this makes exponential function necessary. This serves as one of the clue about the necessity of the  $e^{it}$  wave function form in QM.

3. The explanations above merely state that “complex numbers are a matter of mathematical convenience,” but they do not prove that complex numbers are truly necessary. One can still ask:

**Question:** *Are complex numbers really essential for quantum mechanics?*

Indeed, the answer is NO for classical physics like oscillatory phenomenon and wave functions. So, it might be natural to hope that complex numbers are not essential for quantum mechanics as well. However, for quantum mechanics, the answer to this question is far more nuanced. Even in the past five years, both experimental and theoretical research has continued to explore this topic. Eventually, the answer is YES! We will not dive into further details. For more insights, I recommend reading [Quantum Mechanics Must Be Complex](#), by Alessio Avella.

### 1.5.2 Complex Numbers in Simple Harmonic Oscillation

The restoring force in a simple harmonic oscillator is given by Hooke's law:

$$\mathbf{F} = -k\mathbf{x},$$

where:

- $\mathbf{F}$  is the restoring force,
- $k$  is the spring constant (a measure of stiffness),
- $\mathbf{x}$  is the displacement from equilibrium.

The negative sign indicates that the force acts in the opposite direction to the displacement.

We apply Newton's second law:

$$\mathbf{F} = m\mathbf{a},$$

where

$$\mathbf{a} = \ddot{\mathbf{x}}$$

(the second derivative of  $x$  with respect to time, i.e., acceleration). Substituting  $F = -kx$ , we get:

$$m\ddot{\mathbf{x}} = -k\mathbf{x}.$$

Rearranging, we obtain the simple harmonic oscillator equation:

$$\ddot{\mathbf{x}} + \frac{k}{m}\mathbf{x} = 0.$$

Define

$$\omega^2 = \frac{k}{m},$$

where  $\omega$  is the *angular frequency* of the oscillator. The equation becomes:

$$\ddot{\mathbf{x}} + \omega^2\mathbf{x} = 0 \quad (1.5)$$

**Solving the Equation.** Equation (1.5) is a second-order differential equation. Most of you have learned the following solution

$$x(t) = C \cos(\omega t + \phi),$$

where:

- $C$  is the amplitude (maximum displacement),
- $\phi$  is the phase angle (determined by the initial conditions).

This form is often more intuitive, as it directly describes the motion in terms of amplitude, frequency, and phase.

However, it is not hard to see that  $x(t) = C \cdot e^{i(\omega t + \phi)}$  is also a solution to the simply harmonic oscillation equation. Indeed,  $C \cos(\omega t + \phi)$  is nothing but the real part of  $C \cdot e^{i(\omega t + \phi)}$ , by Euler's formula. The form  $x(t) = C \cdot e^{i(\omega t + \phi)}$  enjoys the following advantages:

1. On the complex plane, the real part of  $C \cdot e^{i(\omega t + \phi)}$  can be perceived as the “shadow” (formally, the projection) of  $C \cdot e^{i(\omega t + \phi)}$  on the  $x$  axis while  $x(t) = C \cdot e^{i(\omega t + \phi)}$  moves as a circle. That is, while  $x(t)$  is rotating around the origin, the real part  $\text{Re}(x(t)) = |C| \cos(\omega t + \phi)$  is a point bouncing between  $-|C|$  and  $|C|$  on the real line, as illustrated in Figure 1.2.
2. Exponential functions are much more convenient for many calculations, such as differentiation and integration.

### 1.5.3 Complex Numbers in Electromagnetic Wave Equations

The electromagnetic wave equations in free space is given as:

$$\begin{aligned} \nabla^2 \mathbf{E} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} &= 0, \\ \nabla^2 \mathbf{B} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} &= 0. \end{aligned}$$

These equations describe how the electric field ( $\mathbf{E}$ ) and magnetic field ( $\mathbf{B}$ ) propagate in free space. Let us break down and explain the terms:

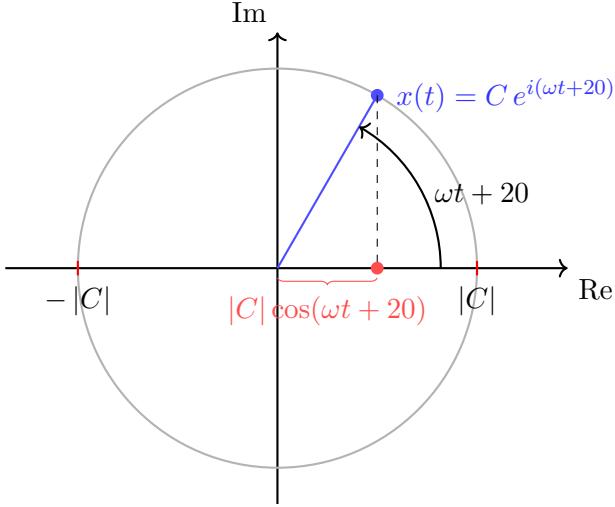


Figure 1.2: Example for  $x(t) = C \cdot e^{i(\omega t + \phi)}$  with  $\phi = 20$

1. **E:** The **electric field**, a vector field that represents the force experienced by a charged particle in the presence of an electric field. It is a function of position ( $\mathbf{r}$ ) and time ( $t$ ):

$$\mathbf{E} = \mathbf{E}(\mathbf{r}, t).$$

In an electromagnetic wave,  $\mathbf{E}$  oscillates perpendicularly to the direction of propagation and to the magnetic field  $\mathbf{B}$ .

2. **B:** The **magnetic field**, a vector field that represents the force on a moving charged particle in the presence of a magnetic field. It is also a function of position and time:

$$\mathbf{B} = \mathbf{B}(\mathbf{r}, t).$$

In an electromagnetic wave,  $\mathbf{B}$  oscillates perpendicularly to the direction of propagation and to the electric field  $\mathbf{E}$ .

3.  $\nabla^2$ : The **Laplacian operator**, which represents the spatial variation of the field. In Cartesian coordinates, it is defined as:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

For example,  $\nabla^2 \mathbf{E}$  describes how the electric field varies in space.

4.  $\frac{\partial^2}{\partial t^2}$ : The **second time derivative**, which describes how the field (either  $\mathbf{E}$  or  $\mathbf{B}$ ) varies with time. For example,  $\frac{\partial^2 \mathbf{E}}{\partial t^2}$  describes the temporal change of the electric field.

5.  $\mu_0$ : The **permeability of free space**, a physical constant that describes how magnetic fields interact with the vacuum. Its value is:

$$\mu_0 = 4\pi \times 10^{-7} \text{ N/A}^2.$$

6.  $\epsilon_0$ : The **permittivity of free space**, a physical constant that describes how electric fields

interact with the vacuum. Its value is:

$$\varepsilon_0 \approx 8.854 \times 10^{-12} \text{ F/m.}$$

7.  $\mu_0\varepsilon_0$ : This combination determines the **speed of light** in free space, given by:

$$c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}.$$

Numerically:

$$c \approx 3 \times 10^8 \text{ m/s.}$$

## **Wave Solutions**

The solutions to the electromagnetic wave equation are sinusoidal traveling waves. For the electric field **E**, the solution can be written as:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \phi),$$

where:

- $\mathbf{E}_0$ : The amplitude of the electric field,
- $\mathbf{k}$ : The wavevector, which gives the direction of propagation and the wavenumber ( $k = 2\pi/\lambda$ , where  $\lambda$  is the wavelength),
- $\omega$ : The angular frequency ( $\omega = 2\pi f$ , where  $f$  is the frequency),
- $\phi$ : The phase constant, determined by initial conditions,
- $\mathbf{r}$ : The position vector,  $(x, y, z)$ ,
- $t$ : Time.

Similarly, for the magnetic field **B**, the solution is:

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{B}_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \phi).$$

### **Supplementary Box 1.5.1: Relation between Electric and Magnetic Fields**

The electric field (**E**) and magnetic field (**B**) in an electromagnetic wave are related as follows:

1. **E**, **B**, and the wavevector **k** (direction of wave propagation) are mutually perpendicular:

$$\mathbf{E} \perp \mathbf{B} \perp \mathbf{k}.$$

2. The magnitudes of **E** and **B** are related by:

$$|\mathbf{B}| = \frac{|\mathbf{E}|}{c},$$

where  $c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$  is the speed of light.

# Chapter 2

## Postulate 1: State Space

Let us start by stating the first postulate:

**Postulate (1/4):** An (isolated) physical system is completely described by its state vector, which is a *unit* vector in some Hilbert space.

Recall that we defined a Hilbert space in [Section 1.4](#). In particular, we stated that we will focus exclusively on a specific Hilbert space, namely  $\mathbb{C}^n$  for some dimension  $n$ . Under this convention, **Postulate 1** states that for any given physical system, there exists a dimension parameter  $n$  such that the state of the system is completely described by a vector  $\mathbf{v} \in \mathbb{C}^n$  with  $\|\mathbf{v}\| = 1$ .

Some important highlights:

- Recall that the norm  $\|\mathbf{v}\|$  on  $\mathbb{C}^n$  is defined to be the inner product induced norm

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\mathbf{v}^\dagger \mathbf{v}}.$$

- The dimension  $n$  depends on the physical system under consideration. Different physical systems may require different values of  $n$  in  $\mathbb{C}^n$  to describe their states.
- We emphasize the *unit-norm* requirement: although there are infinitely many vectors in the vector space  $\mathbb{C}^n$ , only those with length 1 (i.e., unit vectors) can be used to represent quantum states.

### 2.1 Two-Dimensional Quantum Systems—A Qubit

The simplest quantum system is when  $n = 2$ , i.e., the space  $\mathbb{C}^2$ .

The state of such a 2D quantum system can always be fully described by a unit vector  $v \in \mathbb{C}^2$ , and every unit vector  $v \in \mathbb{C}^2$  could possibly be corresponding to a state of the system.

Such a 2D quantum system is referred to as a *qubit*.

Before we dive into more mathematical discussions of such 2D quantum systems, let us first see two examples to help us imagine the them.

**Remark 2.1.1** (How does the microscopic world look like?). *Before we begin, a quick clarification: we will sometimes use simplified, intuitive diagrams that depict atoms or other microscopic particles as little dots or balls moving in fixed orbits. Please keep in mind that this is not accurate. In the microscopic realm, our everyday intuitions about position and shape no longer apply. It is a quantum world, where familiar quantities such as position, speed, and kinetic energy are represented by wave functions.*

*For example, you should not imagine an electron as a tiny ball located at a precise point in space. Rather, an electron is better described by a probability distribution: before measurement, it is spread out and “exists” everywhere that its wave function has support, and upon observation it*

appears at a particular location with a probability determined by that wave function. For a more accurate visualization of these ideas, see [this YouTube video](#).

### 2.1.1 Example 1: Spin- $\frac{1}{2}$ in a Magnetic Field

Nothing from [Section 2.1.1](#) will be tested on quizzes or exams!

The first example involves the physical concept of *spin*.

**What Spin Is.** Many particles (electrons, protons, some nuclei) behave as if each carries a built-in, permanent tiny bar magnet. This built-in magnetism is called *spin*. Spin is as basic to a particle as its mass or electric charge. It's not something added later; the particle simply has it. In a magnetic field, this tiny magnet cannot take on any arbitrary orientation with any value; only specific outcomes appear when you measure<sup>1</sup> along a chosen direction. For the simplest case (spin- $\frac{1}{2}$ , like the electron), measurement along a given axis yields one of two results: "up" or "down" relative to that axis.

**What Spin Is Not.** The particle is not a tiny ball physically spinning in space. Modeling the electron as a spinning sphere leads to contradictions (like surface speeds faster than light). "Spin" is a quantum property that *acts like* angular momentum and magnetism.

Spin is neither a literal rotation nor a hidden arrow you can simply peek at. We draw an arrow for intuition, but measuring along different directions yields outcomes that no single pre-existing classical arrow can fully explain. That is the quantum part.

#### How We Know Spin Is Real:

- **Stern-Gerlach experiment.** A beam of atoms sent through a non-uniform magnetic field splits into two distinct spots, not a smear. This matches the idea of a tiny magnet with only two allowed outcomes along the field direction.
- **Magnetic resonance.** In a magnetic field, applying an oscillating, sideways field at just the right frequency flips spins in a controlled way. This is the basis of NMR (nuclear magnetic resonance) and MRI.

#### How Spin Interacts with Magnetic Fields

- **Energy preference.** In a magnetic field, "aligned with the field" and "against the field" have slightly different energies. The difference grows with stronger fields.
- **A natural resonance frequency.** That energy difference corresponds to a natural rhythm (Larmor frequency). Driving the system with a sideways oscillating magnetic field at this frequency moves the spin between the two orientations.

**Measuring Spin.** A measurement along a chosen axis yields one of the allowed discrete results (for spin- $\frac{1}{2}$ : "up" or "down"). Importantly, note that *measure along a different axis and the outcomes and their probabilities change*.

<sup>1</sup>Technically, we have not yet introduced the concept of "measurement" in a formal way; we will do so in a later chapter. For now, you can think of "measuring" intuitively as taking some action to detect the current value of a particle's spin.

## 2.1.2 Example 2: Two-Level Atom (or “Artificial Atom”)

Nothing from [Section 2.1.2](#) will be tested on quizzes or exams!

**What a Two-Level Atom Is.** Many physical systems (real atoms, trapped ions, and “artificial atoms” such as superconducting circuits) have many possible energy levels. In practice, we often choose *two* of these—a lower-energy *ground* level and a higher-energy *excited* level—and arrange our experiment so that everything we do mainly affects just this pair.

In that regime, the atom effectively behaves like a system with two options, much like a switch that can be “ground” or “excited.” The energy gap between these two levels sets a specific color of light (for optical transitions) or a specific microwave frequency (for artificial atoms) that most strongly interacts with the pair.

**What a Two-Level Atom Is Not.** It is not that the atom truly has only two levels; rather, we *select* two and isolate them from the rest by careful tuning. We avoid accidentally using light at other colors (or microwaves at other frequencies) that would drive unwanted transitions. So, “two-level” is a well-engineered approximation: extremely accurate when we respect the conditions that keep other levels quiet.

### How We Know This Picture Works:

- **Resonant absorption and emission.** When we shine light at the right color (or apply the right microwave frequency), the system efficiently moves *population*<sup>2</sup> between ground and excited and can emit light back at that same color. Off-color light barely does anything by comparison.
- **Rabi flopping (controlled swapping).** If we turn on the right-frequency drive for a controlled time, the population sloshes back and forth between ground and excited in a regular, predictable way. Turning the drive off at the right moment lets us stop in ground, excited, or any desired mix of the two.

### How the Two-Level Atom Interacts with Light (or Microwaves).

- **Energy gap sets the “right” frequency.** The energy difference between ground and excited corresponds to a specific frequency. Driving at this frequency most effectively couples the two states; driving far away from it has little effect.
- **Pulse duration and phase control the outcome.** Short bursts nudge population partway; longer bursts can complete a swap. Adjusting the timing and the *phase* of the drive changes how the two states are mixed, giving flexible control.

**Measuring a Two-Level Atom.** We often arrange things so the two states respond differently to our probe:

---

<sup>2</sup>In this context, the term “population” means how many atoms (or what fraction/probability for a single atom) are in each of the two states: ground vs. excited. For a single atom, “population” means the chance you would find it in ground or in excited if you measured it right now. For many identical atoms, it is literally how many are in each state.

- **Atoms/ions (optical readout).** A detection laser makes only one state glow (“bright”), while the other stays dark. Counting the glow tells us which state the atom was in.
- **Artificial atoms (circuit readout).** A nearby microwave resonator acts like a sensitive listener whose pitch shifts slightly depending on whether the system is in ground or excited; we infer the state from that shift.

**Everyday Intuition and Classroom Analogy.** Think of a playground swing. It naturally moves back and forth at its own rhythm. A gentle push at exactly that rhythm transfers energy efficiently, carrying the swing from low to high and back. In the same way, light (or microwaves) at the right frequency efficiently moves population between ground and excited. Pushing off-rhythm does much less.

The above intuitive analogy is useful for beginners. But we do feel obligated to highlight the following caveats:

- Two levels vs a continuum: A swing moves through a continuum of heights; a two-level atom has only two energy levels.
- Probability, not position: “Population” means probability (or fraction) in ground/excited, not a physical angle like a swing.
- Superposition and phase: The atom can be in a superposition with a phase; the swing has no true analog of quantum phase.
- “Measurement” matters: Measuring the atom changes its state; watching a swing does not affect anything.
- Off-resonance still works: Light slightly off the exact frequency can still drive the atom (less efficiently); the swing analogy suggests almost no effect.

## 2.2 The Math for a Single Qubit

From now on, we will set aside the physics and take a purely mathematical view of a qubit: it is simply a unit vector in  $\mathbb{C}^2$ , corresponding to some quantum physical system (like those discussed in [Section 2.1](#)) that we no longer need to consider.

**Analogy to a classical bit.** Classically, we think of a bit  $b \in \{0, 1\}$  as something stored in a register within the CUP. Note that this is also a mathematical abstraction:  $b$  is logical data, and the register is some piece of computer hardware whose implementation details we can ignore.

A *qubit* is just a *quantum bit*. By the same analogy, it is logical data encoded as a unit vector  $\mathbf{v} \in \mathbb{C}^2$ , stored in a “quantum register.” The “quantum register” is some quantum computer hardware whose implementation we do not need to specify (which is exactly the underlying 2D quantum physical system).

Ultimately, the only thing we care about is how the math for a bit (and now, a qubit) works: how to describe it and how to use it to perform computation.

**How to specify a single qubit.** It is a vector. So, do what we will do for vectors:

- Find a basis and write it as a linear combination of the basis vectors.

Recall that the linear algebra theory is invariant to the choice of basis. Let us start with our favorite basis for  $\mathbb{C}^2$ : the standard basis

$$\mathbf{e}_0 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_1 := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(Notice that computer scientists prefer that the subscript starts from 0.)

Using the standard basis, a qubit  $\mathbf{v}$  can always be written as

$$\mathbf{v} = \alpha_0 \cdot \mathbf{e}_0 + \alpha_1 \cdot \mathbf{e}_1,$$

where  $\alpha_0, \alpha_1 \in \mathbb{C}$  and  $|\alpha_0|^2 + |\alpha_1|^2 = 1$ . The coefficients  $\alpha_b$  ( $b \in \{0, 1\}$ ) are referred to as *amplitudes*.

#### Teaching Suggestions

Derive the  $|\alpha_0|^2 + |\alpha_1|^2 = 1$  condition from the “unit” requirement

$$1 = \|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\mathbf{v}^\dagger \mathbf{v}}.$$

(Help students recall that the definition of inner product in  $\mathbb{C}^n$  requires the *conjugate transpose*  $\mathbf{v}^\dagger$ .)

This condition is often known as the *normalization condition* for state vectors.

Note that if you choose a different basis, the value of the amplitudes will change. You have complete freedom to choose the basis. We typically prefer the standard basis for its mathematical convenience, but for certain calculations a different basis may be more convenient. **Always be aware of the basis you are working in!**

**Data encoding by a single qubit.** Note that we can use a classical bit to encode a binary bit of either 0 or 1, and do some computation over it.

We do similar things on a quantum computer. We use a qubit to encode data and will operate the qubit to finish some computation to our interest.

Note that there are infinitely many unit vectors in  $\mathbb{C}^2$ . This immediately leads to a strange thing:

- In contrast to a classical bit, it seems a qubit can encode an infinite amount of information?!

Yes, it does! However, this does not mean we have infinite encoding or computational power. Quantum mechanics imposes quite stringent restrictions on how we can operate on a qubit and how we can read the data encoded in it. We will learn these rules as the following postulates. For now, the main takeaway is:

- *A single qubit can, in principle, encode an infinite amount of information, but quantum mechanics restricts how we can access and use this unbounded resource in precise ways that we will study later!*

## 2.3 Quantum Systems of Higher Dimensions

We have already learned a great deal about two-dimensional (2D) quantum systems. Before moving on, it is helpful to recognize that higher-dimensional quantum systems also exist and are widely used.

While providing concrete physical examples would require additional physics background beyond the scope of this course, we can still develop the mathematics cleanly. In fact, such systems are naturally understood as straightforward generalizations of the 2D case to an arbitrary dimension parameter  $n$ :

- **State space.** An  $n$ -dimensional quantum system corresponds to the Hilbert space  $\mathbb{C}^n$ . That is, its (pure) state can be completely specified by a unit vector  $\mathbf{v} \in \mathbb{C}^n$ , and each unit vector in this space can in principle describe a possible state of the system under appropriate conditions (for instance, at different times or after different operations).
- **Remark on physical realizations.** Such  $n$ -dimensional systems do exist in the real world; however, they are not our focus here, so we will not provide specific examples.
- **Coordinates via a basis.** To describe the state of an  $n$ -dimensional system concretely, we choose a basis of  $\mathbb{C}^n$  and express the state vector  $\mathbf{v}$  as a linear combination of the basis vectors, exactly as we did in the 2D case.
- **A convenient choice: the standard basis.** A particularly convenient choice is the standard basis for  $\mathbb{C}^n$ :

$$\mathbf{e}_0 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_1 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Here,  $\mathbf{e}_j$  (for  $j \in \{0, 1, \dots, n-1\}$ ) is the vector whose  $j$ -th entry is 1 and all other entries are 0. (Note that we follow the computer-science convention: the first entry of a vector is indexed as the 0-th entry.)

- **Expression of a unit vector:** In the standard basis, a unit vector  $\mathbf{v} \in \mathbb{C}^n$  can always be written as

$$\mathbf{v} = \sum_{j=0}^{n-1} \alpha_j \cdot \mathbf{e}_j,$$

where  $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{C}$  satisfy the normalization condition  $\sum_{j=0}^{n-1} |\alpha_j|^2 = 1$ .

## 2.4 Dirac Notation

Dirac<sup>3</sup> introduced a concise and powerful notation that is extremely convenient for quantum calculations. This notation has become the standard in quantum mechanics and quantum computing, although it is not widely adopted by mathematicians. One reason is that the notation is tailored to Hilbert spaces, whereas mathematicians often work in more general linear-algebraic settings. In spaces that are not Hilbert (for example, where no inner product is defined), Dirac's notation is less natural or not directly applicable.

Fix a positive integer  $n$  and consider the space  $\mathbb{C}^n$ .

### The “ket” for column vectors

---

<sup>3</sup>Paul A. M. Dirac (1902–1984) was a pioneering theoretical physicist and one of the founders of quantum mechanics and quantum field theory. He formulated the Dirac equation for relativistic electrons, predicted the existence of antimatter, and introduced the bra–ket (Dirac) notation that is standard in quantum mechanics. He shared the 1933 Nobel Prize in Physics.

We use the “ket” notation  $|\psi\rangle$  to denote a **column vector** in  $\mathbb{C}^n$ . The symbol  $|\psi\rangle$  is read as “ket  $\psi$ .”

For example, in  $\mathbb{C}^4$  we may assign

$$|\psi\rangle = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 5 \end{bmatrix}.$$

This is merely a change of symbols: instead of the familiar  $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 5 \end{bmatrix}$ , we write  $|\psi\rangle$ . The label  $\psi$

in the notation  $|\psi\rangle$  is a variable, just as  $v$  is in  $\mathbf{v}$ ; you are free to choose any letter inside the ket. However, the “ $|\cdot\rangle$ ” part indicates that the object is a vector, analogous to using boldface  $\mathbf{v}$  to signal a vector.

### Notation for an orthonormal basis

In  $\mathbb{C}^n$ , we denote an **orthonormal basis** by

$$\{|0\rangle, |1\rangle, \dots, |n-1\rangle\}.$$

This notation specifies only that these symbols represent some orthonormal basis; it does not, by itself, fix which orthonormal basis is intended. When you use this notation, you should state explicitly which concrete orthonormal basis the symbols refer to, to avoid ambiguity.

For example, in  $\mathbb{C}^4$  you might take the standard basis

$$|0\rangle = \mathbf{e}_0 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |1\rangle = \mathbf{e}_1 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |2\rangle = \mathbf{e}_2 := \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |3\rangle = \mathbf{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Alternatively, you could also use the following orthonormal basis:

$$|0\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad |1\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad |2\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \quad |3\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix}.$$

We emphasize again that  $\{|0\rangle, \dots, |n-1\rangle\}$  is just a naming convention for whichever orthonormal basis you choose. You have total freedom to associate it to a specific orthonormal basis.

In many textbooks and papers on quantum computing, if no basis is explicitly specified, the default convention is the standard basis:

$$|0\rangle = \mathbf{e}_0, \quad |1\rangle = \mathbf{e}_1, \dots, |n-1\rangle = \mathbf{e}_{n-1}.$$

Once a basis is fixed, any vector  $|\psi\rangle$  can be expressed as a linear combination

$$|\psi\rangle = \alpha_0 |0\rangle + \dots + \alpha_{n-1} |n-1\rangle,$$

where the coefficients (amplitudes)  $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{C}$ .

### The “bra” for conjugate transpose

The “bra”  $\langle \psi |$  denotes the conjugate transpose of  $|\psi\rangle$ ; it is a row vector. For example, if

$$|\psi\rangle = \begin{bmatrix} 1 \\ 2+3i \\ 5i \\ 7 \end{bmatrix},$$

then

$$\langle \psi | = [1 \ 2-3i \ -5i \ 7].$$

### Standard linear algebra in Dirac notation

Dirac notation is simply an alternative symbolic system for the same linear-algebraic objects. All computations follow the usual rules of linear algebra.

A few operations to emphasize:

- **Inner product:**  $\langle \phi | \cdot | \psi \rangle$  is the inner product of  $|\phi\rangle$  and  $|\psi\rangle$ . It is often abbreviated as  $\langle \phi | \psi \rangle$ .
- **Ket–bra products:** The operator  $|u\rangle \cdot \langle v|$  (often abbreviated as  $|u\rangle \langle v|$ ) is defined as the outer product of  $|u\rangle$  and  $\langle v|$ , i.e., the matrix whose  $(i,j)$ -entry is  $u_i \bar{v}_j$ .
- **Matrix multiplication:** Matrix multiplication with kets/bras behaves as expected. For example, for any  $|a\rangle, |b\rangle, |c\rangle \in \mathbb{C}^n$  and  $\mathbf{M} \in \mathbb{C}^{n \times n}$ , the following equations hold:

$$\begin{aligned} |a\rangle \langle b| |c\rangle &= \langle b | c \rangle |a\rangle, \\ \langle c | |a\rangle \langle b| &= \langle c | a \rangle \langle b |, \\ |a\rangle \langle b| \mathbf{M} &= |a\rangle (\langle b | \mathbf{M}), \\ \mathbf{M} |a\rangle \langle b| &= (\mathbf{M} |a\rangle) \langle b|. \end{aligned}$$

These identities are direct consequences of associativity of matrix multiplication and the definitions of inner and outer products.

#### 2.4.1 Worked Examples: Calculations in Dirac Notation

##### Example 1: Computing a bra from a ket

Let

$$|\psi\rangle = \begin{bmatrix} 1 \\ 2+3i \\ -i \\ 0 \end{bmatrix}.$$

Find  $\langle \psi |$ .

**Solution.** Take the conjugate transpose (transpose and conjugate entries):

$$\langle \psi | = [1 \ 2-3i \ i \ 0].$$

### **Example 2: Inner products and norms**

Let

$$|\phi\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad |\psi\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}.$$

Compute  $\langle \phi | \psi \rangle$  and  $\| |\psi\rangle \|$ .

**Solution.**

$$\langle \phi | = \frac{1}{2} [1 \ 1 \ 1 \ 1], \quad \langle \phi | \psi \rangle = \frac{1}{4} (1 \cdot 1 + 1 \cdot i + 1 \cdot (-1) + 1 \cdot (-i)) = \frac{1}{4} (1 + i - 1 - i) = 0.$$

Thus  $|\phi\rangle$  and  $|\psi\rangle$  are orthogonal.

For the norm,

$$\| |\psi\rangle \|^2 = \langle \psi | \psi \rangle = \frac{1}{4} (1 + |i|^2 + |-1|^2 + |-i|^2) = \frac{1}{4} (1 + 1 + 1 + 1) = 1,$$

so  $\| |\psi\rangle \| = 1$ .

### **Example 3: Expanding a vector in the computational basis**

Let the (standard) orthonormal basis of  $\mathbb{C}^4$  be  $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$  with  $|0\rangle = \mathbf{e}_0, \dots, |3\rangle = \mathbf{e}_3$ . Given

$$|\psi\rangle = \begin{bmatrix} 2 \\ -i \\ 0 \\ 1 \end{bmatrix},$$

write  $|\psi\rangle$  as a linear combination of  $|0\rangle, |1\rangle, |2\rangle, |3\rangle$ .

**Solution.** Read off coordinates:

$$|\psi\rangle = 2|0\rangle - i|1\rangle + 0 \cdot |2\rangle + 1 \cdot |3\rangle = 2|0\rangle - i|1\rangle + |3\rangle.$$

### **Example 4: Outer products (rank-1 operators)**

Let

$$|u\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |v\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Compute the matrix  $|u\rangle\langle v|$  and its action on  $|v\rangle$ .

**Solution.**

$$\langle v | = \frac{1}{\sqrt{2}} [1 \ i], \quad |u\rangle\langle v| = \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \left( \frac{1}{\sqrt{2}} [1 \ i] \right) = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix}.$$

Action on  $|v\rangle$ :

$$|u\rangle\langle v| |v\rangle = \langle v | v \rangle |u\rangle = 1 \cdot |u\rangle = |u\rangle,$$

since  $|v\rangle$  is normalized.

### Example 5: Projectors and their properties

Fix a  $|a\rangle \in \mathbb{C}^n$ . Define a matrix  $P := |a\rangle\langle a|$ . Such a  $P$  corresponds to the projection operation on the “direction”<sup>4</sup> specified by  $|a\rangle$ .

Show that  $P^2 = P$  and  $P^\dagger = P$ .

**Solution.** First,

$$P^2 = |a\rangle\langle a| |a\rangle\langle a| = |a\rangle \underbrace{\langle a|a\rangle}_{=1} \langle a| = |a\rangle\langle a| = P.$$

Second,

$$P^\dagger = (|a\rangle\langle a|)^\dagger = |a\rangle\langle a|,$$

since  $(|\cdot\rangle)^\dagger = \langle\cdot|$  and  $(\langle\cdot|)^\dagger = |\cdot\rangle$ .

### Example 6: Orthogonal decomposition with a projector

Let  $|a\rangle$  be a unit vector and  $P = |a\rangle\langle a|$ . For any  $|\psi\rangle$ , decompose  $|\psi\rangle$  into components parallel and orthogonal to  $|a\rangle$ . That is, write  $|\psi\rangle$  as a linear combination of two vectors  $|\psi\rangle_{\text{para}}$  and  $|\psi\rangle_{\text{orth}}$  such that  $|\psi\rangle_{\text{para}}$  is in parallel with  $|a\rangle$  and  $|\psi\rangle_{\text{orth}}$  is orthogonal to  $|a\rangle$ .

**Solution.** The parallel component is

$$|\psi\rangle_{\text{para}} = P|\psi\rangle = |a\rangle\langle a| |\psi\rangle = \langle a|\psi\rangle |a\rangle.$$

The orthogonal component is

$$|\psi\rangle_{\text{orth}} = (I - P)|\psi\rangle.$$

### Example 7: Change of basis

In  $\mathbb{C}^n$ , let  $\{|f_0\rangle, \dots, |f_{n-1}\rangle\}$  be a set of orthonormal basis. Let  $U$  be a matrix whose columns are  $|f_0\rangle, \dots, |f_{n-1}\rangle$  in order. Then, for any vector  $|\psi\rangle \in \mathbb{C}^n$ , its coordinates in the  $\{|f_k\rangle\}$  basis are exactly the elements of the vector

$$|\tilde{\psi}\rangle = U^\dagger |\psi\rangle.$$

That is,  $|\psi\rangle$  can be written as

$$|\psi\rangle = \sum_{k=0}^{n-1} \tilde{\psi}_k |f_k\rangle,$$

where  $\tilde{\psi}_k$ 's are elements of the vector  $|\tilde{\psi}\rangle = \begin{bmatrix} \tilde{\psi}_0 \\ \tilde{\psi}_1 \\ \vdots \\ \tilde{\psi}_{n-1} \end{bmatrix}$ .

---

<sup>4</sup>This is exactly the subspace  $\text{span}\{|a\rangle\} \subset \mathbb{C}^n$ .

**Worked example.** Consider the following orthonormal basis of  $\mathbb{C}^4$

$$|f_1\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad |f_2\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \quad |f_3\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad |f_4\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix},$$

and let  $U$  be a matrix whose columns are  $|f_1\rangle, \dots, |f_4\rangle$ .

Let  $|\psi\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = |0\rangle$ . Then

$$\tilde{\psi} = U^\dagger |0\rangle = \begin{bmatrix} \langle f_0 | \\ \langle f_1 | \\ \langle f_2 | \\ \langle f_3 | \end{bmatrix} |0\rangle = \begin{bmatrix} \langle f_0 | 0 \rangle \\ \langle f_1 | 0 \rangle \\ \langle f_2 | 0 \rangle \\ \langle f_3 | 0 \rangle \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Hence  $|0\rangle = \frac{1}{2} \sum_{k=0}^3 |f_k\rangle$ .

### Example 8: Expectation value of an “observable”

Let

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We define a value  $\mathbf{A}_\psi$  as follows:

$$\mathbf{A}_\psi = \langle \psi | \mathbf{A} | \psi \rangle.$$

This value is called the *expectation* of the *observable*  $\mathbf{A}$  in the state  $|\psi\rangle$ . Do not worry if the terms “expectation” and “observable” are not yet familiar: they arise from the theory of measurements of quantum states. We will introduce these concepts in later chapters.

For now, let us just compute the value  $\mathbf{A}_\psi$ .

**Solution.**

$$\langle \psi | = \frac{1}{\sqrt{2}} [1 \quad -i], \quad \mathbf{A} |\psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Thus

$$\mathbf{A}_\psi = \frac{1}{2} (1 \cdot i + (-i) \cdot 1) = \frac{1}{2}(i - i) = 0.$$

### Example 9: Matrix elements in a basis

Let  $\{|0\rangle, |1\rangle\}$  be the standard basis of  $\mathbb{C}^2$ . Given the operator

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

compute its matrix elements  $H_{ij} = \langle i | H | j \rangle$ , and verify  $H = \sum_{i,j \in \{0,1\}} H_{ij} |i\rangle\langle j|$ .

**Solution.**

$$H_{00} = \langle 0 | H | 0 \rangle = 1, \quad H_{01} = \langle 0 | H | 1 \rangle = 0, \quad H_{10} = \langle 1 | H | 0 \rangle = 0, \quad H_{11} = \langle 1 | H | 1 \rangle = -1.$$

Therefore

$$H = |0\rangle\langle 0| - |1\rangle\langle 1|.$$

---

**Example 10: Verifying orthonormality**

---

Let

$$|f_1\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad |f_2\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad |f_3\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ i \\ -i \end{bmatrix}, \quad |f_4\rangle = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -i \\ i \end{bmatrix}.$$

Show that  $\{|f_k\rangle\}_{k=1}^4$  is an orthonormal basis.

**Solution (sketch).** Each vector has norm 1 since each has four entries of magnitude 1/2. Pairwise inner products vanish by direct calculation; e.g.,

$$\langle f_3 | f_4 \rangle = \frac{1}{4} (1 \cdot 1 + (-1) \cdot (-1) + i \cdot (-i) + (-i) \cdot i) = \frac{1}{4} (1 + 1 + 1 - 1) = 0.$$

Thus the set is orthonormal and spans  $\mathbb{C}^4$ .

# Chapter 3

## Postulate (2/4): State Evolution

**Postulate 2:** The evolution of a closed quantum system is described by a unitary operator (aka matrix). Notation-wise,  $|\psi'\rangle = U|\psi\rangle$ , where  $U$  is a unitary operator (or transformation or matrix).

**Why unitaries?** The use of unitary operators follows from certain physics reasons: it follows directly from Schrödinger's equation. We will not derive this here; you can simply treat it as an axiom put by the Nature. Interested readers can consult the discussion in [NC10, Section 2.2.2].

### 3.1 Commonly Used Unitaries

#### The Pauli operators: bit flip $X$ , phase flip $Z$ , and $Y$

The Pauli matrices<sup>1</sup>  $\{X, Y, Z\}$  act on a single qubit (i.e., on  $\mathbb{C}^2$ ):

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

Key properties:

- All three are unitary:  $X^\dagger = X$ ,  $Y^\dagger = Y$ ,  $Z^\dagger = Z$ , and  $X^\dagger X = Y^\dagger Y = Z^\dagger Z = I$ .
- $X$  is the bit-flip:  $X|0\rangle = |1\rangle$ ,  $X|1\rangle = |0\rangle$ .
- $Z$  is the phase-flip:  $Z|0\rangle = |0\rangle$ ,  $Z|1\rangle = -|1\rangle$ .
- $Y$  combines a bit and phase flip:  $Y|0\rangle = i|1\rangle$ ,  $Y|1\rangle = -i|0\rangle$ .
- Eigenstructure:

$$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle; \quad X|\pm\rangle = \pm|\pm\rangle, \text{ where } |\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle).$$

- Commutation and multiplication:

$$XZ = -ZX = iY, \quad YZ = -ZY = iX, \quad XY = -YX = iZ, \quad X^2 = Y^2 = Z^2 = I.$$

---

<sup>1</sup>Named after the physicist Wolfgang Pauli (1900–1958), a pioneer of quantum mechanics best known for the Pauli exclusion principle. Pauli introduced these  $2 \times 2$  matrices to describe spin- $\frac{1}{2}$  systems; they generate the Lie algebra  $SU(2)$  and are central to qubit dynamics. He was awarded the 1945 Nobel Prize in Physics for the discovery of the exclusion principle.

## The Hadamard operator

The  $2 \times 2$  Hadamard operator<sup>2</sup>

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

is unitary. It maps the *computational* basis (another name for standard basis) to the following *Hadamard* basis

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

Check that:

$$H|0\rangle = |+\rangle, \quad H|1\rangle = |-\rangle, \quad H|+\rangle = |0\rangle, \quad H|-\rangle = |1\rangle.$$

It also interchanges  $X$  and  $Z$  by *conjugation* (i.e., multiply the target matrix from both the left and the right):

$$HZH = X, \quad HXH = Z.$$

Also note that  $H^2 = I$ .

## An interesting unitary on $\mathbb{C}^4$ : the 4-point Quantum Fourier Transform

Unitary matrices are not only of size  $2 \times 2$ . As a higher-dimension example, consider the so-called 4-point Quantum Fourier Transform (QFT) matrix  $F_4$  defined by:

$$F_4 := \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}.$$

This unitary maps the computational basis

$$\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$$

to the so-called “Fourier basis”

$$\{\tilde{|0\rangle}, \tilde{|1\rangle}, \tilde{|2\rangle}, \tilde{|3\rangle}\}$$

that are defined as follows:

$$\forall k \in \{0, 1, 2, 3\}, \quad \tilde{|k\rangle} := F_4 |k\rangle = \frac{1}{2} \sum_{j=0}^3 \omega^{jk} |j\rangle,$$

where  $\omega = e^{\frac{2\pi i}{4}}$  is the principle 4-th root of unity.

It is unitary since  $F_4^\dagger F_4 = I$ , and it plays a central role in quantum algorithms (e.g., phase estimation, period finding). We will return to this unitary later in the course. For now, we use it simply to illustrate a unitary operator acting on vectors in  $\mathbb{C}^n$  with  $n > 2$ .

---

<sup>2</sup>Named after the French mathematician Jacques Hadamard (1865–1963). Hadamard studied matrices with entries  $\pm 1$  whose rows are mutually orthogonal—now called Hadamard matrices—which maximize determinant among  $\pm 1$  matrices and appear in coding theory, design theory, and signal processing. In quantum computing, the  $2 \times 2$  Hadamard matrix  $H$  creates and interconverts computational and  $|\pm\rangle$  bases, serving as a basic “mixing” operation.

---

**There are many more unitaries**

---

Beyond these, there are numerous commonly used unitaries: phase gates ( $S, T$ ), rotations ( $R_x, R_y, R_z$ ), two-qubit gates such as CNOT, CZ, the SWAP gate, controlled-phase gates, and more. We will introduce them as needed throughout the course rather than all at once.

# Chapter 4

## Postulate (3/4): Measurement (Born Rule)

### 4.1 What a Quantum Measurement Means (Physical Intuition)

In addition to applying unitaries, we can also perform *measurements* on quantum states. Postulate 3, which we are about to introduce, fully characterizes how quantum measurements operate.

But before giving the mathematical formulation of Postulate 3, it helps to build physical intuition for what “measurement” means in quantum mechanics.

#### A measurement is an interaction with a device

A quantum measurement is not just “looking” at a system; it is a controlled physical interaction between the system (e.g., the qubit) and a macroscopic apparatus (detector, photodiode, superconducting resonator, etc.). During this interaction:

- information about some property of the system (referred to an *observable*) is transferred to a classical readout (a voltage pulse, a bright/dark spot, a pointer position),
- the system is generally disturbed by the interaction,
- the outcome is intrinsically probabilistic, even when the apparatus is ideal and free of engineering errors.

#### From everyday measurements to the microscopic world

We already know how measurements work in everyday life:

- **Length of a pencil:** a ruler gives a number (e.g., 12.3 cm).
- **Mass of a ball:** a scale reads off a weight.
- **Speed of a car:** a radar gun displays the speed.

In each case, you use a device that interacts with the object and produces a classical readout (a number on a display or a mark on a ruler).

**Key idea:** We can do the same in the microscopic world. We build devices that interact with tiny systems (like single electrons or atoms) and produce classical outcomes we can record. The overall logic:

$$\text{system} + \text{device} \Rightarrow \text{readout}$$

is the same.

#### Examples for microscopic measurements

Let us revisit the single-qubit systems we have introduced as concrete microscopic examples for measurements:

- **Electron spin (Stern–Gerlach–type idea).** An electron's spin can be “up” or “down” along a chosen axis (our qubit basis). To measure it, we let the electron pass through a region where the spin affects the path (conceptually like a Stern–Gerlach magnet) and place detectors at the two possible exit spots. One detector click = “spin up,” the other = “spin down.” This is directly analogous to the ruler/scale/gun giving a definite readout. If we repeat the experiment on many identically prepared electrons, the fraction of “up” vs. “down” clicks stabilizes to well-defined probabilities.
- **Atomic energy level (bright/dark detection).** Consider an atom that can be in a ground state  $|g\rangle$  or an excited state  $|e\rangle$  (another qubit). Shine a laser that makes  $|e\rangle$  fluoresce. If the atom is in  $|e\rangle$ , we see photons (bright); if it is in  $|g\rangle$ , we see none (dark). Bright vs. dark is the classical outcome shown by the detector, just like a scale showing a number. Repeating on many atoms prepared the same way yields stable bright/dark frequencies.

### **What's similar?**

- There is always a *device* that interacts with the system and produces a *classical readout* (click/no click, bright/dark, 0/1).
- Repeating the same procedure on identically prepared systems gives *stable statistics* for the outcomes.

### **What's different (specifically quantum)?**

- **Intrinsic randomness:** Even with a perfect device, a single microscopic outcome can be random. The probabilities come from the quantum state, not from hidden imperfections of the apparatus.
- **State update (“collapse”):** After you read “spin up,” the electron’s state is now aligned with “up.” After a bright detection, the atom is effectively in  $|e\rangle$  (or quickly pumped into  $|g\rangle$  depending on the scheme). The measurement changes what the system will do next.
- **Choice of question matters:** Measuring spin along  $Z$  (up/down) answers a different question than measuring along  $X$  (left/right). Answering one precisely disturbs the other. This basis-dependence has no direct everyday analog for length, mass, or speed measured at one moment.

### **Summary**

The above examples prepare us for the formal Postulate 3, also known as the Born rule<sup>1</sup>. It is a precise recipe that connects the quantum state, the chosen measurement (the “question”), and the probabilities of each classical outcome.

---

<sup>1</sup>Named after the German physicist Max Born (1882–1970), a founder of quantum mechanics. Born formulated the probabilistic interpretation of the wavefunction—now called the Born rule—identifying  $|\psi(x)|^2$  (or, in finite dimensions,  $|\langle\phi|\psi\rangle|^2$ ) as a probability density (or probability) for measurement outcomes. For this fundamental contribution to quantum theory, Born was awarded the 1954 Nobel Prize in Physics.

## 4.2 Postulate 3: the Formal Statement

**Postulate 3: Measurement (Born Rule).** Besides applying unitaries, we can perform *measurements* on quantum states.

Quantum measurements obey the following rules:

- **Measurement specification:** a quantum measurement is specified by a collection  $\{M_m\}$  of *measurement operators* satisfying the *completeness* condition:

$$\sum_m M_m^\dagger M_m = I. \quad (4.1)$$

These operators acting on the state space of the system being measured. The index  $m$  refers to the measurement outcomes that may occur in the experiment.

- **Measurement outcome:** if the state of the quantum system is  $|\psi\rangle$  immediately before the measurement, then the probability that result  $m$  occurs is given by

$$\Pr[\text{outcome} = m] = \langle \psi | M_m^\dagger M_m | \psi \rangle, \quad (4.2)$$

- **Post-measurement state:** the state of the system after the measurement is

$$\frac{M_m |\psi\rangle}{\|M_m |\psi\rangle\|} \quad \left( = \frac{M_m |\psi\rangle}{\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}} \right). \quad (4.3)$$

Remarks regarding this postulate:

- We emphasize that  $\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}$  is just  $\|M_m |\psi\rangle\|$ , where  $\|\cdot\|$  is the inner-product induced norm.

Thus, the probability is the inner product between the vector  $M_m |\psi\rangle$  and itself. Note that this probability must be positive. Think why.

- The completeness equation [Equation \(4.1\)](#) expresses the fact that probabilities sum to one. Prove it during the lecture.
- Also note that the post-measurement state is really a valid quantum state. That is, [Expression \(4.3\)](#) is consistent with Postulate 1, the condition for being a valid quantum state. Prove it during the lecture.

## 4.3 Example: Measuring a Single Qubit

We now illustrate the Born rule with concrete single-qubit measurements, written in the general  $\{M_m\}$  (measurement-operator) framework we just learned.

### 4.3.1 Case 1: measurement in the computational ( $Z$ ) basis

Let the qubit be  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  with  $|\alpha|^2 + |\beta|^2 = 1$ . Specify the measurement by two operators

$$M_0 = |0\rangle\langle 0|, \quad M_1 = |1\rangle\langle 1|.$$

They satisfy completeness:

$$M_0^\dagger M_0 + M_1^\dagger M_1 = |0\rangle\langle 0| + |1\rangle\langle 1| = I.$$

Born-rule probabilities:

$$\Pr[0] = \langle\psi|M_0^\dagger M_0|\psi\rangle = |\alpha|^2, \quad \Pr[1] = \langle\psi|M_1^\dagger M_1|\psi\rangle = |\beta|^2.$$

Post-measurement states (conditioned on the outcome), where we use  $|\psi_b\rangle$  to denote the post-measurement states corresponding to outcome  $b$  ( $b \in \{0, 1\}$ ):

$$|\psi_0\rangle = \frac{M_0|\psi\rangle}{\|M_0|\psi\rangle\|} = |0\rangle, \quad |\psi_1\rangle = \frac{M_1|\psi\rangle}{\|M_1|\psi\rangle\|} = |1\rangle.$$

In summary, the measurement results are stated as the following:

- with probability  $|\alpha|^2$ , we observe the measurement outcome 0, and the state collapses to  $|0\rangle$ ;
- with probability  $|\beta|^2$ , we observe the measurement outcome 1, and the state collapses to  $|1\rangle$ ;

#### Geometric picture in the 2D computational plane

For this simple measurement of a single qubit, we have an elegant geometric picture for it. See [Figure 4.1](#).

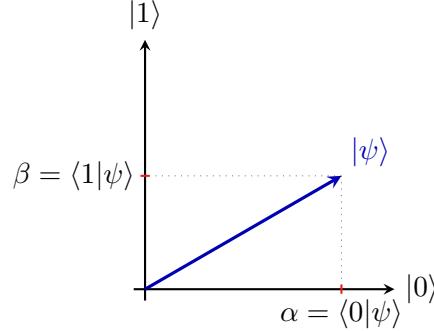


Figure 4.1: Geometric picture in the 2D space

Consider the 2D vector space spanned by  $\{|0\rangle, |1\rangle\}$ . Write the state as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Think of  $|0\rangle$  and  $|1\rangle$  as two orthonormal axes in this plane. The number  $\alpha = \langle 0|\psi\rangle$  is the component of  $|\psi\rangle$  along the  $|0\rangle$  axis, and  $\beta = \langle 1|\psi\rangle$  is the component along the  $|1\rangle$  axis. The measurement “collapses” onto one axis, with probabilities given by the squared lengths of these components:

$$\Pr[0] = |\alpha|^2 = |\langle 0|\psi\rangle|^2, \quad \Pr[1] = |\beta|^2 = |\langle 1|\psi\rangle|^2.$$

If  $\alpha$  and  $\beta$  are real and nonnegative, you can visualize the state vector forming an angle  $\theta$  with the  $|0\rangle$  axis so that  $\alpha = \cos \theta$  and  $\beta = \sin \theta$ ; then the probabilities are the squared projections onto each axis,  $\cos^2 \theta$  and  $\sin^2 \theta$ . In the general complex case, the same idea holds: you still project onto the two orthonormal axes and take the squared magnitudes of those projections to get the outcome probabilities.

### On the naming

In the 2D picture above, this measurement is a projection of the state onto the orthonormal axes  $\{|0\rangle, |1\rangle\}$ . For this reason, it is called a measurement in the computational basis (or simply a computational measurement).

Furthermore, the Pauli  $Z$  operator has  $|0\rangle$  and  $|1\rangle$  as eigenstates (i.e., eigenvectors) with eigenvalues  $+1$  and  $-1$ . Thus, measuring in  $\{|0\rangle, |1\rangle\}$  is equivalently a projection to the eigenstates of  $Z$ . This is why it is also referred to as measurement in the  $Z$  basis.

Later in the course, we will explore in detail how measurements relate to the eigenstates of a special class of matrices called *observables*. For now, you need not worry about these details. Interested readers can consult [Appendix A](#).

#### 4.3.2 Case 2: measurement in the Hadamard ( $X$ ) basis

Let the qubit be  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  with  $|\alpha|^2 + |\beta|^2 = 1$ . Define the Hadamard-basis ( $X$ -basis) states

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

Specify the measurement by two operators

$$M_+ = |+\rangle\langle +|, \quad M_- = |-\rangle\langle -|.$$

They satisfy completeness:

$$M_+^\dagger M_+ + M_-^\dagger M_- = |+\rangle\langle +| + |-\rangle\langle -| = I.$$

Born-rule probabilities (expressing  $|\psi\rangle$  in the  $\{|+\rangle, |-\rangle\}$  basis):

$$\langle +|\psi\rangle = \frac{\alpha+\beta}{\sqrt{2}}, \quad \langle -|\psi\rangle = \frac{\alpha-\beta}{\sqrt{2}},$$

$$\Pr[+] = \langle \psi | M_+ | \psi \rangle = \left| \frac{\alpha+\beta}{\sqrt{2}} \right|^2, \quad \Pr[-] = \langle \psi | M_- | \psi \rangle = \left| \frac{\alpha-\beta}{\sqrt{2}} \right|^2.$$

Post-measurement states (conditioned on the outcome), using  $|\psi_b\rangle$  for  $b \in \{+, -\}$ :

$$|\psi_+\rangle = \frac{M_+ |\psi\rangle}{\|M_+ |\psi\rangle\|} = |+\rangle, \quad |\psi_-\rangle = \frac{M_- |\psi\rangle}{\|M_- |\psi\rangle\|} = |-\rangle.$$

In summary, the measurement results are:

- with probability  $\left| \frac{\alpha+\beta}{\sqrt{2}} \right|^2$ , we observe the outcome “+” and the state collapses to  $|+\rangle$ ;
- with probability  $\left| \frac{\alpha-\beta}{\sqrt{2}} \right|^2$ , we observe the outcome “-” and the state collapses to  $|-\rangle$ .

## **Geometric picture in the same 2D plane**

Remain in the 2D vector space spanned by  $\{|0\rangle, |1\rangle\}$ , but now view  $\{|+\rangle, |-\rangle\}$  as an orthonormal pair of axes obtained by a  $45^\circ$  rotation of the computational axes. The components of  $|\psi\rangle$  along these axes are  $\langle +|\psi\rangle = (\alpha + \beta)/\sqrt{2}$  and  $\langle -|\psi\rangle = (\alpha - \beta)/\sqrt{2}$ . The measurement collapses onto one of these rotated axes with probabilities given by the squared magnitudes of the corresponding projections:

$$\Pr[+] = |\langle +|\psi\rangle|^2, \quad \Pr[-] = |\langle -|\psi\rangle|^2.$$

## **On the naming**

This measurement is in the Hadamard or  $X$  basis because the Pauli  $X$  operator has  $|+\rangle$  and  $|-\rangle$  as eigenstates with eigenvalues  $+1$  and  $-1$ :  $X|+\rangle = +|+\rangle$ ,  $X|-\rangle = -|-\rangle$ . Equivalently, applying a Hadamard gate  $H$  maps the computational basis to the  $X$  basis ( $H|0\rangle = |+\rangle$ ,  $H|1\rangle = |-\rangle$ ), so “measure in  $X$ ” is the same procedure as “apply  $H$ , then measure in  $Z$ ”.

# Bibliography

- [NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, USA, 10th edition, 2010. [39](#)

# Appendix A

## From matrices to measurable properties: why $Z$ represents an observable

This is our first concrete example in which a matrix—the Pauli  $Z$ —stands for a physical property that one can actually measure. Here is the conceptual bridge:

- **States as vectors.** A qubit's state is a unit vector  $|\psi\rangle$  in a two-dimensional complex vector space spanned by  $\{|0\rangle, |1\rangle\}$ .
- **Physical questions as operators.** A *yes/no* or *which-of-two* question about the system (e.g., “is the spin up or down along the  $Z$  direction?” or “is the qubit in  $|0\rangle$  or  $|1\rangle$ ?”) is encoded by a Hermitian operator  $A$  acting on that same space. Hermitian matrices have real eigenvalues and an orthonormal eigenbasis—exactly what we need to define possible outcomes and the states we collapse to.
- **Eigenvalues = possible outcomes; eigenvectors = post-measurement states.** For a Hermitian operator  $A$ , the spectral theorem says

$$A = \sum_m a_m \Pi_m, \quad \Pi_m = |a_m\rangle\langle a_m|,$$

where  $a_m$  are real eigenvalues (the *numbers* you can read on a device) and  $\Pi_m$  are projectors onto the corresponding eigenvectors  $|a_m\rangle$  (the *states* you end up in after getting outcome  $a_m$ ).

- **Born rule in observable language.** If the system is  $|\psi\rangle$ , then

$$\Pr[\text{outcome } a_m] = \langle \psi | \Pi_m | \psi \rangle \quad \text{and} \quad |\psi | a_m \rangle = \frac{\Pi_m |\psi\rangle}{\sqrt{\langle \psi | \Pi_m | \psi \rangle}}.$$

The operator  $A$  thus determines both the *set of outcomes* and the *state update rule*.

### Specializing to the Pauli $Z$

The Pauli  $Z$  matrix is

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

with eigenpairs

$$Z|0\rangle = (+1)|0\rangle, \quad Z|1\rangle = (-1)|1\rangle.$$

Thus:

- The *possible outcomes* are the real numbers  $+1$  and  $-1$ .
- The *associated projectors* are  $\Pi_+ = |0\rangle\langle 0|$  and  $\Pi_- = |1\rangle\langle 1|$ .

- Measuring the observable  $Z$  is the same operational procedure as measuring in the  $\{|0\rangle, |1\rangle\}$  basis: outcome  $+1$  corresponds to reporting “ $|0\rangle$ ”, and outcome  $-1$  corresponds to “ $|1\rangle$ ”.

### **Why this is physically meaningful**

A matrix like  $Z$  is not merely a computational device; it *labels a concrete experimental question*:

- It specifies which directions in state space are the “definite answers” (its eigenvectors).
- It determines what numbers are read out on the meter (its eigenvalues).
- Through the projectors  $\{\Pi_m\}$ , it fixes the Born-rule probabilities and the post-measurement states, connecting linear algebra to experimental statistics and dynamics.

### **Computational vs. physical labels**

In quantum computing, we often report the outcome as a classical bit  $0/1$ . In the observable viewpoint, the same experiment can be labeled by  $+1/-1$  (the eigenvalues of  $Z$ ). These are two notational skins over the same process:

$$\text{“bit} = 0\text{”} \equiv \text{“eigenvalue } +1 \text{ of } Z\text{”} \equiv \text{project onto } |0\rangle,$$

$$\text{“bit} = 1\text{”} \equiv \text{“eigenvalue } -1 \text{ of } Z\text{”} \equiv \text{project onto } |1\rangle.$$

Seeing  $Z$  as an observable crystallizes the idea that *matrices encode measurable questions*, with their eigenstructures defining what can be observed and how the state changes when we observe it.