Quantum Computing
An introduction

Mathematical foundations • Quantum mechanics • Information in QM
Quantum gates Programming • Quantum Entanglement • Quantum Protocols
Quantum Algorithms • Hybrid Quantum-classical algorithms • Hardware

Quantum Computing
An introduction

Mathematical foundations • Quantum mechanics • Information in QM
Quantum gates Programming • Quantum Entanglement • Quantum Protocols
Quantum Algorithms • Hybrid Quantum-classical algorithms • Hardware

Alcides Montoya Cañola

Associate Professor, Department of Physics,
Faculty of Sciences, Universidad Nacional de Colombia,
Sede Medellín

UNIVERSIDAD
NACIONAL
DE COLOMBIA

Medellín, Colombia, 2026

004.1
M79 Montoya Cañola, Alcides
Quantum Computing: An introduction /
Montoya. – First edition. – Medellín, Colombia : Universidad
Nacional
de Colombia. Faculty of Sciences, 2026.
1 online resource (XXX pages) : illustrations, diagrams

ISBN: XXXXX

1. QUANTUM COMPUTING 2. QUANTUM MECHANICS
3. QUANTUM ALGORITHMS 4. QUANTUM INFORMATION
I. Montoya Cañola, Alcides
II. Title.

Cataloging-in-Publication Data, Universidad Nacional de Colombia. Sede Medellín

Quantum Computing: An introduction

Editorial Center of the Faculty of Sciences

Digital ISBN: XXXXX
Medellín, Colombia
First edition: August 2026

Cover design: XXXXX

Printed and made in Medellín, Colombia.

To all the women who made science
in laboratories where no chair awaited them,
who persevered against closed doors,
unsigned papers and unspoken names,
and built, step by step,
the luminous path we walk today.

To Marie Curie and Lise Meitner;
to Emmy Noether, who revealed that symmetry
is the deepest grammar of the universe;
to Grete Hermann, who dared to question
a proof that no one else questioned;
to Maria Goeppert Mayer and Chien-Shiung Wu;
to Ada Lovelace and Grace Hopper,
who first imagined that a machine could think.

And to those who are writing the present:
Dorit Aharonov, Barbara Terhal,
Michelle Simmons, Stephanie Wehner,
Eleanor Rieffel, Krysta Svore, Maria Schuld,
and Ana María Rey, who carried the light
of these same mountains to the coldest atoms.

Acknowledgments

My gratitude goes first to the Universidad Nacional de Colombia, Sede Medellín, for providing the academic space where these ideas could germinate and grow. This institution has been the garden where the seed of this project found fertile soil to flourish.

I am especially grateful to the **Centro de Excelencia en Computación Cuántica e Inteligencia Artificial** of the Universidad Nacional de Colombia, a space where quantum computing and artificial intelligence meet, and where the conviction that these two fields must be taught together has taken shape. Much of what appears in these pages was first discussed, questioned, and refined there.

To Professor Mario Trujillo Vargas, whose teaching has transformed the lives of his students. He taught us that physics is not a collection of formulas to be memorized but a way of looking at the world, and that a teacher's true measure is not what he knows, but what he awakens in others. His vision continues to shape those of us who had the fortune of learning from him.

To the young minds of Engineering Physics, who with their genuine questions and tireless curiosity have shaped every page of this text. Their doubts became clarities, their difficulties became opportunities to explain things better. This book is, in reality, an extended dialogue with all those restless souls seeking to understand.

Finally, to the students of the quantum computing course, who read these pages with a critical eye and helped correct the errors that inevitably inhabit a work of this length, and to all those who have helped and will help to improve it: you are a fundamental part of this collective process of building knowledge. In every correction, in every suggestion, lives the spirit of collaboration that makes the advancement of science possible.

Contents

Chapter 1

Mathematical Foundations

The applications explored in Chapter 1—from simulating molecular systems and optimizing financial portfolios to breaking cryptographic codes—reveal quantum computing's transformative potential. Yet realizing these applications requires mastering a mathematical framework fundamentally different from classical computing. Where classical bits occupy definite states (0 or 1), quantum bits exist in superpositions of states, described by complex probability amplitudes rather than simple probabilities. Where classical information processing follows deterministic Boolean logic, quantum operations are represented by unitary transformations in complex vector spaces. Understanding this quantum formalism is not merely a mathematical exercise but the essential foundation for implementing quantum algorithms and interpreting their results.

This chapter develops the mathematical machinery necessary for quantum computing, progressing systematically from elementary concepts to the structures of quantum mechanics. We begin with state representation through vectors, introducing complex numbers and their role in describing quantum amplitudes. Linear algebra provides the core framework: vector spaces, inner products, and linear operators become the language for expressing quantum states and their transformations. Dirac notation, developed by Paul Dirac specifically for quantum mechanics, offers elegant shorthand that simplifies calculations and clarifies physical meaning. The Pauli matrices and the Stern–Gerlach experiment illustrate how abstract mathematics connects directly to measurable physical phenomena. Eigenvalues and eigenvectors reveal the possible outcomes of quantum measurements, while Hermitian operators ensure these outcomes are physically meaningful real numbers. Finally, we introduce Hilbert space, the mathematical arena where quantum mechanics achieves full rigor.

The mathematical structures presented here are not arbitrary abstractions imposed on quantum physics but emerge naturally from experimental observations. The Stern–Gerlach experiment, conducted in 1922 with silver atoms, revealed that the magnetic moment of an atom is quantized into discrete values—a phenomenon later understood in terms of spin and described by 2×2 matrices, the Pauli matrices that appear throughout quantum computing. Complex probability amplitudes, rather than classical probabilities, are necessary to explain interference patterns in double-slit experiments. Unitary evolution reflects the conservation of total probability and the reversible nature of quantum dynamics. Each mathematical structure has physical justification, connecting abstract formalism to laboratory reality.

Readers already familiar with linear algebra will recognize many concepts but encounter them in new contexts: matrices represent not just coordinate transformations but quantum gates; eigenvalues correspond not merely to scaling factors but to measurement outcomes; tensor products describe not geometric combinations but composite—and possibly entangled—quantum systems. Those approaching linear algebra for the first time should work carefully through definitions and examples, as every concept introduced here will be used extensively in subsequent chapters.

1.1 From Bits to Quantum States

The fundamental building block of computation is the two-state system, known as a bit. In computer science, the 'bit' is the basic unit of information, allowing us to discern between two states: 0 or 1. Bits are grouped to form binary words; for example, an 8-bit word, called a byte, can represent up to 256 distinct states.

Consider a coin with two sides, commonly called 'heads' and 'tails'. When flipping the coin, there are two possibilities: landing on the heads side or landing on the tails side. While the coin is in the air, we cannot determine with certainty which side it will land on; for a fair coin, each outcome has probability $\frac{1}{2}$.

We can encode the two definite outcomes of the system as vectors: the 'heads' state as $\mathbf{heads} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the 'tails' state as $\mathbf{tails} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Our incomplete knowledge about the coin in the air can then be summarized by a *probability vector*,

$$\frac{1}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \tag{1.1}$$

whose components are the probabilities of the two outcomes and therefore sum to one. It is essential to be precise about what this vector represents: a *classical statistical mixture*. The coin is always in one definite state—heads or tails—and the vector merely encodes our ignorance about which one. This is *not* a quantum superposition. A quantum superposition, as we will see, is a genuinely new kind of state in which the system does not have a definite value of the measured property prior to measurement, and whose components are complex *amplitudes* rather than probabilities. Keeping the distinction between classical mixtures (probabilities that sum to one) and quantum superpositions (amplitudes whose squared moduli sum to one) firmly in mind is one of the most important conceptual habits the reader can develop; confusing the two is perhaps the single most common error in popular accounts of quantum computing.

Why are classical probabilities not enough to describe nature at the microscopic scale? Consider an electron passing through a double-slit apparatus: if we do not observe which slit it traverses, an interference pattern emerges on the detection screen—alternating bands of high and low probability that cannot be explained by classical probability theory. Amplitudes for the two paths can cancel, producing zero probability at locations where classical reasoning predicts the electron should appear. This destructive interference requires probability amplitudes that can be negative or, more generally, complex-valued, allowing cancellation through destructive interference and reinforcement through constructive interference.

Classical probabilities are non-negative real numbers that simply add: if two mutually exclusive alternatives have probabilities p_A and p_B, the probability of one or the other is $p_A + p_B$. Quantum mechanics, by contrast, assigns each possible outcome a complex probability amplitude. These amplitudes can interfere—adding constructively when their phases align or destructively when phases oppose—before being squared to yield observable probabilities. This amplitude-based description is not a mathematical convenience but a necessity imposed by experimental observations of interference phenomena that pervade quantum physics.

1.1.1 Complex Amplitudes and the Born Rule

The transition from classical probability to quantum probability amplitudes represents one of the conceptual leaps distinguishing quantum mechanics from classical physics. In quantum theory, the state of a system is described by a wavefunction ψ, which assigns a complex number—a probability amplitude—to each possible measurement outcome. For a particle in three-dimensional space, the wavefunction $\psi(x, y, z)$ depends on position coordinates and contains complete information about the particle's spatial state.

The connection between these abstract complex amplitudes and measurable probabilities is provided by the Born rule, formulated by Max Born in 1926 [3]. This principle establishes that the probability density of finding a particle at position (x, y, z) is the squared modulus of the wavefunction at that point:

$$P(x, y, z) = |\psi(x, y, z)|^2 = \psi^*(x, y, z)\, \psi(x, y, z) \tag{1.2}$$

Here ψ^* denotes the complex conjugate of ψ. The squared modulus $|\psi|^2$ is always a non-negative real number, as required for a probability, even though ψ itself may be complex. For the total probability of finding the particle somewhere in space to equal one, the wavefunction must satisfy the *normalization condition*:

$$\int_{\mathbb{R}^3} |\psi(x, y, z)|^2\, dx\, dy\, dz = 1. \tag{1.3}$$

This seemingly simple formula has profound implications: quantum mechanics is fundamentally probabilistic, not due to incomplete knowledge but as an intrinsic feature of nature. Born's proposal—that the wavefunction does not directly describe physical reality, but rather the probability of finding the particle in a particular configuration—profoundly transformed twentieth-century physics and earned him the Nobel Prize in Physics in 1954. In his own words: *"The motion of particles follows the laws of probability, but probability itself propagates according to the law of causality"* [3].

The Born rule also reveals why complex numbers are essential in quantum mechanics. If we denote the amplitude for an electron to reach point x via slit 1 as $\psi_1(x)$ and via slit 2 as $\psi_2(x)$, the total amplitude is $\psi(x) = \psi_1(x) + \psi_2(x)$, and the observed probability is:

$$P(x) = |\psi_1(x) + \psi_2(x)|^2 = |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\,\mathrm{Re}[\psi_1^*(x)\psi_2(x)] \tag{1.4}$$

The final term, absent in classical probability, represents quantum interference. It can be positive (constructive interference) or negative (destructive interference), producing the characteristic bright and dark fringes of the interference pattern. Complex amplitudes enable this interference through their phase relationships, making them not merely convenient but necessary for describing quantum phenomena.

Understanding the Born rule is essential for quantum computing. A qubit in state $\alpha\,|0\rangle + \beta\,|1\rangle$, where α and β are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$, will be measured as $|0\rangle$ with probability $|\alpha|^2$ and as $|1\rangle$ with probability $|\beta|^2$. The complex phases of α and β, though not directly observable in a single measurement, determine how qubits interfere during quantum computations, enabling quantum algorithms to amplify correct answers and suppress incorrect ones through carefully orchestrated interference.

1.1.2 Complex Numbers

The history of the quadratic equation and its relationship with complex numbers has its roots in ancient Arabic mathematics. The general method for solving quadratic equations, such as $x^2 - 10x + 34 = 0$, has its origin in the works of **Muhammad ibn Musa al-Khwarizmi (c. 780 – c. 850)**, who in his work **Al-Kitab al-Mukhtasar fi Hisab al-Jabr wal-Muqabala** developed systematic methods for solving quadratic equations. This work, which introduced the concept of **al-jabr** (from which the word "algebra" derives), was fundamental for the development of algebraic theory in the Islamic world and later in Europe, after being translated in the 12th century [15].

To solve a quadratic equation of the form

$$x^2 - 10x + 34 = 0, \tag{1.5}$$

the well-known quadratic formula is used:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{1.6}$$

Applying it to our equation, the possible values of x are:

$$x = \frac{-(-10) \pm \sqrt{(-10)^2 - 4(1)(34)}}{2(1)} = \frac{10 \pm \sqrt{100 - 136}}{2} = \frac{10 \pm \sqrt{-36}}{2} \tag{1.7}$$

The question is: how do we interpret the quantity $\sqrt{-36}$? There is no real number whose square is negative; thus, this equation has no real solutions. However, we can write $\sqrt{-36} = \sqrt{36 \cdot (-1)} = 6\sqrt{-1}$. If a new entity i is introduced, defined by the property $i^2 = -1$, the equation becomes solvable: the *imaginary unit* is defined as $i = \sqrt{-1}$, so that $\sqrt{-36} = 6i$ and the two solutions of the quadratic are $x = 5 \pm 3i$.

René Descartes (1596–1650) introduced the term **imaginary** to refer to the roots of negative numbers, such as $\sqrt{-36}$. In his work **La Géométrie** (1637), Descartes used this denomination because he considered that such quantities did not have a tangible existence, unlike "real" numbers. Due to this terminology, $\sqrt{-1}$ is known as the imaginary unit, represented by i, and it became a fundamental element for the development of complex numbers, significantly expanding the scope of mathematical analysis [4].

The complex numbers $\mathbb{C}$ will be the numerical field on which we build the state space, and from there we will proceed to Hilbert space, which is a complex vector space [12]. In this section we review complex numbers and their basic properties. To begin, let us define the complex plane and represent this set of numbers there.

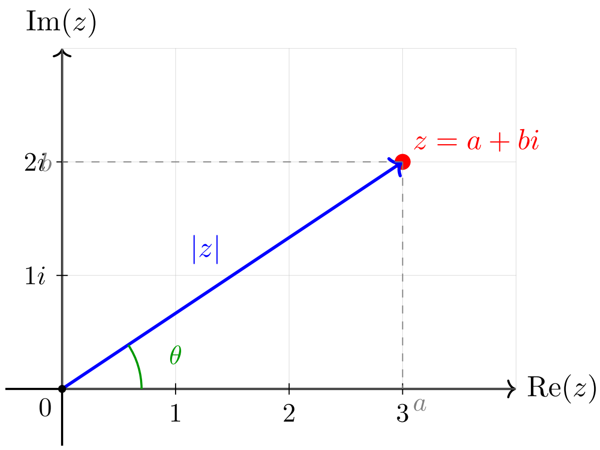

Figure 1.1: Representation of a complex number $z = a + bi$ in the complex plane $\mathbb{C}$. The number is represented as a point (or vector from the origin) with real part a and imaginary part b. The modulus $|z| = \sqrt{a^2 + b^2}$ is the distance from the origin, and the argument $\theta = \arg(z)$ is the angle with the positive real axis. In polar form: $z = |z|e^{i\theta} = |z|(\cos\theta + i\sin\theta)$.

A complex number can be expressed in terms of its Cartesian components in the plane as

$$z = a + ib, \tag{1.8}$$

or in terms of its polar coordinates as

$$z = r(\cos\theta + i\sin\theta), \tag{1.9}$$

which, using Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$, can be written compactly as

$$z = re^{i\theta}. \tag{1.10}$$

For the purposes of this text, the basic operations with complex numbers are:

- Addition and subtraction:

$$(a + ib) \pm (c + id) = (a \pm c) + i(b \pm d) \tag{1.11}$$

- Multiplication:

$$(a+ib)(c+id) = (ac-bd)+i(ad+bc), \qquad (re^{i\theta})(r'e^{i\theta'}) = rr'e^{i(\theta+\theta')} \tag{1.12}$$

- Complex conjugate: $z^* = \overline{z} = a - ib = re^{-i\theta}$
- Absolute value (modulus): $|z| = \sqrt{a^2 + b^2} = r$, with $|z_1 z_2| = |z_1||z_2|$
- Absolute value squared: $|z|^2 = a^2+b^2 = r^2$; an important relation: $|z|^2 = z\overline{z}$
- Inverse: $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$ for $z \neq 0$
- De Moivre's formula: $z^n = |z|^n[\cos(n\theta) + i\sin(n\theta)]$

These basic elements of complex numbers will be sufficient for our development of quantum computing.

1.1.3 Probabilities and Measurements

Einstein, in his criticism of the interpretation of quantum mechanics, stated that "God does not play dice" [2]. Although this phrase is famous for expressing his skepticism toward the probabilistic nature of quantum mechanics, it is important to clarify that probability in this field is not equivalent to the classical probability found in games of chance like dice. Quantum probabilities arise from complex-valued amplitudes, and the observables of the theory are represented by operators that, in general, do not commute with one another. This non-commutativity of observables—absent in classical physics—is one of the features that fundamentally distinguishes quantum mechanics, and it implies, among other things, that quantum probabilities do not always combine according to the classical rules.

The foundations of classical probability theory have their origins in the 17th century, when mathematicians such as **Pierre de Fermat** and **Blaise Pascal** began to analyze problems related to games of chance. Their correspondence in 1654 about the *division of stakes* in interrupted games marked a milestone in the development of this discipline, laying the foundations for what we now know as **classical probability** [9].

Classical probability is based on the following fundamental principles:

1. **Sample space (Ω)**: A non-empty set that includes all possible outcomes of a random experiment.
2. **Events ($E \subset \Omega$)**: Subsets of the sample space that represent specific outcomes or combinations of outcomes.
3. **Probability function (P)**: A function that assigns to each event E a numerical value between 0 and 1 ($0 \leqslant P(E) \leqslant 1$), where $P(E)$ represents the probability that event E occurs.

For this function to be valid, it must satisfy the following fundamental properties:

1. $P(\emptyset) = 0$, indicating that the probability of an impossible event is zero.
2. $P(\Omega) = 1$, since some outcome in the sample space must occur with certainty.
3. If E_1 and E_2 are disjoint events ($E_1 \cap E_2 = \emptyset$), then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$. This property, known as additivity, is essential for calculating probabilities of compound events.

The complete formalization of probability theory occurred almost three centuries later, in **1933**, when **Andrey Kolmogorov** published his influential work *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Foundations of the Theory of Probability*). Kolmogorov introduced an axiomatic approach that extended classical ideas to encompass both finite and infinite sample spaces, providing the rigorous basis that today underpins all modern applications of probability, from statistics to quantum physics [13].

In quantum mechanics, the **probability amplitude** is a complex number that encapsulates information about the state of a quantum system. Its interpretation, proposed by Max Born in 1926 and discussed in the previous subsection, establishes that the squared modulus of this amplitude corresponds to a probability (or probability density, for continuous variables). This concept is one of the pillars of the **Copenhagen interpretation**, which emphasizes the inherent probabilistic nature of quantum phenomena and the impossibility of predicting individual outcomes with certainty.

In this framework, the state of any quantum system is described by a vector $|\psi\rangle$ in an abstract complex vector space, known as **Hilbert space**. This space provides the mathematical setting where linear operators represent physical observables, and the temporal evolution of the system is governed by the **Schrödinger equation**, which describes how $|\psi\rangle$ changes as a function of time. Although this equation is deterministic in its mathematical form, predictions about the outcomes of individual measurements are intrinsically probabilistic.

The probabilistic interpretation was not accepted without controversy. Physicists of the stature of Einstein, Schrödinger, and de Broglie questioned whether nature could be fundamentally probabilistic, and proposed that the apparent randomness might reflect our ignorance of underlying *hidden variables*. This deep and historically rich debate—from de Broglie's pilot wave and Bohm's deterministic reformulation to Bell's theorem and its experimental verification—is presented in detail in Chapter 3, where the postulates of quantum mechanics provide the proper context for it. For now, we note only its conclusion: no local hidden-variable theory can reproduce all the predictions of quantum mechanics, and the confirmed non-locality of quantum correlations is precisely what makes entanglement a genuinely non-classical computational resource.

In this book we adopt the matrix formulation of quantum mechanics, and therefore we will develop the fundamental concepts of linear algebra. Readers already fluent in linear algebra may skim the next section and continue with Section 1.4.

1.2 Review of Linear Algebra Concepts

Vector spaces and linear operators are fundamental concepts in linear algebra. In quantum mechanics, the states of a quantum system are represented by vectors in a complex vector space endowed with an inner product—a Hilbert space. The transformations that the system undergoes, such as operations on quantum states, are described by linear operators, which act on these vectors according to the laws of quantum mechanics.

1.2.1 Matrices

We denote the set of $n \times m$ complex matrices by $\mathbf{M}_{n,m}(\mathbb{C})$; a generic element has the form:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{bmatrix} \tag{1.13}$$

We write $A_{ij} = a_{ij}$ for the entry of matrix A in row i and column j. We now review the basic operations with matrices required in quantum computing.

To add two or more matrices they must have the same size, that is, the same number of rows and columns. Similarly, matrices can be multiplied by a scalar. If $A, B \in \mathbf{M}_{n,m}(\mathbb{C})$ and $c \in \mathbb{C}$, then:

$$(A + B)_{ij} = A_{ij} + B_{ij} \tag{1.14}$$

and for the scalar multiple:

$$(cA)_{ij} = c\,A_{ij} \tag{1.15}$$

As examples we have:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix} \tag{1.16}$$

$$c\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{bmatrix} \tag{1.17}$$

Conjugate Transpose of a Matrix

For an element $A \in \mathbf{M}_{n,m}(\mathbb{C})$ we define the following operations:

- Complex conjugate: A^* (also written $\overline{A}$). The conjugate of a complex matrix is obtained by taking the complex conjugate of each entry:

$$(A^*)_{ij} = (A_{ij})^* \tag{1.18}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}^* = \begin{bmatrix} a_{11}^* & a_{12}^* & a_{13}^* \\ a_{21}^* & a_{22}^* & a_{23}^* \end{bmatrix} \tag{1.19}$$

  An important property of this operation is that it is an involution: $(A^*)^* = A$, i.e., the conjugate of the conjugate is the original matrix.

- Transpose: A^T. It is the matrix that results from exchanging rows with columns:

$$(A^T)_{ij} = A_{ji} \tag{1.20}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} \tag{1.21}$$

  The transpose has the following properties:

$$\begin{aligned} (\mathbf{M}^T)^T &= \mathbf{M} \\ (\mathbf{M}+\mathbf{N})^T &= \mathbf{M}^T + \mathbf{N}^T \\ (\mathbf{M}\mathbf{N})^T &= \mathbf{N}^T\mathbf{M}^T \\ (k\mathbf{M})^T &= k\mathbf{M}^T \end{aligned} \tag{1.22}$$

  If $\mathbf{M}^T = \mathbf{M}$ then $\mathbf{M}$ is called a symmetric matrix. If $\mathbf{M}^T = -\mathbf{M}$ then $\mathbf{M}$ is called an antisymmetric matrix.

- Conjugate transpose (adjoint): $A^\dagger = (A^*)^T$. The conjugate transpose or adjoint matrix is obtained by taking the transpose and then the complex conjugate of each entry (the two operations commute). The adjoint of a matrix containing only real numbers equals its transpose.

$$(A^\dagger)_{ij} = (A_{ji})^* \tag{1.23}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}^\dagger = \begin{bmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \\ a_{13}^* & a_{23}^* \end{bmatrix} \tag{1.24}$$

Throughout this book we use the dagger † exclusively for the adjoint, both of matrices and of abstract operators.

1.2.2 Inner Product of Vectors and Matrix Multiplication

The history of the use of matrices in quantum mechanics has its roots in fundamental mathematical developments throughout the 19th and early 20th centuries. In 1801, **Carl Friedrich Gauss**, in his work *Disquisitiones Arithmeticae*, presented concepts related to "combinations of substitutions," a mathematical operation that we recognize today as matrix multiplication. In 1843, **William Rowan Hamilton** revolutionized algebra by discovering **quaternions**, a hypercomplex number system that extended real and complex numbers and enabled describing rotations in three-dimensional space [10]. Five years later, in 1848, the English mathematician **James Joseph Sylvester** coined the term **"matrix"**, derived from the Latin word for "womb," referring to this structure as an algebraic container that can be systematically manipulated. Later, **William Clifford** combined Grassmann's ideas about vector algebra with theories of hypercomplex systems, developing what are now known as **Clifford algebras**, essential tools in theoretical physics and geometry [14].

The use of matrices in quantum physics was consolidated with Werner Heisenberg, who in 1925 formulated matrix quantum mechanics in his pioneering article [11]. Heisenberg developed a new approach to describing the behavior of subatomic particles based on arrays of numbers combined by what he had, in effect, reinvented: matrix multiplication, an operation that was not part of the standard training of physicists at the time. Shortly thereafter, Max Born and Pascual Jordan recognized that the operations introduced by Heisenberg were indeed matrix multiplications, and contributed to formalizing the theory.

Inner Product of Vectors

We define a vector as a matrix that has a single row or a single column, thus:

Row vector:

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix} \tag{1.25}$$

Column vector:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} \tag{1.26}$$

In this text, these vectors are considered to have complex components. A row vector and a column vector can be multiplied, provided they have the same number of entries; the result of this operation is a number. This is a special case of a matrix product between these two vectors:

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 + \dots + a_n b_n = \sum_{i=1}^{n} a_i b_i \tag{1.27}$$

We extend this concept to matrix multiplication as follows. If $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ is an $m \times l$ matrix, then $\mathbf{C} = \mathbf{AB}$ is an $n \times l$ matrix with entries:

$$C_{ik} = \sum_{j=1}^{m} A_{ij} B_{jk} \tag{1.28}$$

This multiplication is only possible if the number of columns of $\mathbf{A}$ coincides with the number of rows of $\mathbf{B}$.

Matrix multiplication has the following properties:

- Associativity: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC}$
- Left distributivity: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- Right distributivity: $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
- Scalar multiplication: $\lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$
- Complex conjugate: $(\mathbf{AB})^* = \mathbf{A}^*\mathbf{B}^*$
- Transpose: $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- Conjugate transpose: $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$

Matrix multiplication is, in general, not commutative: $\mathbf{AB} \neq \mathbf{BA}$ except in special cases (for instance, when one of the matrices is the identity, or when both are diagonal). This non-commutativity, far from being a nuisance, encodes deep physics: as we will see, the failure of observables to commute is at the heart of the uncertainty relations.

1.2.3 Bra-ket: Dirac Notation

Paul Adrien Maurice Dirac (1902–1984) was one of the most influential physicists of the 20th century and a pioneer in the development of quantum mechanics and quantum electrodynamics. Born in Bristol, England, Dirac combined logical rigor with deep physical intuition. In 1928, he formulated the **Dirac equation**, which described the relativistic behavior of electrons and predicted the existence of antimatter; this led him to win the Nobel Prize in Physics in 1933, shared with Erwin Schrödinger.

Dirac consolidated the general formalism of quantum mechanics in his influential treatise *The Principles of Quantum Mechanics*, first published in 1930 [6]. The compact bra-ket notation that now bears his name was introduced somewhat later, in a short 1939 paper [7], and was incorporated into the third edition of the *Principles* (1947), from which it spread to become the universal standard for describing quantum states. It is worth noting that when quantum mechanics was being created, linear algebra was not a widely taught subject, even among physicists and mathematicians—as mentioned above, Heisenberg had to rediscover matrix multiplication on his own. The success of quantum mechanics contributed significantly to the inclusion of linear algebra in academic curricula in subsequent decades.

The basic idea is that a ket, denoted as $|\psi\rangle$, is a vector in a Hilbert space, a complex vector space used to describe quantum states. If we have N basis vectors $|e_i\rangle$ with $i = 1, 2, 3, \ldots, N$, then any vector $|v\rangle$ in this space can be expressed as a linear combination of the basis vectors:

$$|v\rangle = \sum_i v_i \, |e_i\rangle$$

The vector $|v\rangle$ can be represented in column form:

$$|v\rangle = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_N \end{bmatrix}$$

This representation is analogous to classical vector notation in 3 dimensions, where a vector $\vec{a}$ in three-dimensional space is expressed in terms of its Cartesian components as:

$$\vec{a} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

A "bra", denoted as $\langle\psi|$, is the dual vector (linear functional) corresponding to the ket $|\psi\rangle$. The bra $\langle a|$ is obtained by taking the Hermitian adjoint of the corresponding ket $|a\rangle$, which in the column-vector representation means taking the conjugate transpose:

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

$$|a\rangle = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \quad \Rightarrow \quad \langle a| = \begin{bmatrix} a_x^* & a_y^* & a_z^* \end{bmatrix}$$

Let us give a physical interpretation using Dirac notation. In quantum mechanics, the states of a system are represented by kets $|\psi\rangle$ in a Hilbert space. For example, if a particle has a definite momentum, we can represent this state as $|p\rangle$, where p is the eigenvalue of the momentum operator. More specifically, $|p = 7\rangle$ represents the eigenstate in which the particle has momentum of exactly 7 kg·m/s. Similarly, $|x = 2.33\rangle$ represents the state in which the particle is localized at position $x = 2.33$ m.[1]

The scalar (inner) product of a bra and a ket, written as $\langle\phi|\psi\rangle$, is a complex number. Physically, $\langle\phi|\psi\rangle$ represents the probability amplitude that a system prepared in state $|\psi\rangle$ will be found in state $|\phi\rangle$ upon measurement. The square of its modulus, $|\langle\phi|\psi\rangle|^2$, gives the corresponding transition probability.

To illustrate this, $\langle x|\psi\rangle$ corresponds to the *probability amplitude* that a system in state $|\psi\rangle$ is found at position x. This amplitude is a function of x, denoted as $\psi(x) = \langle x|\psi\rangle$, and is known as the *wavefunction in position space*. The *probability density* of finding the particle at position x is given by $|\psi(x)|^2 = |\langle x|\psi\rangle|^2$, so that the probability of finding the particle in an infinitesimal interval dx around x is $|\psi(x)|^2 dx$.

Dirac notation allows expressing the quantum state in different representations. While $|\psi\rangle$ is an abstract description of the state (independent of the basis), its projection onto the position eigenstates $|x\rangle$ gives us the wavefunction $\psi(x) = \langle x|\psi\rangle$. Analogously, the projection onto the momentum eigenstates $|p\rangle$ defines the wavefunction in momentum space:

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{-ipx/\hbar} dx \tag{1.29}$$

where we have used the notation $\tilde{\psi}(p)$ to distinguish it from $\psi(x)$.

For example, for a particle in the ground state of a one-dimensional box of length $L = 1$, the wavefunction in position space is $\psi(x) = \sqrt{2}\sin(\pi x)$ for $0 \leq x \leq 1$, and zero outside this interval [8]. The corresponding wavefunction in momentum space would be:

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{0}^{1} \sqrt{2}\sin(\pi x) e^{-ipx/\hbar} dx \tag{1.30}$$

A fundamental principle of quantum mechanics establishes that the state $|\psi\rangle$ contains all the physical information of the system. Knowing $\psi(x)$ or $\tilde{\psi}(p)$ is equivalent, since both are representations of the same abstract state $|\psi\rangle$ in different bases.

It is important to distinguish clearly: a *probability amplitude* is a complex number that can interfere constructively or destructively, while a *probability* is a non-negative real number obtained as the modulus squared of the amplitude. This distinction is fundamental for understanding quantum phenomena such as interference and entanglement.

[1] Strictly speaking, position and momentum eigenstates are idealized, non-normalizable states; they are nevertheless extremely useful as a basis for expanding physical states. For the finite-dimensional spaces of quantum computing this subtlety does not arise.

A *state vector* $|\psi\rangle$ can have complex coefficients when expanded in a basis. For example, for a two-level system (qubit):

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{1.31}$$

where $\alpha = a + ib$ and $\beta = c + id$ are complex amplitudes with $a, b, c, d \in \mathbb{R}$.

The corresponding bra is:

$$\langle\psi| = \alpha^*\,\langle 0| + \beta^*\,\langle 1| = (a - ib)\,\langle 0| + (c - id)\,\langle 1| \tag{1.32}$$

The inner product of the state with itself gives:

$$\langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2 = (a^2 + b^2) + (c^2 + d^2) \tag{1.33}$$

For the state to be normalized, we require $\langle\psi|\psi\rangle = 1$.

1.2.4 Inner Product in Dirac Notation

In Dirac notation, the inner product, written as $\langle u|v\rangle$, is a complex number; physically it represents the transition probability amplitude between states $|v\rangle$ and $|u\rangle$. Let $|u\rangle$ and $|v\rangle$ be vectors in $\mathbb{C}^n$:

$$|u\rangle = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}, \quad |v\rangle = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

The inner product $\langle u|v\rangle$ is calculated as:

$$\langle u|v\rangle = \begin{bmatrix} a_1^* & a_2^* & a_3^* & \cdots & a_n^* \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^{n} a_i^* b_i$$

The inner product has the following fundamental properties:

- Hermitian symmetry: $\langle u|v\rangle = \overline{\langle v|u\rangle}$.
- $\langle u|u\rangle = \sum_{i=1}^{n} |a_i|^2$ is a real and non-negative number.
- $\langle u|u\rangle \geq 0$, and $\langle u|u\rangle = 0$ if and only if $|u\rangle = 0$.

A crucial structural property is the behavior of the inner product under linear combinations. Throughout this book we follow the physics convention: *the inner product is linear in the ket (second argument) and antilinear in the bra (first argument)*. Linearity in the ket means:

$$\langle u|c_1v_1 + c_2v_2\rangle = c_1 \langle u|v_1\rangle + c_2 \langle u|v_2\rangle$$

for any complex constants c_1 and c_2. Antilinearity in the bra means that the constants come out conjugated:

$$\langle c_1u_1 + c_2u_2|v\rangle = c_1^* \langle u_1|v\rangle + c_2^* \langle u_2|v\rangle$$

Both properties follow directly from the coordinate formula $\langle u|v\rangle = \sum_i a_i^* b_i$: the components of the bra enter conjugated, those of the ket do not. A form with these two properties is called *sesquilinear* ("one-and-a-half linear").[2]

The antilinearity in the bra gives us a simple rule to pass from a ket to its corresponding bra, and vice versa:

$$|v\rangle = \alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle \;\leftrightarrow\; \langle v| = \alpha_1^* \langle a_1| + \alpha_2^* \langle a_2|$$

Two vectors $|u\rangle$ and $|v\rangle$ are orthogonal if their inner product is zero, that is, if $\langle u|v\rangle = 0$. The norm of a vector $|v\rangle$ is defined as $\| \, |v\rangle \, \| = \sqrt{\langle v|v\rangle}$. The Cauchy–Schwarz inequality [1] establishes that, for any pair of vectors $|u\rangle$ and $|v\rangle$:

$$|\langle u|v\rangle| \leq \| \, |u\rangle \, \| \, \| \, |v\rangle \, \|$$

Now consider an orthonormal basis $\{|e_i\rangle\}$, where the orthonormality condition is expressed as $\langle e_i|e_j\rangle = \delta_{ij}$. A vector $|v\rangle$ can be expanded in this basis as:

$$|v\rangle = \sum_i \alpha_i |e_i\rangle$$

The coefficients α_i are extracted using the inner product and linearity in the ket:

$$\langle e_k|v\rangle = \langle e_k| \Big(\sum_i \alpha_i |e_i\rangle \Big) = \sum_i \alpha_i \langle e_k|e_i\rangle = \alpha_k$$

Thus, we can write any vector in terms of its expansion in an orthonormal basis:

$$|v\rangle = \sum_i |e_i\rangle \langle e_i|v\rangle$$

an identity that, as we will see, expresses the *completeness* of the basis: $\sum_i |e_i\rangle \langle e_i| = I$.

[2] Mathematics texts often adopt the opposite convention, with linearity in the first argument. The physics convention is universal in quantum mechanics and quantum computing, and we use it consistently.

Norm and Orthogonality Using Dirac Notation

The norm of a vector $|u\rangle$ is a non-negative real number defined as:

$$\||u\rangle\| = \sqrt{\langle u|u\rangle} = \sqrt{\sum_{i=1}^{n} |a_i|^2}$$

A vector $|u\rangle$ is said to be *normalized* (or a *unit vector*) if its norm equals 1, which implies that:

$$\sum_{i=1}^{n} |a_i|^2 = 1$$

(The adjective *unitary* is reserved in this book for operators and matrices, to avoid confusion.)

In the case of a qubit in the state $\alpha |0\rangle + \beta |1\rangle$, the normalization condition is expressed as:

$$|\alpha|^2 + |\beta|^2 = 1$$

This ensures that the qubit is in a valid quantum state: when measured in the computational basis $\{|0\rangle, |1\rangle\}$, the probabilities of the two outcomes sum to one.

The requirement that state vectors be normalized is a fundamental principle in quantum mechanics. It arises from the Born rule: the squared moduli of the expansion coefficients of the state are probabilities, and the total probability over all possible outcomes must equal one. In quantum computing, normalization is crucial because qubits must be in states that yield valid probabilities for quantum operations and measurements to produce physically meaningful outcomes.

1.2.5 Bases: Standard and Computational

In the vector space $\mathbb{C}^n$, a basis is a collection of linearly independent vectors $\{|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle\}$ such that any vector $|v\rangle \in \mathbb{C}^n$ can be expressed as a linear combination of them:

$$|v\rangle = \alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle + \cdots + \alpha_n |v_n\rangle$$

where the coefficients $\alpha_i \in \mathbb{C}$ are complex numbers. Linear independence means that no vector $|v_i\rangle$ can be expressed as a linear combination of the others.

The number n of vectors in a basis determines the dimension of the vector space; for example, the dimension of $\mathbb{C}^n$ is n. Given a basis, any vector $|v\rangle$ in $\mathbb{C}^n$ is uniquely represented by the ordered tuple of its n coordinates $\alpha_1, \alpha_2, \ldots, \alpha_n$.

A basis is called orthonormal if it is composed of mutually orthogonal unit vectors, which is expressed by the condition:

$$\langle v_i|v_j\rangle = \delta_{ij}$$

where δ_{ij} is the Kronecker delta, which equals 1 if $i = j$ and 0 otherwise.

As an example, in the vector space $\mathbb{C}^2$, two possible orthonormal bases are:

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad \text{or also} \quad \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

In $\mathbb{C}^n$, the standard basis—called the *computational basis* in quantum computing—is given by the vectors:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, |2\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots \quad |n-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Note that the labeling starts at zero: this is the universal convention in quantum computing, chosen so that for a single qubit ($n = 2$) the two basis states are $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, matching the values of a classical bit. Any vector $|v\rangle$ in $\mathbb{C}^n$ can be expanded in this computational basis as:

$$|v\rangle = \sum_{i=0}^{n-1} a_i |i\rangle$$

where the coefficients a_i are the coordinates of $|v\rangle$ in the computational basis.

1.2.6 Outer Product and Projection

We saw that the inner product multiplies a bra (row vector) by a ket (column vector) and produces a number. The matrix multiplication rule remains valid if we multiply in the opposite order—a column vector by a row vector. This operation produces an $n \times n$ matrix instead of a number, and is called the *outer product*:

$$|\phi\rangle \langle\psi| = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} \begin{bmatrix} \psi_1^* & \psi_2^* & \dots & \psi_N^* \end{bmatrix} = \begin{bmatrix} \phi_1\psi_1^* & \phi_1\psi_2^* & \dots & \phi_1\psi_N^* \\ \phi_2\psi_1^* & \phi_2\psi_2^* & \dots & \phi_2\psi_N^* \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N\psi_1^* & \phi_N\psi_2^* & \dots & \phi_N\psi_N^* \end{bmatrix}$$

Note that the components of the bra enter conjugated, consistent with $\langle\psi| = |\psi\rangle^\dagger$. The outer product is a linear operator: it can act on other vectors. Applying it to an arbitrary ket $|\chi\rangle$, and using the associativity of matrix multiplication:

$$(|\phi\rangle \langle\psi|) |\chi\rangle = |\phi\rangle \langle\psi|\chi\rangle = \langle\psi|\chi\rangle \, |\phi\rangle$$

The operator $|\phi\rangle\langle\psi|$ thus takes any vector $|\chi\rangle$ and returns the vector $|\phi\rangle$ scaled by the amplitude $\langle\psi|\chi\rangle$. The name "outer" signals that it is the complementary construction to the inner product.

One of the main uses of the outer product is to construct *projection operators*. A projection is a linear transformation $\mathbf{P}$ from a vector space to itself such that $\mathbf{P}^2 = \mathbf{P}$. This property is called *idempotence*: applying $\mathbf{P}$ twice gives the same result as applying it once.

A classic example is the orthogonal projection of $\mathbb{R}^3$ onto the xy plane, which maps each point (x, y, z) to the point $(x, y, 0)$, represented by the matrix:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The action of this matrix on an arbitrary vector is:

$$\mathbf{P}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

which immediately shows its idempotence:

$$\mathbf{P}^2\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{P}\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

Using Dirac notation and the outer product, an orthogonal projection onto the subspace spanned by an orthonormal set $\{|\psi_1\rangle, \ldots, |\psi_r\rangle\}$ can be written as:

$$\mathbf{P} = \sum_{j=1}^{r} |\psi_j\rangle\langle\psi_j|$$

In the simplest case, the projector onto the direction of a single normalized vector $|\psi\rangle$ is $\mathbf{P} = |\psi\rangle\langle\psi|$. Its idempotence is verified directly:

$$\mathbf{P}^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \mathbf{P}$$

where we used $\langle\psi|\psi\rangle = 1$. Projectors of this form model quantum measurements, as we will formalize in the postulates of Chapter 3: measuring a system in state $|\chi\rangle$ and obtaining the outcome associated with $|\psi\rangle$ leaves the system in (the normalization of) $\mathbf{P}|\chi\rangle$, with probability $|\langle\psi|\chi\rangle|^2$.

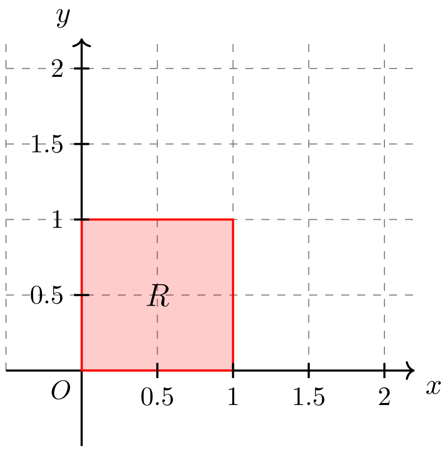

Figure 1.2: Representation of a square of side 1 in a coordinate system with grid. The Cartesian axes, origin, and main divisions are shown.

1.3 Eigenvalues, Hermitian and Unitary Matrices

1.3.1 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are fundamental concepts in the study of linear operators. An eigenvector of a matrix is a non-zero vector that, when transformed by that matrix, results in a scalar multiple of itself: the operator does not change its direction, but scales it by a factor known as the eigenvalue. In physics, identifying these invariants is crucial: as we will see, eigenvalues and eigenvectors provide the possible results of measurements and the states associated with them.

Let us look at a concrete example. Consider the matrix:

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

We want to analyze the effect of this matrix on vectors in $\mathbb{R}^2$. Multiplying $\mathbf{A}$ by a generic vector:

$$\mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x + 2y \\ x + 4y \end{bmatrix}$$

Consider a square of side 1 located at the origin of the coordinate system, as shown in Figure 1.2.

What happens to this square under the action of matrix $\mathbf{A}$? Let us see how its vertices are transformed:

$$\mathbf{A} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

Graphically, the square is transformed into the following parallelogram:

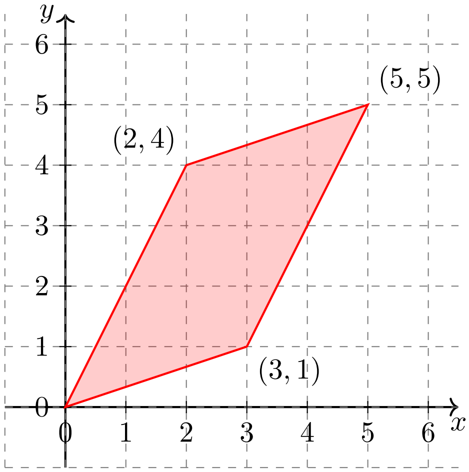

Figure 1.3: Transformation of a square of side 1 into a parallelogram through the linear transformation $\mathbf{A}$. The resulting parallelogram has vertices at $(0, 0)$, $(3, 1)$, $(5, 5)$ and $(2, 4)$.

Do there exist vectors that, after transformation, preserve their direction? Let us analyze the results:

- The vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ was transformed to $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. It does not preserve its direction.
- The vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ was transformed to $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$. It does not preserve its direction.
- The vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ was transformed to $\begin{bmatrix} 5 \\ 5 \end{bmatrix}$. It preserves its direction, with a dilation factor of 5.

The vector that maintained its direction is an *eigenvector*, and the factor by which it was dilated is the corresponding *eigenvalue*:

$$\mathbf{A} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This matrix has a second eigenvalue, which the four vertices of the square happened not to reveal: the vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ satisfies $\mathbf{A} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, so $\lambda = 2$ is also an eigenvalue, with eigenvector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. We will see below how to find *all* the eigenvalues of a matrix systematically.

Eigenvalues and Eigenvectors in Dirac Notation

In Dirac notation, an eigenvector of a matrix $A \in \mathbf{M}_{n,n}(\mathbb{C})$ is a non-zero vector $|v\rangle \in \mathbb{C}^n$ such that:

$$A\,|v\rangle = \lambda\,|v\rangle$$

where the number λ is the eigenvalue corresponding to the eigenvector $|v\rangle$.

The eigenvalues of a matrix A are obtained as the roots of the characteristic polynomial:

$$\det(A - \lambda I) = 0$$

where det denotes the determinant and I is the identity matrix, which in Dirac notation can be written as $I = \sum_i |e_i\rangle\langle e_i|$ for any orthonormal basis $\{|e_i\rangle\}$. By the fundamental theorem of algebra, the characteristic polynomial of an $n \times n$ complex matrix always has at least one root in $\mathbb{C}$; therefore every square complex matrix has at least one eigenvalue and one non-trivial eigenvector.

For the particular case of 2×2 matrices, the characteristic polynomial takes the convenient form:

$$p(\lambda) = \lambda^2 - \lambda\ \mathrm{tr}(A) + \det(A)$$

The roots of this polynomial are the eigenvalues λ_k, with $k = 1, 2$, and the corresponding eigenvectors are the non-trivial solutions of $(A - \lambda_k I)\,|v\rangle = 0$.

Example: Eigenvalues of a Non-Hermitian Matrix

Given the matrix $A = \begin{bmatrix} i & 1 \\ 0 & -1+i \end{bmatrix}$, let us find its eigenvalues and eigenvectors. First we calculate the trace and determinant of A:

$$\mathrm{tr}(A) = i + (-1+i) = -1+2i, \qquad \det(A) = i(-1+i) - (0 \cdot 1) = -1-i$$

The characteristic polynomial is:

$$p(\lambda) = \lambda^2 - (\mathrm{tr}\,A)\lambda + \det A = \lambda^2 - (-1+2i)\lambda + (-1-i)$$

Solving this quadratic (or simply noting that A is upper triangular, so its eigenvalues are its diagonal entries), we obtain:

$$\lambda_1 = -1+i, \quad \lambda_2 = i$$

1. Eigenvector corresponding to $\lambda_1 = -1+i$. We solve $(A - \lambda_1 I)\,|u_1\rangle = 0$:

$$A - (-1+i)I = \begin{bmatrix} i+1-i & 1 \\ 0 & -1+i-(-1+i) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Therefore:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad |u_1\rangle = \alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where α is an arbitrary non-zero complex scalar.

2. Eigenvector corresponding to $\lambda_2 = i$. Similarly, we solve $(A-\lambda_2 I)\,|u_2\rangle = 0$:

$$A - iI = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$$

Which gives us:

$$\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad |u_2\rangle = \beta \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

where β is an arbitrary non-zero complex scalar. Note that this matrix is not Hermitian: its eigenvalues are complex, and its eigenvectors are not orthogonal ($\langle u_1 | u_2 \rangle = 1 \neq 0$). The contrast with the Hermitian case, developed next, is precisely the point of this example.

1.3.2 The Stern–Gerlach Experiment and Spin

The Stern–Gerlach experiment [17], conducted in 1922 by Otto Stern and Walther Gerlach, was a fundamental milestone in the understanding of quantum mechanics. In this experiment, a beam of silver atoms was passed through a non-homogeneous magnetic field. According to classical physics, the magnetic moments of the atoms should be oriented at random, producing a continuous distribution of deflections on the detection screen, as shown in Figure 1.4. The result was surprising: the beam split into two discrete components, indicating that the projection of the atomic magnetic moment along the field direction is quantized, as illustrated in Figure 1.5.

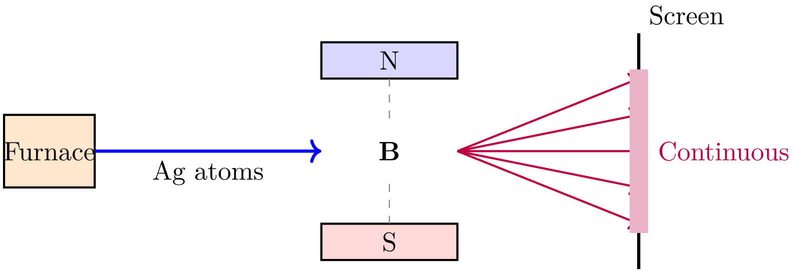

Figure 1.4: Classical expectation of the Stern–Gerlach experiment. A beam of unpolarized silver atoms passing through a non-uniform magnetic field was expected to produce a continuous distribution of impacts on the detection screen, reflecting a gradual variation of magnetic moments.

This result provided clear evidence that certain physical quantities, such as angular momentum, cannot take arbitrary values, but are restricted to discrete

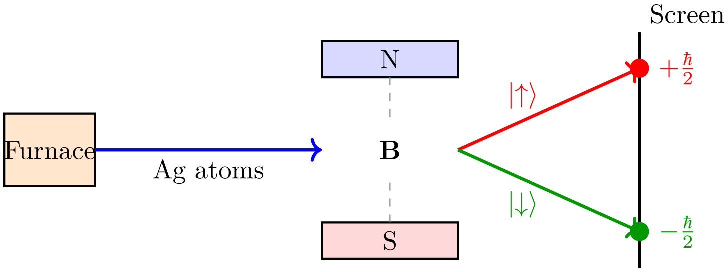

Figure 1.5: Stern–Gerlach experiment illustrating spin quantization. A beam of unpolarized silver atoms splits into two distinct trajectories ($|\uparrow\rangle$ and $|\downarrow\rangle$) when passing through a non-uniform magnetic field, detected as two separate spots on the screen, corresponding to spin projections $\pm\hbar/2$.

ones. The modern interpretation came a few years later: in 1925, George Uhlenbeck and Samuel Goudsmit proposed that the electron possesses an intrinsic angular momentum—**spin**—to explain anomalies in atomic spectra, and the two spots of the Stern–Gerlach experiment were then understood as the two possible projections of the spin of the unpaired valence electron of silver, $\pm\hbar/2$. Spin has no exact analog in classical physics, where angular momentum arises from the physical rotation of extended objects: it is an internal, purely quantum degree of freedom. Its importance is hard to overstate—it underlies the magnetism of materials and magnetic resonance imaging, and the two-dimensional state space of a spin-$\frac{1}{2}$ particle, spanned by the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, is the prototype of the qubit.

The binary character of spin-$\frac{1}{2}$ measurements led Wolfgang Pauli to a compact mathematical formalization through the **Pauli matrices** [16]: three 2×2 Hermitian matrices that represent the spin observables (in units of $\hbar/2$) along the x, y and z directions:

$$\sigma_1 = \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.34}$$

Both numerical $(\sigma_1, \sigma_2, \sigma_3)$ and Cartesian $(\sigma_x, \sigma_y, \sigma_z)$ subscripts are common in the literature; we will use them interchangeably, always with the definitions of Equation (1.34). Each Pauli matrix has eigenvalues ± 1, which correspond to the two possible outcomes of a spin measurement along the given direction (in units of $\hbar/2$). Their Hermiticity guarantees that these eigenvalues are real, as required for measurement outcomes—a general fact we prove below.

1.3.3 Worked Example: Eigenvalues and Eigenvectors of the Pauli Matrices

Rather than repeating the same computation three times, we work through the most instructive case in full detail—σ_2, whose complex entries exercise all the machinery—and then summarize the results for the three matrices, leaving σ_1 and σ_3 as (recommended) exercises.

Consider a generic state $|\psi\rangle$ in $\mathbb{C}^2$, written as a linear combination of the computational basis states:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}, \qquad |0\rangle = \begin{bmatrix}1\\ 0\end{bmatrix}, \quad |1\rangle = \begin{bmatrix}0\\ 1\end{bmatrix}$$

Applying σ_2 to $|\psi\rangle$:

$$\sigma_2\,|\psi\rangle = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix}-i\beta\\ i\alpha\end{bmatrix}$$

For $|\psi\rangle$ to be an eigenvector, $\sigma_2\,|\psi\rangle = \lambda\,|\psi\rangle$ must hold:

$$\begin{bmatrix}-i\beta\\ i\alpha\end{bmatrix} = \lambda\begin{bmatrix}\alpha\\ \beta\end{bmatrix} \quad\Longrightarrow\quad -i\beta = \lambda\alpha, \qquad i\alpha = \lambda\beta$$

For a non-trivial solution to exist, λ must be a root of the characteristic polynomial

$$p(\lambda) = \lambda^2 - \lambda\ \mathrm{tr}(\sigma_2) + \det(\sigma_2) = \lambda^2 - 1,$$

which gives $\lambda = \pm 1$, the expected eigenvalues of a Pauli matrix. We now find the eigenvector for each.

Case 1: $\lambda = 1$. The conditions become $-i\beta = \alpha$ and $i\alpha = \beta$, both equivalent to $\beta = i\alpha$. The corresponding normalized eigenvector is:

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix}$$

Case 2: $\lambda = -1$. The conditions become $-i\beta = -\alpha$ and $i\alpha = -\beta$, both equivalent to $\beta = -i\alpha$. The corresponding normalized eigenvector is:

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -i\end{bmatrix}$$

As a check of the general theory: the eigenvalues are real (Hermiticity) and the eigenvectors are orthogonal,

$$\langle\psi_+|\psi_-\rangle = \frac{1}{2}\begin{bmatrix}1 & -i\end{bmatrix}\begin{bmatrix}1\\ -i\end{bmatrix} = \frac{1}{2}(1 + i^2) = \frac{1}{2}(1-1) = 0$$

(note the conjugation of the components of the bra). The analogous computations for σ_1 and σ_3 are simpler, since their entries are real; the results for the three matrices are collected in Table 1.1.

Matrix	Eigenvalue +1	Eigenvalue −1	Common names
$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\lvert+\rangle\,,\ \lvert-\rangle$
$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$	$\lvert+i\rangle\,,\ \lvert-i\rangle$
$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\lvert0\rangle\,,\ \lvert1\rangle$

Table 1.1: Eigenvalues and normalized eigenvectors of the Pauli matrices. Each pair of eigenvectors forms an orthonormal basis of $\mathbb{C}^2$; the eigenbasis of σ_3 is the computational basis.

Pauli Matrices and Observables

The operator σ_3 is the spin operator in the z direction (in units of $\hbar/2$): its eigenvalues ±1 are the possible results of a spin measurement along that direction, indicating whether the spin is found aligned or anti-aligned with the field, and its eigenvectors $|0\rangle\,, |1\rangle$ are the states in which the system is left after the corresponding outcome is obtained. It is worth being precise with the terminology: the *observable* is the physical quantity, represented by the Hermitian operator; the *eigenvalues* are its possible measured values; and the *eigenvectors* are the post-measurement states.

The Pauli matrices are Hermitian, which guarantees that their eigenvalues are real, and together with the identity matrix $\sigma_0 = I$, they form a basis for the real vector space of 2×2 Hermitian matrices—a fact we exploit below. In quantum computing, the Pauli matrices play an essential role not only in describing spin, but also as quantum gates (X, Y, Z) and as the building blocks of more general operators. Generalizations of the Pauli matrices to higher-dimensional systems (qudits) will be studied in later chapters.

1.3.4 Hermitian and Normal Matrices

We now formalize the two properties repeatedly observed above. A matrix $A \in \mathbf{M}_{n,n}(\mathbb{C})$ is called **Hermitian** if it equals its adjoint:

$$A = A^\dagger \tag{1.35}$$

More generally, a matrix is called **normal** if it commutes with its adjoint:

$$AA^\dagger = A^\dagger A \tag{1.36}$$

Every Hermitian matrix is normal, but not conversely (unitary matrices, discussed next, are normal without being Hermitian in general). The importance of normality is the *spectral theorem*: a matrix is diagonalizable by a unitary change of basis—equivalently, it admits an orthonormal basis of eigenvectors—if and only if it is normal. A matrix can be diagonalizable (with n linearly independent eigenvectors) without being normal; in that case its eigenvectors exist but need not be orthogonal, as in the triangular example above.

Hermitian matrices enjoy two additional properties of central physical importance, which we now prove.

Theorem: The Eigenvalues of a Hermitian Matrix are Real

Let $A = A^\dagger$ and let $|v\rangle$ be a normalized eigenvector with eigenvalue λ, so $A|v\rangle = \lambda|v\rangle$. Taking the inner product with $\langle v|$:

$$\langle v|A|v\rangle = \lambda\langle v|v\rangle = \lambda \tag{1.37}$$

Now take the complex conjugate of this number. Using the Hermitian symmetry of the inner product and $A^\dagger = A$:

$$\langle v|A|v\rangle^* = \left\langle v|A^\dagger|v\right\rangle = \langle v|A|v\rangle \tag{1.38}$$

Hence $\lambda^* = \lambda$: the eigenvalue is real. ■

Theorem: Eigenvectors of a Hermitian Matrix with Different Eigenvalues are Orthogonal

Let $A = A^\dagger$, and let

$$A|\phi_m\rangle = \lambda_m|\phi_m\rangle, \qquad A|\phi_n\rangle = \lambda_n|\phi_n\rangle \tag{1.39}$$

with $\lambda_m \neq \lambda_n$. Taking the inner product of the first equation with $\langle\phi_n|$:

$$\langle\phi_n|A|\phi_m\rangle = \lambda_m\langle\phi_n|\phi_m\rangle \tag{1.40}$$

On the other hand, letting A act to the left on $\langle\phi_n|$: since A is Hermitian and—by the previous theorem—λ_n is real, $\langle\phi_n|A = \langle\phi_n|A^\dagger = (A|\phi_n\rangle)^\dagger = \lambda_n^*\langle\phi_n| = \lambda_n\langle\phi_n|$, so:

$$\langle\phi_n|A|\phi_m\rangle = \lambda_n\langle\phi_n|\phi_m\rangle \tag{1.41}$$

Subtracting the two expressions:

$$(\lambda_m - \lambda_n)\langle\phi_n|\phi_m\rangle = 0 \tag{1.42}$$

Since $\lambda_m \neq \lambda_n$, we conclude $\langle\phi_n|\phi_m\rangle = 0$: the eigenvectors are orthogonal. ■

This theorem allows us to identify orthogonal states in quantum systems without performing additional calculations, such as integrals or symmetry analyses. Note how the proof used the reality of the eigenvalues—the two theorems work together.

Examples and Applications

Consider the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. This matrix is diagonalizable (its eigenvalues 1 and 2 are distinct), but it is not normal, since $AA^\dagger \neq A^\dagger A$; accordingly, its eigenvectors are not orthogonal. On the other hand, the matrix $\begin{bmatrix} 3 & 2i \\ -2i & 1 \end{bmatrix}$ is Hermitian, since it equals its adjoint, and consequently all its eigenvalues are real and its eigenvectors orthogonal.

In quantum mechanics, physical observables, such as a particle's position, momentum, or energy, are modeled by Hermitian operators in a Hilbert space. The eigenvalues of these operators correspond to the possible measurement results—which is precisely why their reality matters: any physical measurement must yield a real number.

Hermitian Matrices as a Real Vector Space

Hermitian matrices form a real vector space: the sum of two Hermitian matrices is Hermitian, and multiplication by a *real* scalar preserves Hermiticity. For example, the 2×2 Hermitian matrices are exactly those of the form:

$$\begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \tag{1.43}$$

where $a, c \in \mathbb{R}$ and $b \in \mathbb{C}$; counting real parameters (a, c, and the two components of b) shows that this space has real dimension 4. A natural basis for it is given by the identity together with the three Pauli matrices:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.44}$$

so that any 2×2 Hermitian matrix can be written as $H = h_0\sigma_0 + h_x\sigma_x + h_y\sigma_y + h_z\sigma_z$ with real coefficients. This decomposition will be fundamental in the study of single-qubit gates and Hamiltonians in later chapters.

1.3.5 Unitary Matrices

In quantum mechanics and quantum computing, unitary matrices play a role as central as that of Hermitian matrices: while Hermitian operators represent what we can *measure*, unitary operators represent what we can *do*—the transformations a quantum state can undergo.

Definition: A matrix $\mathbf{U} \in \mathbf{M}_{n,n}(\mathbb{C})$ is said to be unitary if it satisfies:

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}, \tag{1.45}$$

where $\mathbf{I}$ is the identity matrix of order $n \times n$. Equivalently, $\mathbf{U}^{-1} = \mathbf{U}^\dagger$: the inverse of a unitary matrix is simply its adjoint.

Properties of Unitary Matrices

Unitary matrices are normal (they commute with their adjoint), and are therefore diagonalizable with orthonormal eigenvectors; their eigenvalues are complex numbers of unit modulus, $e^{i\theta}$. Their defining physical property is the preservation of the inner product: if $|u'\rangle = \mathbf{U}|u\rangle$ and $|v'\rangle = \mathbf{U}|v\rangle$, then:

$$\langle u'|v'\rangle = (\mathbf{U}|u\rangle)^\dagger(\mathbf{U}|v\rangle) = \left\langle u|\mathbf{U}^\dagger\mathbf{U}|v\right\rangle = \langle u|v\rangle\,. \tag{1.46}$$

In particular, taking $|u\rangle = |v\rangle$, unitary operators preserve the norm:

$$\|\mathbf{U}|v\rangle\|^2 = \left\langle v|\mathbf{U}^\dagger\mathbf{U}|v\right\rangle = \langle v|v\rangle = \||v\rangle\|^2. \tag{1.47}$$

Norm preservation is crucial in quantum computing: it ensures that the total probability of a quantum state remains equal to one throughout the computation, which is necessary for measurements to yield physically meaningful results. In practical terms, all quantum operations on closed systems are represented by unitary matrices and are therefore *reversible*—unlike many operations in classical computing, such as the AND and OR gates, which destroy information.

Another important property is that the product of two unitary matrices is also unitary. Considering two unitary matrices $\mathbf{U}_1$ and $\mathbf{U}_2$:

$$(\mathbf{U}_1\mathbf{U}_2)^\dagger(\mathbf{U}_1\mathbf{U}_2) = \mathbf{U}_2^\dagger\mathbf{U}_1^\dagger\mathbf{U}_1\mathbf{U}_2 = \mathbf{U}_2^\dagger\mathbf{I}\mathbf{U}_2 = \mathbf{U}_2^\dagger\mathbf{U}_2 = \mathbf{I}. \tag{1.48}$$

This closure under products means that any sequence of quantum gates is itself a single (composite) unitary operation—the mathematical backbone of the quantum circuit model.

Examples of Unitary Matrices

1. Rotation matrix by $\pi/2$ about the x axis of $\mathbb{R}^3$:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Real orthogonal matrices ($\mathbf{R}\mathbf{R}^T = \mathbf{I}$) are the real special case of unitary matrices; the verification $\mathbf{R}\mathbf{R}^\dagger = \mathbf{I}$ is immediate.

2. Complex unitary matrix:

$$\mathbf{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$

This matrix is unitary (verify it!), and appears in various quantum transformations; it is sometimes called $\sqrt{X}$, since $\mathbf{U}^2 = iX$ up to a global phase.

3. Discrete Fourier Transform (DFT): One of the most important unitary matrices is the Discrete Fourier Transform, used to analyze discrete signals in terms of their frequency components:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi kn/N}, \qquad k = 0, 1, \ldots, N-1.$$

Collecting the coefficients into a matrix and including the normalization factor $1/\sqrt{N}$ (required for unitarity), the entries are $W_{kn} = \omega^{kn}/\sqrt{N}$ with $\omega = e^{-i2\pi/N}$. For $N = 4$, where $\omega = -i$:

$$\mathbf{W}_{4\times 4} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}.$$

This matrix is unitary and symmetric ($\mathbf{W}^T = \mathbf{W}$). Its quantum analogue, the Quantum Fourier Transform, is the engine of Shor's factoring algorithm, which reduces integer factorization—requiring superpolynomial time with the best known classical algorithms—to a problem solvable in polynomial time, thus threatening the security of cryptographic systems like RSA.

Quantum Gates

In quantum computing, operations on qubits are performed through quantum gates: linear operations represented mathematically by unitary matrices. A quantum gate acts on one or more qubits and transforms their state into a new state, preserving the norm and hence the total probability. Some fundamental examples:

- **Hadamard Gate (H):** The Hadamard gate is used to create superpositions. It is represented by the unitary matrix:

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

  This gate transforms the basis state $|0\rangle$ into the equal superposition $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and the state $|1\rangle$ into $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

- **Pauli-X Gate:** Also known as the quantum NOT gate, it exchanges the basis states, transforming $|0\rangle$ into $|1\rangle$ and $|1\rangle$ into $|0\rangle$. It is represented by the matrix:

$$\mathbf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

  which is none other than the Pauli matrix σ_x—simultaneously Hermitian and unitary.

- **CNOT Gate (Controlled-NOT):** This gate operates on two qubits, one acting as the control and the other as the target. If the control qubit is in state $|1\rangle$, the gate inverts the state of the target qubit; otherwise it does nothing. Its unitary representation is:

$$\mathbf{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Reversibility and the No-Cloning Theorem

Since $\mathbf{U}^{-1} = \mathbf{U}^\dagger$ always exists, every quantum gate can be undone: quantum computation on closed systems is intrinsically reversible, and no information is lost during unitary processing. This reversibility interacts with another fundamental restriction of quantum information: the **no-cloning theorem**, formulated in 1982 by William Wootters and Wojciech Zurek [18] and independently by Dennis Dieks [5], which establishes that it is impossible to create an exact copy of an arbitrary unknown quantum state. Unlike classical information, which can be copied freely, quantum information cannot be duplicated; this guarantees, for example, the security of quantum key distribution protocols like BB84 (an eavesdropper cannot copy the transmitted quantum states without disturbing them), and it implies that quantum algorithms cannot rely on making backup copies of intermediate states or amplifying a state by duplication. Both restrictions—reversibility and no-cloning—shape quantum circuit design in ways that have no classical analogue.

Temporal Evolution in Quantum Systems

Another central application of unitary matrices is the temporal evolution of quantum systems. According to the Schrödinger equation, the evolution of a closed quantum system is described by a unitary operator: given an initial state $|\psi(0)\rangle$, the state at time t is

$$|\psi(t)\rangle = \mathbf{U}(t)\,|\psi(0)\rangle\,, \qquad \mathbf{U}(t) = e^{-i\mathbf{H}t/\hbar},$$

where $\mathbf{H}$ is the system Hamiltonian—a Hermitian operator. Here the two families of matrices of this section meet: every Hermitian operator generates, by exponentiation, a unitary evolution. This relation will be developed in detail in Chapter 3.

1.3.6 Diagonal Representation and Spectral Decomposition

A linear operator $A \in \mathbf{M}_{n,n}(\mathbb{C})$ is diagonalizable if it can be represented in a basis in which its action reduces to scalar multiplication of each component. This property is of particular importance in quantum mechanics, where the operators describing physical observables are Hermitian and hence, by the spectral theorem, always diagonalizable in an orthonormal basis.

We say that a matrix A admits a *spectral decomposition* if there exists an orthonormal basis of eigenvectors $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ such that:

$$A = \sum_{i=1}^{n} \lambda_i |v_i\rangle \langle v_i|, \tag{1.49}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues associated with the eigenvectors. By the spectral theorem, exactly the normal matrices admit such a decomposition. In this representation, A acts on each eigenvector $|v_i\rangle$ by scaling it by λ_i.

To diagonalize A, we seek a unitary matrix U, whose columns are the normalized eigenvectors of A, such that:

$$U^\dagger A U = D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \tag{1.50}$$

In the eigenvector basis, the action of the operator on any vector is simply the independent scaling of each component by its eigenvalue, which is especially useful in temporal evolution calculations and in solving the Schrödinger equation.

Application in Quantum Mechanics

Spectral decomposition is central to the interpretation of observables. A physical observable, represented by a Hermitian operator A, decomposes as $A = \sum_i \lambda_i |v_i\rangle \langle v_i|$: the eigenvalues λ_i are the possible measurement results, and the projectors $|v_i\rangle \langle v_i|$ implement the corresponding measurements, in the sense formalized by the postulates of Chapter 3.

For example, consider the Hamiltonian H, which represents the total energy of a quantum system. Diagonalizing H yields the system's stationary states: the eigenvalues are the possible energies, and the eigenvectors are the states that—apart from a phase—do not change with time.

Example: Diagonalization of a Hermitian Operator

Consider the Hermitian operator $H = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Its characteristic polynomial is:

$$\det(H - \lambda I) = (3-\lambda)^2 - 1 = (\lambda - 4)(\lambda - 2) = 0, \tag{1.51}$$

with eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$. Solving $(H - \lambda_i I)\,|v_i\rangle = 0$ and normalizing:

$$|v_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad |v_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

which are orthogonal, as the theory demands. The spectral decomposition is then:

$$H = 4\,|v_1\rangle\,\langle v_1| + 2\,|v_2\rangle\,\langle v_2| = 4 \cdot \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + 2 \cdot \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \checkmark$$

1.3.7 Tensor Products

In quantum computing, the tensor product is the fundamental operation used to combine individual quantum states into a joint state of a composite system. This is essential when working with systems of multiple qubits.

Definition of the Tensor Product

The tensor product of two vectors $\mathbf{x} \in \mathbb{C}^2$ and $\mathbf{y} \in \mathbb{C}^2$ is a new vector in $\mathbb{C}^4$, given by:

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \otimes \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 y_0 \\ x_0 y_1 \\ x_1 y_0 \\ x_1 y_1 \end{bmatrix}. \tag{1.52}$$

This product extends to more vectors; for example, for three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^2$, the tensor product is a vector in $\mathbb{C}^8$:

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \otimes \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \otimes \begin{bmatrix} z_0 \\ z_1 \end{bmatrix} = \begin{bmatrix} x_0 y_0 z_0 \\ x_0 y_0 z_1 \\ x_0 y_1 z_0 \\ x_0 y_1 z_1 \\ x_1 y_0 z_0 \\ x_1 y_0 z_1 \\ x_1 y_1 z_0 \\ x_1 y_1 z_1 \end{bmatrix}. \tag{1.53}$$

In general, the tensor product of a vector in $\mathbb{C}^n$ with a vector in $\mathbb{C}^m$ lives in $\mathbb{C}^{nm}$: dimensions *multiply*, which is the origin of the exponential growth 2^n of the state space of n qubits.

Examples of Tensor Products

Concrete examples with real values:

$$\begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}3\\4\end{bmatrix} = \begin{bmatrix}3\\4\\6\\8\end{bmatrix}, \qquad \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\0\end{bmatrix}. \tag{1.54}$$

The second example is, in qubit language, $|1\rangle \otimes |1\rangle \otimes |0\rangle = |110\rangle$: the tensor product of computational basis states is the computational basis state whose label is the concatenation of the individual labels (and, read as a binary number, $110_2 = 6$ indicates the position of the single 1 in the column vector, counting from 0).

Application in Quantum Computing

In quantum computing, the tensor product combines individual qubit states into the joint state of the complete system. If one qubit is in state $|0\rangle$ and another in state $|1\rangle$, the combined state of the two qubits is $|0\rangle \otimes |1\rangle = |01\rangle$. This operation is fundamental for describing multi-qubit systems and is used extensively in the construction of quantum circuits.

A crucial fact, developed in later chapters, is that *not* every state of $\mathbb{C}^4$ is a tensor product of two states of $\mathbb{C}^2$: states that cannot be so factored are called *entangled*, and they are the key resource behind many quantum algorithms and protocols (see Exercise 15).

1.4 Hermitian Operators

According to the postulates of quantum mechanics, which we will address in the next chapter, physical properties such as position, momentum, energy, and spin are each assigned an operator. In general terms, an operator $\hat{O}$ is a rule that transforms one function into another: $\hat{O}f = g$. A classic example is differentiation: $\frac{d}{dx}f(x) = \hat{D}f(x)$, where the differential operator $\hat{D}$ transforms the function $f(x)$ into its derivative $f'(x)$. So far in this chapter we have worked with operators on the finite-dimensional spaces $\mathbb{C}^n$, where they are represented by matrices; the operator language extends the same ideas to infinite-dimensional function spaces.

Operators can be combined to form other operators. They are distributive with respect to addition, $(\hat{A}+\hat{B})f = \hat{A}f + \hat{B}f$, and they can be composed, $(\hat{A}\hat{B})f = \hat{A}(\hat{B}f)$. Another example is the integral operator, such as the Laplace transform.

A fundamental operation between operators is the commutator, defined as:

$$[\hat{A},\hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{1.55}$$

Some important properties of the commutator, which will be useful throughout this book, are:

$$[\hat{B},\hat{A}] = -[\hat{A},\hat{B}] \tag{1.56}$$

$$[\hat{A},\hat{B}+\hat{C}] = [\hat{A},\hat{B}] + [\hat{A},\hat{C}] \tag{1.57}$$

$$[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}] \tag{1.58}$$

Linear operators are the most common in quantum mechanics. An operator is linear if it satisfies $\hat{A}(cf) = c\hat{A}f$ and $\hat{A}(f+g) = \hat{A}f + \hat{A}g$, where c is a scalar, and f and g are arbitrary functions.

For a linear operator, the equation

$$\hat{A}\,U_n = a_n U_n \tag{1.59}$$

is the operator version of the eigenvalue equation we studied for matrices: the scalars a_n are the eigenvalues, and the functions U_n are the associated eigenfunctions. The set of all eigenvalues forms the *spectrum* of the operator.

An operator $\hat{A}$ is **Hermitian** if it satisfies the condition $\hat{A}^\dagger = \hat{A}$, where $\dagger$ denotes the adjoint of the operator (the conjugate transpose in the matrix case).

Expectation Value of a Hermitian Operator

To relate quantum mechanical calculations to quantities measured in the laboratory, one computes the *expectation value* of an operator: the average of all possible measurement results, weighted by their probabilities. For a system described by a normalized wavefunction $\psi(r)$, the expectation value of $\hat{A}$ is:

$$\langle A \rangle = \int \psi(r)^*\,\hat{A}\,\psi(r)\,dr = \left\langle \psi|\hat{A}|\psi \right\rangle \tag{1.60}$$

Measured quantities must be real numbers, so we require $\langle A \rangle = \langle A \rangle^*$ for every state ψ. Writing out this condition:

$$\int \psi(r)^*\,\hat{A}\,\psi(r)\,dr = \left(\int \psi(r)^*\,\hat{A}\,\psi(r)\,dr\right)^* = \int (\hat{A}\psi(r))^*\,\psi(r)\,dr \tag{1.61}$$

Operators satisfying this condition for all admissible states are precisely the Hermitian operators.[3] They are characterized by the following properties:

1. Their eigenvalues are real numbers.

2. Their eigenvectors (or eigenfunctions) corresponding to distinct eigenvalues are orthogonal, and they can be chosen to form a complete orthonormal set.

These are exactly the two theorems we proved for Hermitian matrices, now stated in operator language. The reason why quantum observables are represented by Hermitian operators is that this guarantees that every possible measured value—every eigenvalue—is a real number, as any physical measurement must be.

1.5 Hilbert Space

In linear algebra, a vector space is *a set closed under vector addition and scalar multiplication*. A common example is the Euclidean space of n dimensions, $\mathbb{R}^n$, where each element is represented by a list of n real numbers, scalars are real numbers, vector addition is performed component by component, and scalar multiplication multiplies each component by a real number.

To establish the mathematical structure of quantum mechanics on a rigorous foundation, an adequate vector space is required. In 1912, David Hilbert developed key work on integral equations, in which the infinite-dimensional spaces now bearing his name first arose. Between 1927 and 1932, the mathematician John von Neumann provided the precise axiomatic formulation of these spaces and of quantum mechanics upon them, coining the term **Hilbert space** in the process. Three features of his formulation are worth highlighting:

1. The components of vectors are not limited to real numbers, but are complex numbers.

2. The space carries an inner product $\langle\cdot|\cdot\rangle$ with values in $\mathbb{C}$; the inner product of a vector with itself, $\langle x|x\rangle$, is a non-negative real number, which allows defining lengths and probabilities.

3. The space may be infinite-dimensional, and is required to be *complete* (a technical condition explained below).

Hermann Weyl and Norbert Wiener studied and extended Hilbert's ideas—Weyl wrote one of the first texts on quantum mechanics and group theory using this framework—but it was von Neumann who established the complete axiomatic structure. Dirac's bra-ket notation, introduced in 1939 [7], later provided

[3] In infinite dimensions the precise term is *self-adjoint*, which involves domain conditions beyond the symmetry expressed here; for the finite-dimensional spaces of quantum computing, Hermitian and self-adjoint coincide.

the compact language in which von Neumann's framework is universally written today.

In quantum mechanics, the states of a system are vectors in a Hilbert space, observables are Hermitian operators, symmetries are represented by unitary operators, and measurements are modeled through orthogonal projections, as we have seen in this chapter. These relationships gave a significant boost to group theory and to the development of the spectral theorem for self-adjoint operators; von Neumann continued this line of work in the 1930s, developing the operator algebras that today bear his name. By 1930 it had even become clear that classical mechanics could be reformulated in terms of Hilbert spaces, in what is known as the Koopman–von Neumann formulation.

Formal Definition of Hilbert Space

We now define precisely the arena in which quantum computing operates. This definition synthesizes the concepts developed throughout this chapter—complex numbers, inner products, orthonormal bases, and linear operators—into a coherent mathematical structure.

Definition: Let $\mathbb{H}$ be a complex vector space. An inner product on $\mathbb{H}$ is a function $\langle\cdot|\cdot\rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{C}$ that satisfies the following properties:

1. **Linearity in the second argument:** $\langle x|ay + bz\rangle = a\,\langle x|y\rangle + b\,\langle x|z\rangle$, where $a, b \in \mathbb{C}$ and $x, y, z \in \mathbb{H}$. (Together with property 2, this implies antilinearity in the first argument, consistent with the convention adopted in this book.)
2. **Hermitian symmetry:** $\overline{\langle x|y\rangle} = \langle y|x\rangle$.
3. **Positive definiteness:** $\langle x|x\rangle \geq 0$ for all $x \in \mathbb{H}$, and $\langle x|x\rangle = 0$ if and only if $x = 0$.

A **Hilbert space** is a complete inner product space—meaning that every Cauchy sequence of vectors converges to a vector in the space. This completeness property, while essential for rigorous analysis in infinite dimensions, will not concern us in the finite-dimensional spaces ($\mathbb{C}^n$) used for qubit systems, where every inner product space is automatically complete.

Why Hilbert Space?

The choice of Hilbert space as the mathematical setting for quantum mechanics is not arbitrary but dictated by physical requirements. Consider what we need to describe quantum systems:

Complex amplitudes: Quantum interference requires probability amplitudes that can be positive or negative, and more generally, complex-valued. The complex vector space $\mathbb{C}^n$ provides exactly this structure.

Superposition: Quantum states combine linearly—if $|\psi_1\rangle$ and $|\psi_2\rangle$ are valid states, so is any (normalized) linear combination $\alpha\,|\psi_1\rangle + \beta\,|\psi_2\rangle$. Vector spaces naturally accommodate this superposition principle through their closure under addition and scalar multiplication.

Measurement probabilities: The Born rule requires computing $|\langle\phi|\psi\rangle|^2$ to obtain transition probabilities. The inner product provides exactly this capability.

Orthogonality of distinct outcomes: Mutually exclusive measurement outcomes correspond to orthogonal quantum states, where $\langle\phi|\psi\rangle = 0$. The inner product enables defining and testing orthogonality.

Unitarity and norm preservation: Quantum evolution must preserve total probability, requiring transformations that preserve inner products. Unitary operators provide exactly this structure: $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle$.

Observable values: Physical observables correspond to Hermitian operators, whose eigenvalues (measurement outcomes) are guaranteed to be real. The spectral theorem, valid in Hilbert space, ensures such operators can be diagonalized with orthonormal eigenvectors.

Hilbert Space in Quantum Computing

For quantum computing with n qubits, the relevant Hilbert space is $\mathbb{C}^{2^n}$, the space of column vectors with 2^n complex components. A single qubit inhabits $\mathbb{C}^2$, spanned by the computational basis $\{|0\rangle, |1\rangle\}$. Two qubits require $\mathbb{C}^4$, with basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ formed by tensor products of single-qubit states. The exponential growth of Hilbert space dimension with qubit count—2^n dimensions for n qubits—underlies both the power and the challenge of quantum computing: quantum computers can manipulate exponentially large state spaces, but this same exponential scaling makes classical simulation intractable for even modest qubit counts.

Quantum gates are unitary matrices acting on this Hilbert space. A single-qubit gate is a 2×2 unitary matrix; a two-qubit gate is 4×4; an n-qubit gate is $2^n \times 2^n$. The algebra of quantum circuits reduces to matrix multiplication in Hilbert space, making the mathematical machinery developed in this chapter directly applicable to quantum algorithm design and analysis.

Chapter Synthesis and Looking Ahead

This chapter has constructed the mathematical foundation for quantum computing from the ground up. We began with simple state representations using vectors, took care to distinguish classical statistical mixtures from quantum superpositions, and motivated the necessity of complex amplitudes through interference. We then developed the machinery of linear algebra—matrices, inner products, eigenvalues, and spectral decomposition—expressed throughout in Dirac notation. The Pauli matrices illustrated how abstract mathematics connects to physical observables like electron spin. Hermitian operators emerged as the representatives of measurable quantities, with real eigenvalues and orthogonal eigenvectors, while unitary operators describe reversible quantum evolution. All these elements find their natural home in Hilbert space, the mathematical arena where quantum mechanics achieves full rigor.

The next chapter builds upon this foundation by presenting quantum mechanics itself—its historical development, its dynamical equations, and its postulates: the fundamental principles that govern how quantum systems behave, evolve, and respond to measurement. Where this chapter answered "what is the mathematical language of quantum mechanics?", the next addresses "what are the physical rules by which quantum systems operate?" Together, they provide the complete conceptual and technical foundation for understanding and implementing quantum algorithms.

Every concept introduced in this chapter will be used repeatedly: tensor products combine qubits into composite systems; unitary matrices implement quantum gates; spectral decomposition enables the analysis of observables and evolutions; the Born rule determines measurement probabilities. The investment in mathematical formalism made here is practical preparation for reading quantum circuits, analyzing quantum algorithms, and ultimately programming quantum computers: quantum computing is, at its core, applied linear algebra in complex Hilbert space.

Summary

- **Matrix Multiplication**: Given $\mathbf{A}$ and $\mathbf{B}$, the product $\mathbf{AB} = \mathbf{C}$ is defined as $C_{ik} = \sum_j A_{ij} B_{jk}$, where dimensions must satisfy $[n \times m] \cdot [m \times l] = [n \times l]$. It is associative and distributive, but in general not commutative.

- **Conjugate Transpose (Adjoint)**: The conjugate transpose of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^\dagger = \overline{\mathbf{A}}^T$, and satisfies $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$.

- **Dirac Notation**:
  - A state vector $|\psi\rangle$ has its corresponding dual (bra) $\langle\psi| = |\psi\rangle^\dagger$.
  - The inner product between two states is $\langle\psi|\phi\rangle$; it is linear in the ket and antilinear in the bra.

- **Orthogonality**: Two vectors are orthogonal if $\langle\psi|\phi\rangle = 0$.

- **Vector Norm**: $\| \, |\psi\rangle \, \| = \sqrt{\langle\psi|\psi\rangle}$; quantum states are normalized, $\| \, |\psi\rangle \, \| = 1$.

- **Orthonormal Bases**: A set of vectors $\{|v_i\rangle\}$ is orthonormal if $\langle v_i|v_j\rangle = \delta_{ij}$. The computational basis of $\mathbb{C}^n$ is $\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\}$.

- **Outer Product**: $|v\rangle\langle u|$ is an operator, acting as $(|v\rangle\langle u|)\,|w\rangle = \langle u|w\rangle\,|v\rangle$. Projectors have the form $\mathbf{P} = |\psi\rangle\langle\psi|$ with $\mathbf{P}^2 = \mathbf{P}$.

- **Eigenvalues and Eigenvectors**: For a matrix $\mathbf{A}$, an eigenvalue λ and its eigenvector $|v\rangle \neq 0$ satisfy $\mathbf{A}\,|v\rangle = \lambda\,|v\rangle$; the eigenvalues are the roots of $\det(\mathbf{A} - \lambda I) = 0$.

- **Hermitian Matrices**: $\mathbf{A}^\dagger = \mathbf{A}$. Their eigenvalues are real and eigenvectors with distinct eigenvalues are orthogonal; they represent observables.
- **Normal Matrices**: $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$. Exactly the normal matrices admit a spectral decomposition $\mathbf{A} = \sum_i \lambda_i |v_i\rangle\langle v_i|$ with orthonormal $|v_i\rangle$.
- **Unitary Matrices**: $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$. They preserve inner products and norms, are reversible, and represent quantum gates and temporal evolution.
- **Tensor Product**: Combines subsystems into composite systems; dimensions multiply:

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \otimes \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 y_0 \\ x_0 y_1 \\ x_1 y_0 \\ x_1 y_1 \end{bmatrix}$$

1.6 Exercises

1. **Complex Number Operations**

   Given the complex numbers $z_1 = 3 + 4i$ and $z_2 = 1 - 2i$:

   (a) Calculate $z_1 + z_2$ and $z_1 \cdot z_2$
   (b) Find the complex conjugates z_1^* and z_2^*
   (c) Compute the moduli $|z_1|$ and $|z_2|$
   (d) Express z_1 and z_2 in polar form $re^{i\theta}$

2. **Born Rule and Probability Calculation**

   A particle is in a quantum state described by the wavefunction:

$$\psi(x) = \begin{cases} A\sin\left(\frac{\pi x}{L}\right) & 0 \leq x \leq L \\ 0 & \text{otherwise} \end{cases}$$

   (a) Determine the normalization constant A such that $\int_0^L |\psi(x)|^2 dx = 1$
   (b) Calculate the probability of finding the particle in the interval $[0, L/4]$
   (c) Find the most probable position (where $|\psi(x)|^2$ is maximum)

3. **Normalization of Quantum States**

   Consider the following quantum states:

$$|\psi_1\rangle = 2|0\rangle + 3|1\rangle, \quad |\psi_2\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}i|1\rangle$$

   (a) Check if $|\psi_1\rangle$ is normalized. If not, find the normalized version $\left|\tilde{\psi}_1\right\rangle$
   (b) Verify that $|\psi_2\rangle$ is normalized

(c) For $|\psi_1\rangle$, calculate the measurement probabilities for outcomes $|0\rangle$ and $|1\rangle$

4. **Inner Product and Dirac Notation**

   Given vectors $|\psi\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix}$ and $|\phi\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}i\\1\end{bmatrix}$, calculate the inner product $\langle\psi|\phi\rangle$ and verify if the vectors are orthogonal.

5. **Tensor Product**

   Calculate the tensor product of vectors $|u\rangle = \begin{bmatrix}1\\0\end{bmatrix}$ and $|v\rangle = \begin{bmatrix}0\\1\end{bmatrix}$. How can this result be interpreted in the context of qubits?

6. **Matrix Operations and Conjugate Transpose**

   Given the matrix $A = \begin{bmatrix}1+i & 2\\3i & 1-i\end{bmatrix}$:

   (a) Calculate A^* (complex conjugate)
   (b) Calculate A^T (transpose)
   (c) Calculate $A^\dagger$ (conjugate transpose)
   (d) Verify whether $AA^\dagger = A^\dagger A$

7. **Pauli Matrices: Eigenvalues and Physical Interpretation**

   Consider the Pauli matrix $\sigma_x = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$:

   (a) Find the eigenvalues and normalized eigenvectors of σ_x, following the procedure used for σ_y in the text
   (b) Verify that the eigenvectors are orthogonal
   (c) Show that $\sigma_x^2 = I$ (the identity matrix)
   (d) Apply σ_x to the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. What do you observe?

8. **Outer Product and Projection Operators**

   Consider the computational basis states $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}$ and $|1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$:

   (a) Construct the projection operators $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$
   (b) Verify that $P_0^2 = P_0$ and $P_1^2 = P_1$ (idempotence)
   (c) Show that $P_0 + P_1 = I$
   (d) Apply P_0 to the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Interpret the result

9. **Hermitian Matrix**

   Verify if matrix $\mathbf{A} = \begin{bmatrix}4 & 1+i\\1-i & 4\end{bmatrix}$ is Hermitian. If it is, find its eigenvalues.

10. **Eigenvalues and Eigenvectors**

    Find the eigenvalues and eigenvectors of the Hermitian matrix $\mathbf{A} = \begin{bmatrix} 2 & -i \\ i & 3 \end{bmatrix}$. Verify if the eigenvectors are orthogonal.

11. **Unitary Matrix Verification**

    Consider the Hadamard gate $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$:

    (a) Verify that H is unitary by showing $HH^\dagger = I$

    (b) Show that H is also Hermitian (i.e., $H = H^\dagger$)

    (c) Apply H to states $|0\rangle$ and $|1\rangle$. What states do you obtain?

    (d) Verify that H preserves the norm: if $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$, show that $H|\psi\rangle$ is also normalized

12. **Spectral Decomposition**

    Given matrix $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, perform the spectral decomposition of $\mathbf{A}$. Express $\mathbf{A}$ as a sum of orthogonal projectors weighted by their eigenvalues.

13. **Commutators and Uncertainty Relations**

    (a) Calculate the commutator $[\sigma_x, \sigma_y]$ of Pauli matrices

    (b) Show that $[\sigma_x, \sigma_y] = 2i\sigma_z$

    (c) Verify the cyclic property: $[\sigma_y, \sigma_z] = 2i\sigma_x$ and $[\sigma_z, \sigma_x] = 2i\sigma_y$

    (d) What does a non-zero commutator imply about the simultaneous measurement of these observables?

14. **Operator Commutator**

    Consider operators $\hat{A} = \hat{x}^2$ and $\hat{B} = \frac{d}{dx}$ acting on differentiable functions. Calculate the commutator $[\hat{A}, \hat{B}]$ and determine if they commute.

15. **Multi-Qubit Systems and Entanglement**

    Consider the two-qubit state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ (Bell state):

    (a) Express $|\Psi\rangle$ as a vector in $\mathbb{C}^4$

    (b) Show that this state cannot be written as a tensor product $|\psi\rangle \otimes |\phi\rangle$ for any single-qubit states $|\psi\rangle$ and $|\phi\rangle$ (proving it is entangled)

    (c) Calculate the probability of measuring $|00\rangle$ and $|11\rangle$

    (d) If we measure the first qubit and obtain $|0\rangle$, what is the state of the second qubit?