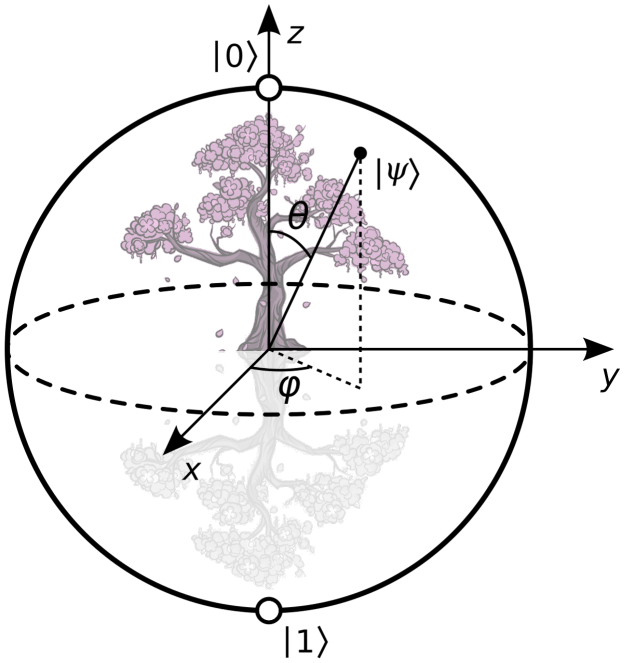
Quantum Information and Computation

2026 Draft

Chinmay Nirkhe and Darin Ershov



University of Washington | CSE 534 Autumn 2026

Acknowledgment

This initial draft was written by Chinmay Nirkhe and Darin Ershov from generative transcription of handwritten notes by Chinmay Nirkhe (using ChatGPT 5.6-Sol). As such, it may have significant errors that we are trying to rectify. Please notify the authors of mistakes or inaccuracies in order to improve the quality of the presentation.

The intention of these notes is to be the first step on a multi-year transition of this course from handwritten to digitized notes. During the year, we will also make the handwritten notes available. Over time, we hope to completely digitize the content of the course till it can stand on its own.

Contents

Chapter 0

Preface

0.1 Computation is a physical process

We start with a simple observation: computation is both a mathematical abstraction and a physical process. The premise of a "classical computer" is that it stores information in physical systems that can be reliably distinguished and manipulated. Present-day digital hardware usually utilizes transistors, but one could equally imagine bits encoded in coins, mechanical switches, DNA, or any other medium with multiple robust states. What matters at the level of computation is that there is a finite set of basic states, which for binary computers are 0 and 1, and that we can apply reliable rules to strings of these states to manipulate the information. This is what we call computation.

Accordingly, classical computation is the manipulation of classical information according to the laws of classical physics. The Church–Turing thesis [1, 2] captures the robustness of this idea: it posits that all "reasonable" classical models of computation appear to be equivalent up to polynomial overhead. This is why the study of algorithms is agnostic to the underlying machine: be it built from vacuum tubes, silicon transistors, or simply pencil-and-paper simulations. This implementation-agnostic approach to computation has been essential to our development of a unified theory of computation.

Quantum computation starts by asking what changes when the information-bearing systems are governed by quantum mechanics. A quantum bit, or *qubit*, has basis states $|0\rangle$ and $|1\rangle$, but it can also be in a superposition

$$\alpha\,|0\rangle + \beta\,|1\rangle\,, \qquad \alpha, \beta \in \mathbb{C}, \qquad |\alpha|^2 + |\beta|^2 = 1. \tag{0.1}$$

For n qubits, a general pure state has the form

$$|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x\,|x\rangle\,, \qquad \sum_{x\in\{0,1\}^n} |\alpha_x|^2 = 1. \tag{0.2}$$

The notation is compact and, in a way, hides a striking fact: the state vector is 2^n complex amplitudes—the mathematical description of a generic n-qubit system grows exponentially with n. This exponential growth is not, by itself, a source of quantum advantage as one cannot simply "read out" all 2^n amplitudes. As we will see, the axioms of quantum mechanics say that we can only access the quantum mechanical state through measurement, a process which only returns ordinary classical data. Instead, this course will show

that the power of quantum computation comes instead from the ability to arrange amplitudes so that wrong computational paths interfere destructively, and right paths interfere constructively.

At the same time, this exponential state space is the first reason quantum computation is a compelling subject. If nature itself uses quantum states, then a classical computer that simulates a generic quantum system appears to face an exponential bookkeeping problem. Physicists have developed many powerful workarounds: perturbation theory, mean-field approximations, Feynman diagrams, density functional methods, tensor networks, Monte Carlo techniques, and other tools for special regimes. These methods are indispensable, but they do not amount to a general efficient simulation of arbitrary quantum dynamics. The question that excites the field of quantum computation is whether a device built from quantum systems can simulate, communicate, and compute in ways that no efficient classical device can match.

0.2 Feynman's challenge

The modern field of quantum computation is often traced to Richard Feynman's 1981 keynote lecture at the first Physics of Computation Conference at MIT Endicott House, the notes of which were later published as "Simulating Physics with Computers" [3]. In this lecture, Feynman emphasized that the difficulty of simulating quantum physics was not just due to a lack of existing algorithms; it reflected a fundamental mismatch between the classical nature of the simulator and the quantum nature of the system being simulated. His central thesis is summarized and (constantly quoted) as:

> *Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical.*

This isn't meant to be a rhetorical statement. Rather, if quantum mechanics is the best microscopic theory of nature, then perhaps the only universal simulators for nature's description are quantum mechanical. This idea is a good place to segue into the role of quantum mechanics in this course. In a standard physics course, quantum mechanics is described as a theory to be solved: given a Hamiltonian, predict measurement outcomes. In our course on quantum computation and information, quantum mechanics is studied as a resource. Superposition, entanglement, measurement disturbance, and interference are not only strange/spooky phenomena; they are ingredients for information processing. The course will repeatedly return to this change in viewpoint. We will isolate a small set of operational axioms—states, measurements, unitary transforms—and use them as the rules of a new model of computation. Then, we will see how far this can take us.

0.3 Early history: from thought experiment to algorithms

Feynman's proposal was motivated by quantum simulation, but the field quickly broadened to other algorithmic questions. David Deutsch formulated a quantum analogue of the universal Turing machine and argued for a quantum version of the Church-Turing principle [4]. Bennett and Brassard introduced quantum key distribution, showing that quantum mechanics could be used not only to compute but also to communicate with security guarantees of a different flavor from classical cryptography [5]. Bernstein and Vazirani gave a formal complexity-theoretic framework for quantum computation and oracle problems,

showing separations between quantum and classical query models [6]. Simon then gave an oracle problem with an exponential quantum speedup, a result that directly influenced the next major breakthrough [7].

That breakthrough was Shor's algorithm. Classical public-key cryptosystems such as RSA rely on computational asymmetry: multiplying two large primes is easy, but factoring their product is believed to be hard for classical computers [8]. Shor showed that a quantum computer can factor integers and compute discrete logarithms in time polynomial in the input length [9]. The result was a turning point because factoring is not merely an artificial oracle problem; it is a concrete, heavily studied problem with direct cryptographic significance. We will see each of these ideas to some extent in this course.

Shor's result forces a useful trilemma. If large, reliable quantum computers can be built, then either quantum mechanics fails at the relevant scale, an efficient classical factoring algorithm exists but has not been discovered, or quantum computers are genuinely more powerful than classical computers for some natural computational tasks. The third possibility is the working hypothesis of most in the field of quantum computation, but it remains tied to some of the deepest open questions in complexity theory. We do not know how to prove, unconditionally, that (in complexity-theoretic terms) BQP is outside BPP, or even how exactly quantum polynomial time relates to many classical complexity classes. Quantum computing is therefore both a model for algorithms and a testing ground for our understanding of computational hardness.

0.4 Quantum information, not only quantum algorithms

Although algorithms such as Shor's are central, quantum information is broader than the study of algorithmic speedups. We shall see that our axioms of quantum mechanics/information change the basic rules, and new protocols, previously impossible, are now unlocked. For example, unknown quantum states cannot be cloned [10], measurements can irreversibly disturb a system, and entanglement produces correlations that cannot be explained by shared classical randomness [11]. To clarify, these are not a new notion of "information theory, " and classical information theory is still present: probability, entropy, coding, communication, and computational complexity remain fundamental. But in the quantum setting these ideas are modified by mathematical formalisms such as linearity, tensor products, and non-commuting measurements.

0.5 The field today

The current state of quantum computing is best described as in the intermediate and noisy regime. Experimental devices are far beyond the few-qubit demonstrations of the 1990s and early 2000s. There are multiple serious hardware platforms, such as superconducting circuits, trapped ions, neutral atoms, photonics, and more speculative topological approaches. These platforms differ in connectivity, coherence, gate speed, control complexity, fabrication methods, and error mechanisms. At the level of idealized computation, these platforms aim to realize computational models equivalent to the standard quantum circuit model, although their physical overheads and architectures may differ substantially. There is no consensus that a single architecture has already won. In that sense, the field resembles an early stage of classical computing, when several physical implementations of computation competed before silicon transistors became dominant.

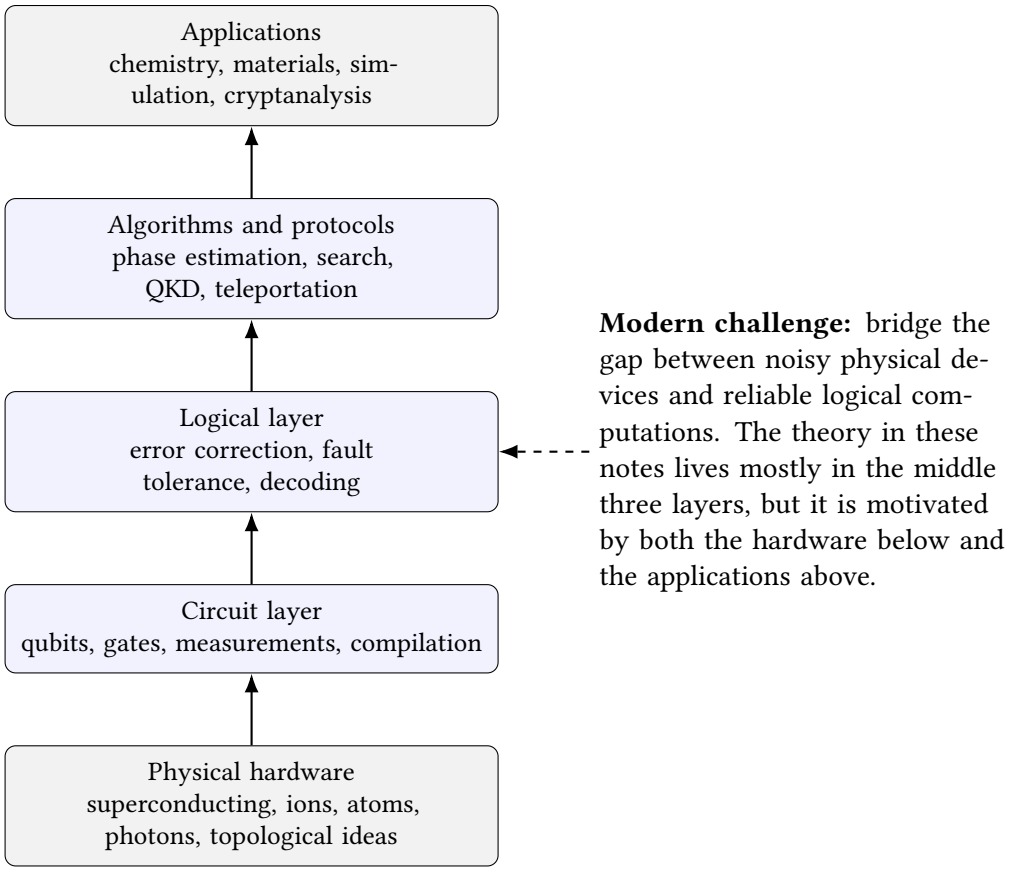


Figure 1: A schematic quantum-computing stack. These notes are implementation must be provided by hardware, error correction, algorithms, and applications.

Present devices are noisy. They can run nontrivial circuits, but noise, imperfect gates, measurement errors, leakage, calibration drift, and limited connectivity restrict circuit depth. This is the regime that Preskill called noisy intermediate-scale quantum computing, or NISQ [12]. NISQ devices are large enough to challenge brute-force classical simulation in some regimes, but not yet reliable enough to run long computations such as large-scale integer factoring or fully fault-tolerant quantum chemistry. The central practical question is therefore not only "How many physical qubits do we have?" but "How reliably can we control them, and can we turn many noisy physical qubits into fewer reliable logical qubits?"

This is why quantum error correction and fault tolerance occupy such a central place in modern quantum computing and why we explore them as the final chapter of this course. Quantum information is fragile, and the no-cloning theorem prevents the most naïve form of redundancy. Nevertheless, quantum error-correcting codes encode logical information into entangled states of many physical qubits [13]. The long-term vision is a fault-tolerant quantum computer whose logical error rates can be made small enough for algorithms of practical scale [14]. Recent experiments on multiple platforms have begun to demonstrate logical qubits, error-corrected operations, and below-threshold behavior in surface-code memories [15, 16].

Theory is evolving in parallel with these engineering advances. Researchers (including some of us at UW) continue to search for new quantum algorithms and for improved versions of known algorithms with lower overhead. We study quantum complexity theory to understand which problems are likely to

remain hard even for quantum computers. We design state-of-the-art error-correcting codes and study classical simulation algorithms to identify the boundary between classically tractable and truly quantum behavior. We use quantum computation as a language for other areas of science, including condensed matter physics, high-energy physics, quantum gravity, and cryptography, and we are starting to run rudimentary experiments on the aforementioned noisy quantum devices.

Chapter 1

Axioms of quantum computation

1.1 The two halves of quantum computation

Quantum computation has two related, but distinct, halves:

1. The first half is the theoretical model of quantum computation. This is the model that we will study in this course. It is an idealized, consistent, and plausible description of reality. It is also implementation-agnostic: the model does not commit to a particular physical substrate. Instead, it gives us a clean mathematical testbed for algorithm design and for proving impossibility results.

2. The second half is the study of quantum computers, i.e., implementations of quantum computation. This includes physical platforms such as ion traps, photonic systems, neutral atoms, and other architectures. These are models in which we hope to simulate the idealized quantum computation model. They will not be the main focus of this course, but they are pertinent to the study and evolution of our model of quantum computation.

The goal of this chapter is to fix the mathematical model of quantum computation. A quantum computation is described by states, transformations, and measurements. The axioms below tell us what these words mean and allow us to discuss quantum computation without having to truly understand quantum mechanics or subscribe to any one physical implementation of a quantum computer.

We will start by introducing a set of axioms for describing one qubit, then we will evolve them to describe multiple qubits, and then we will construct a mathematically equivalent description that allows us to describe probability distributions over quantum systems more effectively.

1.2 Bra-ket notation

Quantum states are vectors in a complex vector space. Dirac's bra-ket notation [17] is a compact notation for vectors, dual vectors, inner products, and outer products. It is used throughout quantum information due to its physics origins, but we hope you will also find that it aligns the linear algebra with the operational meaning of the objects being described.

Definition 1.1 (Kets and bras)

A *ket* is a column vector, written as $|\psi\rangle$. The corresponding *bra* is the conjugate transpose of this vector, written as $\langle\psi| = |\psi\rangle^\dagger$. Thus, if

$$|\psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{d-1} \end{pmatrix} \in \mathbb{C}^d, \text{ then } \langle\psi| = \begin{pmatrix} \alpha_0^* & \alpha_1^* & \cdots & \alpha_{d-1}^* \end{pmatrix} \tag{1.1}$$

where α^* is the complex conjugate of $\alpha \in \mathbb{C}$.

The inner product between two states $|\phi\rangle$ and $|\psi\rangle$ is written as $\langle\phi|\psi\rangle$. This is the same object that would be written in matrix notation as

$$\langle\phi|\psi\rangle = |\phi\rangle^\dagger \cdot |\psi\rangle . \tag{1.2}$$

Note that the order matters: the expression is conjugate-linear in the first argument and linear in the second argument.

Example 1.2 (Inner product)

Let

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \text{and} \qquad |\phi\rangle = \gamma|0\rangle + \delta|1\rangle . \tag{1.3}$$

Then

$$\langle\phi|\psi\rangle = \gamma^*\alpha + \delta^*\beta. \tag{1.4}$$

In particular, the norm of $|\psi\rangle$ is determined by

$$\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{|\alpha|^2 + |\beta|^2} \tag{1.5}$$

A vector $|\psi\rangle$ satisfying $\||\psi\rangle\| = 1$ is called a unit vector.

Bra-ket notation also describes linear operators. The expression $|\phi\rangle\langle\psi|$ is an outer product and can be equivalently expressed as

$$|\phi\rangle\langle\psi| = \begin{pmatrix} \gamma\alpha^* & \gamma\beta^* \\ \delta\alpha^* & \delta\beta^* \end{pmatrix} . \tag{1.6}$$

1.3 A single qubit

Armed with bra-ket notation, we start building up our understanding of quantum systems, beginning with a system of a single qubit. A *qubit*—the portmanteau of quantum and bit [18]—is a model generalizing a bit from classical information theory. In commonplace speak about quantum computation, you will hear people refer to a quantum computer by "how many qubits it has." Soon, we will have a good understanding of what that means and how it relates to the mathematical definition we present here.

In this course, we take an axiomatic approach to defining the manipulation of quantum information and computation. As such, we will take the following as our first axiom of quantum computation.

Axiom (One-qubit system)

A one-qubit system is one whose state can be expressed as a vector $|\psi\rangle \in \mathbb{C}^2$ of unit norm. When expressed as such, we refer to $|\psi\rangle$ as a *one-qubit quantum state*.

There are some particular quantum states that a one-qubit system can be in. Due to their prevalence, they have special names.

Definition 1.3 (Common quantum states)

Define the following two states (vocalized as "ket zero" and "ket one")

$$|0\rangle \stackrel{\text{def}}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle \stackrel{\text{def}}{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.7}$$

These are possible states of a one-qubit system. They are orthogonal and of unit norm and, therefore, form a basis for $\mathbb{C}^2$. Together, this basis is referred to as the *standard basis* or the *Z-basis*. **Additionally, if the state of a quantum system is either $|0\rangle$ or $|1\rangle$, we call the system "classical".**

Define the following two states (vocalized as "ket plus" and "ket minus")

$$|+\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad |-\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{1.8}$$

These are possible states of a one-qubit system. They are orthogonal and of unit norm and, therefore, form a basis for $\mathbb{C}^2$. Together, this basis is referred to as the *Hadamard basis* or the *X-basis*.

The two bases defined give us a convenient way of expressing any state as a linear combination of the basis elements.

Example 1.4

The states $|0\rangle$ and $|1\rangle$ can be expressed as linear combinations of $|+\rangle$ and $|-\rangle$:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) = \frac{|+\rangle + |-\rangle}{\sqrt{2}}, \tag{1.9a}$$

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle) = \frac{|+\rangle - |-\rangle}{\sqrt{2}}. \tag{1.9b}$$

Example 1.5

The state of any one-qubit quantum system can be expressed, without loss of generality, as

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \text{ where } |\alpha|^2 + |\beta|^2 = 1. \tag{1.10}$$

Additionally, we can express $|\psi\rangle$ in terms of the Z-basis and X-basis as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \left(\frac{\alpha+\beta}{\sqrt{2}}\right)|+\rangle + \left(\frac{\alpha-\beta}{\sqrt{2}}\right)|-\rangle. \tag{1.11}$$

1.4 Transformation of a single qubit

How does the state of a quantum system evolve? We will need another axiom to describe transformations of quantum states.

Axiom

Let U be a unitary matrix $\in \mathbb{C}^{2\times 2}$. Then we can transform a one-qubit system by U, meaning that if the state of the system is currently $|\psi\rangle$, after transformation, the state of the system is now $U|\psi\rangle$.

Colloquially, we call this *applying* U.

For this axiom, it is helpful to recall the definition of a unitary matrix:

Definition 1.6 (Unitary matrix)

A matrix $U \in \mathbb{C}^{d\times d}$ is a *unitary matrix* if any of the following equivalent definitions are satisfied.

1. $U^\dagger U = I$.
2. $U^{-1} = U^\dagger$.
3. The columns of U form an orthonormal basis of $\mathbb{C}^d$.

4. The rows of U form an orthonormal basis of $\mathbb{C}^d$.

5. U preserves norms: for all $|\psi\rangle \in \mathbb{C}^d$,

$$\|U|\psi\rangle\| = \||\psi\rangle\|. \tag{1.12}$$

6. U maps every orthonormal basis of $\mathbb{C}^d$ to another orthonormal basis.

Observe that this transformation maps quantum states to quantum states as the dimensionality of the quantum state and the norm are preserved. Second, this axiom only describes the valid transformations on a one-qubit system; it does not describe whether this transformation is efficient. This will come up later when we discuss the computational complexity of manipulating quantum information.

Definition 1.7 (Common unitaries)

Similar to the special states of states, $|0\rangle, |1\rangle, |+\rangle$, and $|-\rangle$, there are some special unitaries worth defining:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \; H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{1.13}$$

.

Observe that the Z-basis: $\{|0\rangle, |1\rangle\}$ are the eigenvectors of Z and the X-basis: $\{|+\rangle, |-\rangle\}$ are the eigenvectors of X. Second, $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$; therefore, H is the unitary that maps the two eigenbases to each other.

Colloquially, the matrices X and Z are known as Pauli-X and Pauli-Z and the matrix H is called the *Hadamard transform.*

1.5 Measurements of a single qubit

Currently, we have described the state of a one-qubit quantum system and how to transform the system unitarily. We also need to describe a classical method for extracting information from a quantum system. This is called *measurement* or *Born's rule* [19].

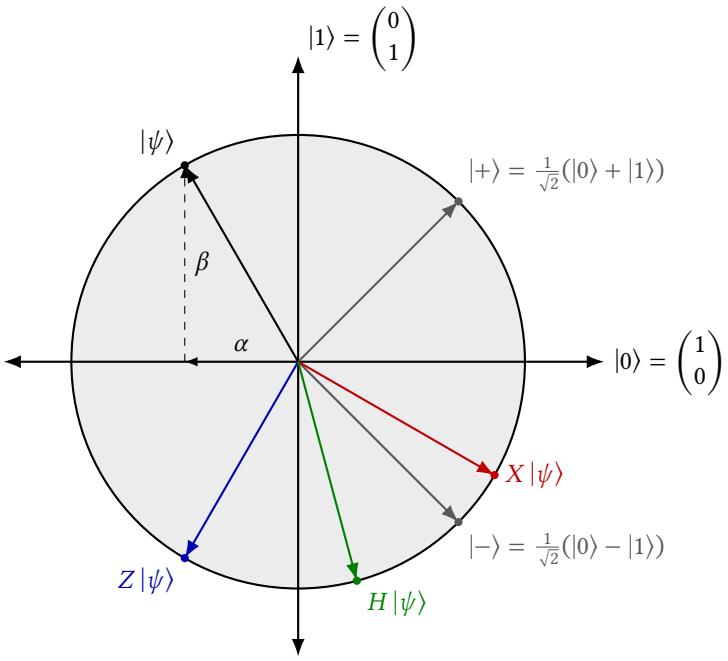


Figure 1.1: A cartoon of the transforms X, Z, H on the state $|\psi\rangle = \cos\frac{2\pi}{3}|0\rangle + \sin\frac{2\pi}{3}|1\rangle$. Observe, that X reflects across $|+\rangle$, Z reflects across $|0\rangle$.

Axiom (Measurement)

Given a quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, we can *measure* the quantum state which transforms the state according to the following rule:

- with probability $|\langle 0|\psi\rangle|^2 = |\alpha|^2$, the state is now $|0\rangle$,
- and with probability $|\langle 1|\psi\rangle|^2 = |\beta|^2$, the state is now $|1\rangle$.

In addition, the classical value "0" or "1" is output in the two scenarios, respectively, and is called the measurement outcome.

This transformation is inherently probabilistic and is a source of randomness/entropy. Let us observe that since $|\psi\rangle$ is of unit norm, this transformation is well-defined as the probabilities add to 1. Second, this gives credence to the definition of $|0\rangle$ and $|1\rangle$ as a "classical" state: Observe that if either $|0\rangle$ or $|1\rangle$ is measured, then the state does not change as one of the two probabilities is equal to 1. The name "classical" comes from the fact that, like classical objects, measurement/observation of the system does not change the state.

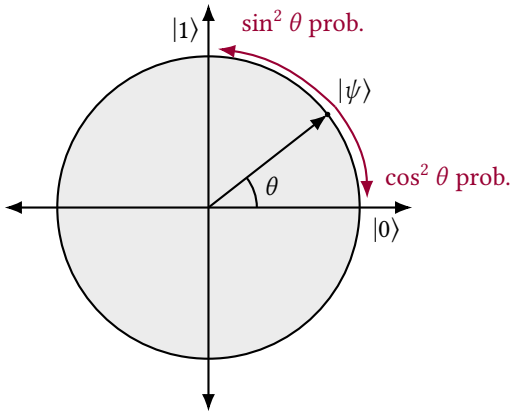


Figure 1.2: Measurement in the Z-basis for a real one-qubit state $|\psi\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$. The angle θ is measured from the $|0\rangle$ axis to $|\psi\rangle$. The state collapses to $|0\rangle$ with probability $\cos^2\theta$ and to $|1\rangle$ with probability $\sin^2\theta$.

Lemma 1.8

Let $\theta \in [0, 2\pi)$. Then for any quantum state $|\psi\rangle \in \mathbb{C}^2$, the states $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ produce the same measurement outcome distribution.

Proof. Consider the probability of the measurement outcome is 0. For the state $e^{i\theta}|\psi\rangle$, the measurement axiom states this is

$$|\langle 0|\, e^{i\theta}|\psi\rangle|^2 = |e^{i\theta}\cdot\langle 0|\psi\rangle|^2 = |e^{i\theta}|^2\cdot|\langle 0|\psi\rangle|^2 = |\langle 0|\psi\rangle|^2\,. \tag{1.14a}$$

This proves that the probability of measuring 0 is the same for both states. A similar argument proves it for outcome 1. □

The term $e^{i\theta}$ is considered the *global phase* as its value does not affect measurements. For this reason, we could technically define quantum states in the set $\mathbb{C}^2/\sim$ where $|\psi\rangle \sim e^{i\theta}|\psi\rangle$ forms equivalence classes.

For our axioms, measurement is the *only* way to extract classical information from a quantum system. This will be one of the central challenges of dealing with quantum information—complex states might be generatable but our interaction with the quantum device will always be classical.

1.6 Random number generator

We need one additional axiom in order to generate something interesting.

Axiom (Initialization)

We can initialize a one-qubit system with a state as $|0\rangle \in \mathbb{C}^2$.

Algorithm 1.9 (Random Number Generator)

1. Initialize a quantum state as $|0\rangle$.
2. Apply the unitary H (i.e., the Hadamard transform).
3. Measure the quantum state for an output $b \in \{0, 1\}$.

Theorem 1.10

The random number generator generates a random bit with uniform probability $\frac{1}{2}$.

Proof. The state after the Hadamard transform is

$$|\psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}. \tag{1.15}$$

Calculate that

$$\langle 0|\psi\rangle = \frac{1}{\sqrt{2}}, \qquad \langle 1|\psi\rangle = \frac{1}{\sqrt{2}}. \tag{1.16}$$

Therefore, by Born's rule, the outcome probabilities of both 0 and 1 are exactly $\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$. □

A one-qubit system, however, is not enough to explore truly interesting quantum effects. We will need to generalize our axioms to multiple qubits. To do that, we first need to review some mathematics.

1.7 Tensor products

Before we can state the many-qubit axioms cleanly, we need the tensor product. The tensor product is the linear-algebraic operation that lets us define a quantum system comprised of two non-interacting quantum systems. In these notes we use the notation $\otimes$ for both matrices and vectors; this is consistent because vectors are just matrices with one column.

Definition 1.11 (Tensor product of matrices)

Let $A \in \mathbb{C}^{m_1 \times n_1}$ and $B \in \mathbb{C}^{m_2 \times n_2}$. The tensor product, also called the Kronecker product, is the matrix $A \otimes B \in \mathbb{C}^{m_1 m_2 \times n_1 n_2}$ obtained by replacing each entry A_{ij} by the block $A_{ij}B$:

$$A \otimes B \stackrel{\text{def}}{=} \begin{pmatrix} A_{11}B & A_{12}B & \cdots & A_{1n_1}B \\ A_{21}B & A_{22}B & \cdots & A_{2n_1}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m_1 1}B & A_{m_1 2}B & \cdots & A_{m_1 n_1}B \end{pmatrix}. \tag{1.17}$$

Lemma 1.12 (Useful tensor identities)

Assume the dimensions are compatible. Then

$$(A \otimes B)(C \otimes D) = AC \otimes BD\,; \tag{1.18}$$

the tensor product is linear in both arguments, so

$$A \otimes \left(\sum_i b_i B_i\right) = \sum_i b_i (A \otimes B_i), \qquad \left(\sum_i a_i A_i\right) \otimes B = \sum_i a_i (A_i \otimes B), \tag{1.19}$$

where a_i, b_i are scalars in $\mathbb{C}$. In addition,

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger \tag{1.20}$$

When A and B are invertible,

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}\,. \tag{1.21}$$

Proof. The first identity follows by checking entries on pure blocks: the row of A contracts with the column of C, while the B and D blocks multiply in the same order. Properties of matrix addition gives the second identity. The adjoint identity follows because conjugating and transposing the block matrix conjugates and transposes both the scalar coefficients from A and the blocks from B. The inverse identity follows by multiplying $(A \otimes B)(A^{-1} \otimes B^{-1})$ and using the first identity. □

Definition 1.13 (Tensor product of vector spaces)

Given two vector spaces V, W, define the space $V \otimes W$ to be the space spanned by tensor products of vectors $v \otimes w$ with $v \in V$ and $w \in W$. Equivalently,

$$V \otimes W = \left\{ \sum_i \alpha_i (v_i \otimes w_i) \,\middle|\, v_i \in V, w_i \in W, \alpha_i \in \mathbb{C} \right\}. \tag{1.22}$$

By linearity of the tensor product, if $\{v_i\}_{i\in[n]}$ forms a basis of V and $\{w_j\}_{j\in[m]}$ forms a basis of W, then $\{v_i \otimes w_j\}_{i\in[n],j\in[m]}$ forms a basis of $V \otimes W$.

It is important to note that this space $V \otimes W$ is *not* simply the set of simple tensors $v \otimes w$ but instead the space of sums of simple tensors. As we will see in more detail in Chapter 2, this is where entanglement comes from.

Definition 1.14 (Computational basis notation)

For $x = (x_1, \dots, x_n) \in \{0,1\}^n$, define

$$|x_1, \dots, x_n\rangle \stackrel{\text{def}}{=} |x_1\rangle \otimes \cdots \otimes |x_n\rangle \ . \tag{1.23}$$

We usually omit the commas and write this vector as $|x\rangle$. Equivalently, if $j \in \{0, \dots, 2^n - 1\}$ and $j_1 \cdots j_n$ is the binary expansion of j, then $|j\rangle$ means $|j_1\rangle \otimes \cdots \otimes |j_n\rangle$. The vectors $\{|x\rangle \ : \ x \in \{0,1\}^n\}$ form the computational basis of $(\mathbb{C}^2)^{\otimes n}$.

Example 1.15 (Tensoring basis states)

We have $|0\rangle \otimes |1\rangle \otimes |0\rangle = |0,1,0\rangle = |010\rangle$ and $|0\rangle \otimes |1\rangle \otimes |1\rangle = |0,1,1\rangle = |011\rangle$. In the standard vector representation of $\mathbb{C}^{2^3}$, these are the basis vectors indexed by the binary strings 010 and 011.

1.8 Multi-qubit quantum systems

The tensor product lets us describe composite systems. If qubit A is in state $|\psi\rangle_A$ and qubit B is in state $|\phi\rangle_B$, the noninteracting joint state is $|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\phi\rangle_B$. However, quantum computation also needs states that are not tensor products. These are the entangled states, and allowing multi-qubit unitary transformations is the mechanism by which the model creates them.

Axiom (Updated axioms for many-qubit systems)

The one-qubit axioms extend to many qubits as follows.

1. **States.** The state of an n-qubit system is a unit vector $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n} \cong \mathbb{C}^{2^n}$. In the computational basis, every state has the form

$$|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle , \tag{1.24}$$

   where $\sum_x |\psi(x)|^2 = 1$. A state is classical if exactly one computational-basis amplitude is nonzero, and it is in superposition if multiple computational-basis amplitudes are nonzero.

2. **Transformations.** For any unitary $U \in \mathbb{C}^{2^n \times 2^n}$, we can transform the state $|\psi\rangle$ to $U|\psi\rangle$.

3. **Measurement of the full state.** Measuring a state

$$|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle \tag{1.25}$$

   yields the classical output x with probability $|\alpha|^2$ and the post-measurement state is $|x\rangle$ for each $x \in \{0,1\}^n$.

4. **Measurement of the first qubit.** Instead of measuring the full state $|\psi\rangle$, we can measure individual portions of the full state. In particular, measuring the first qubit in the computational basis sends

$$|\psi\rangle = \alpha_0 |0\rangle \otimes |\psi_0\rangle + \alpha_1 |1\rangle \otimes |\psi_1\rangle , \tag{1.26}$$

where $\| |\psi_0\rangle \| = \| |\psi_1\rangle \| = 1$, to

$$|j\rangle \otimes |\psi_j\rangle \tag{1.27}$$

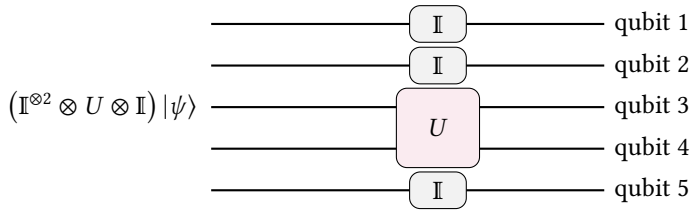
with probability α_j and outputs the classical value j for each $j \in \{0, 1\}$.

5. **Initialization.** Given an n-qubit state $|\psi\rangle$, we can initialize a fresh qubit as $|0\rangle$, producing the unentangled $(n+1)$-qubit state $|\psi\rangle \otimes |0\rangle$.

The transformation axiom as written only allows a gate to act on adjacent qubits. This restriction is not a weakness for the usual circuit model: adjacent one- and two-qubit gates already give a universal model once we know how to move information around with swap operations [20]. The problem set develops this point further.

Definition 1.16 (k-qubit unitaries)

A unitary $U \in \mathbb{C}^{2^k \times 2^k}$ is called a k-qubit unitary. Thus a 2×2 unitary is a single-qubit unitary and a 4×4 unitary is a two-qubit unitary.



U acts only on the adjacent block of qubits 3 and 4

Figure 1.3: A local two-qubit unitary written as a quantum circuit (we will quantum circuits in detail in chapter 5). The horizontal wires are qubits; the box U acts on the two wires it touches, while identities act on the untouched wires.

Example 1.17 (How a two-qubit gate acts on basis states)

Let U be a two-qubit unitary acting on qubits i and $i+1$. On a basis state $|x_1, \ldots, x_n\rangle$, the gate acts only on the two selected bits:

$$\left(\mathbb{I}_2^{\otimes(i-1)} \otimes U \otimes \mathbb{I}_2^{\otimes(n-i-1)}\right) |x_1, \ldots, x_n\rangle \tag{1.28a}$$

$$= |x_1, \dots, x_{i-1}\rangle \otimes U\,|x_i, x_{i+1}\rangle \otimes |x_{i+2}, \dots, x_n\rangle \tag{1.28b}$$

$$= |x_1, \dots, x_{i-1}\rangle \otimes \left(\sum_{k_1,k_2\in\{0,1\}} U_{(k_1,k_2),(x_i,x_{i+1})}\,|k_1, k_2\rangle \right) \otimes |x_{i+2}, \dots, x_n\rangle\,. \tag{1.28c}$$

By linearity, this formula determines the action of the same gate on any state $|\psi\rangle = \sum_{x\in\{0,1\}^n} \psi(x)\,|x\rangle$.

The matrix of a two-qubit gate can also be written using outer products:

$$U = \sum_{k_1,k_2,j_1,j_2\in\{0,1\}} U_{(k_1,k_2),(j_1,j_2)}\,|k_1, k_2\rangle\langle j_1, j_2|\,. \tag{1.29}$$

This notation emphasizes that each matrix entry is the amplitude for the input basis state $|j_1, j_2\rangle$ to be transformed into the output basis state $|k_1, k_2\rangle$.

Remark 1.18 (Joining and separating systems)

The ability to form $|\psi\rangle_A \otimes |\phi\rangle_B$ from separate systems and the ability to separate systems that are already unentangled are not extra axioms. They are consequences of the state-space and initialization axioms together with the tensor-product description of composite systems.

Remark 1.19 (Matrix manipulations)

We have provided the matrix entry-by-entry transformations in the previous example to be pedagogical. However, as you will see in this course, we will typically not deal with the entry-by-entry representation of quantum states or unitaries, but rather deal with them through symbol manipulation. This will be both more succinct and conceptually simpler.

Chapter 2

Introduction to quantum information via examples

The previous chapter fixed the axioms of quantum computation. We now use those axioms to build some small examples where the quantum model behaves fundamentally differently from a purely classical randomized model. These examples will be small and only involve a few qubits.

2.1 Measuring in a different basis

Before we define the first example, let us establish a few additional theorems that will make describing these examples significantly easier. The measurement axiom directly measures a single qubit in the standard basis $\{|0\rangle, |1\rangle\}$. In many situations, we would like to be able to measure in some other orthonormal basis $\{|b_0\rangle, |b_1\rangle\}$ of $\mathbb{C}^2$. Such a measurement does not require us defining an additional axiom. Rather, it follows from the previously stated measurement axiom and the ability to apply single-qubit unitaries.

This simple observation will recur throughout quantum computation. Although our measurement axiom singles out the computational basis, there is nothing fundamentally special about that basis: we can first rotate whichever information we wish to observe into the computational basis and then measure. For example, measuring in the $|+\rangle, |-\rangle$ basis amounts to applying a Hadamard transform before a standard-basis measurement. Later, Bell-basis measurements and many quantum algorithms will use exactly the same principle: apply a unitary that converts otherwise hidden information into a basis in which it can be read out.

Definition 2.1 (Single-qubit basis measurement)

Let $|b_0\rangle, |b_1\rangle \in \mathbb{C}^2$ be orthonormal. Measuring $|\psi\rangle$ in the basis $\{|b_0\rangle, |b_1\rangle\}$ means producing outcome $i \in \{0, 1\}$ with probability $|\langle b_i|\psi\rangle|^2$ and leaving the post-measurement state $|b_i\rangle$.

Lemma 2.2 (Implementing a basis measurement)

Every single-qubit basis measurement can be implemented by unitary evolution, a standard-basis measurement, and then another unitary evolution.

Proof. Define the matrix

$$U \stackrel{\text{def}}{=} |b_0\rangle\langle 0| + |b_1\rangle\langle 1| \ . \tag{2.1}$$

The columns of U are $|b_0\rangle$ and $|b_1\rangle$, which are orthonormal, so U is unitary. The algorithm is drawn as a quantum circuit in Figure 2.1: apply $U^\dagger$, measure in the standard basis, and then apply U to the post-measurement state. After applying $U^\dagger$, the probability of standard-basis outcome i is

$$\left|\langle i|\, U^\dagger\, |\psi\rangle\right|^2 = |\langle b_i|\psi\rangle|^2. \tag{2.2}$$

Conditioned on outcome i, the intermediate state is $|i\rangle$, and applying U sends this to $U|i\rangle = |b_i\rangle$. This is exactly the measurement in Definition 2.1. □

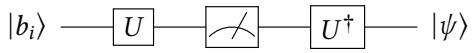


Figure 2.1: A measurement in the basis $\{|b_0\rangle, |b_1\rangle\}$ can be implemented by first applying $U^\dagger$, where $U|i\rangle = |b_i\rangle$, then measuring in the standard basis, and finally applying U. As throughout these notes, the circuit is read from right to left.

2.2 The Elitzur–Vaidman bomb tester

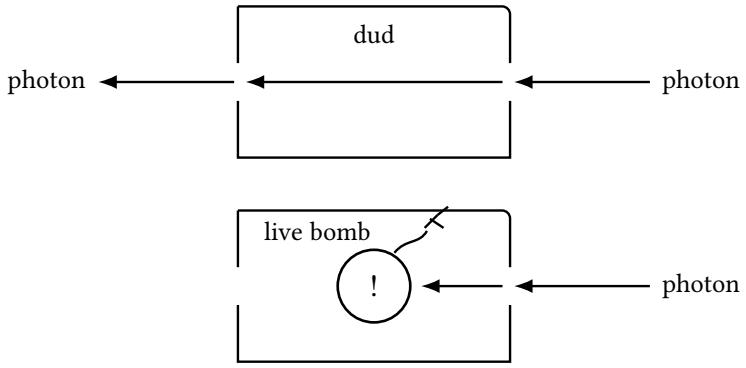


Figure 2.2: The promise in the bomb-testing thought experiment, drawn right to left. A photon passes straight through a dud. A live bomb is sensitive enough that a single photon entering the box will set it off.

The Elitzur–Vaidman bomb tester is a 1993 thought experiment designed to illustrate how differently quantum measurements can behave from ordinary classical observations [21]. For this thought experiment, imagine we have a purported bomb that we are trying to determine if it is real or if it is fake. In the thought experiment, the device is very temparmental and tamper-proof—if it truly is a live bomb, it is so sensitive that a single photon hitting it is enough to trigger an explosion. Whereas a dud lets a photon pass through harmlessly. Ideally, we want to come up with an experiment that determines whether the bomb is live or a

dud with the property that, in the case of a live bomb, the probability of an explosion is small. Indeed, we will see that for any sensitivity parameter of risk, ε, we want to tolerate, we will come up with a protocol which accurately determines if the bomb is live or a dud with at most an ε chance of explosion.

The two possibilities are illustrated in Figure 2.2. Unsurprisingly, we take an inherently quantum approach to the problem. Instead of simply sending a photon at the bomb and hoping for the best, we come up with a strategy that will place a photon in a superposition of two paths, one path that sends it through a safe passage and another path that sends it through the bomb, and then combine the

2.2.1 A two-path experiment

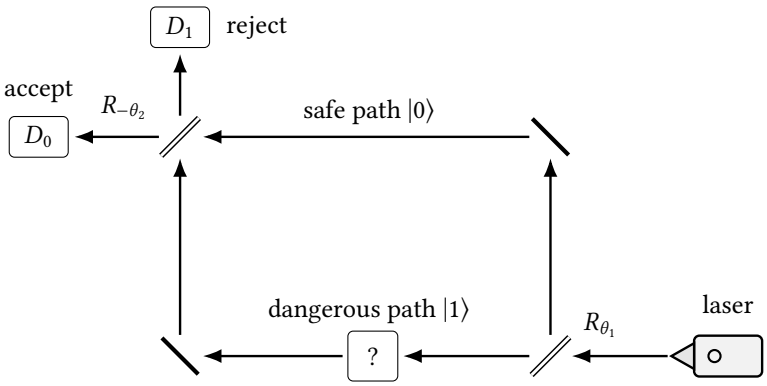


Figure 2.3: The one-shot bomb-testing experiment, drawn right to left. The first beam splitter creates a superposition of the safe and dangerous paths. The second recombines the paths before the photon reaches one of the two detectors.

In order to build the first iteration of the Elitzur–Vaidman bomb test, we are going to use a setup as drawn in Figure 2.3. In this experimental setup, we fire a laser that shoots one photon at a beam-splitter. The beam-splitter sends the photon in a superposition of the upper and lower paths—on the upper path, there is no device whereas the lower path has the purported bomb placed along the path. Based on the angle θ_1 of the beam-splitter, the photon will be in a superposition of the two paths as we discuss shortly. Next, there is a recombining beam-splitter at angle θ_2 before the photon hits one of the two detectors.

To model the experiment, we represent the photon's state as a single qubit whose basis states correspond to the two possible paths. The upper path is safe and is denoted by $|0\rangle$, while the lower path passes through the unknown object and is denoted by $|1\rangle$. The beam splitters are modeled as unitaries R_θ, which are rotation unitiaries, defined below. Then, we consider the two cases. If the bomb is a dud, then the photon will pass through in either case and reach the second beam-splitter. Whereas a live bomb behaves differently: it effectively measures whether the photon occupies the safe or dangerous path. If the outcome is 1, the bomb explodes. If the outcome is 0, the photon survives, and its state has collapsed to $|0\rangle$ as it can only take the safe path. Thus, we can reframe the problem as distinguishing

$$\text{dud} \longleftrightarrow \text{identity}, \tag{2.3a}$$

$$\text{live bomb} \longleftrightarrow \text{standard-basis measurement, with outcome 1 an explosion.} \tag{2.3b}$$

Playing with simple angles Before fully diving into a full analysis, let's consider some simple scenarios. Before using the full interferometer, consider a few simple probes. Sending $|0\rangle$ is perfectly safe but useless: both a dud and a live bomb return $|0\rangle$. Sending $|1\rangle$ goes to the opposite extreme. A dud returns $|1\rangle$, while a live bomb explodes with probability 1.

A more complex situation is to start by considering $\theta_2 = 0$, or equivalently the second beam-splitter doesn't exist, while leaving θ_1 a free variable. The beam-splitters are defined by rotation unitaries:

$$R_\theta \stackrel{\text{def}}{=} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{2.4}$$

Therefore, the state of the qubit immediately after the first beam-splitter is

$$R_{\theta_1} = \cos\theta_1 |0\rangle + \sin\theta_1 |1\rangle . \tag{2.5}$$

It is important that this is a superposition rather than a classical random choice of path. As long as no measurement reveals which path the photon occupies, the two branches may later be recombined and interfere.

Let's examine this particular case. In the case of a dud, the state measured is $|\theta_1\rangle$ and therefore,

$$\Pr[\text{accept} \mid \text{dud}] = \cos^2\theta_1, \tag{2.6a}$$
$$\Pr[\text{reject} \mid \text{dud}] = \sin^2\theta_1, \tag{2.6b}$$
$$\Pr[\text{explode} \mid \text{dud}] = 0. \tag{2.6c}$$

In the case of a bomb,

$$\Pr[\text{accept} \mid \text{bomb}] = \cos^2\theta_1, \tag{2.7a}$$
$$\Pr[\text{reject} \mid \text{bomb}] = 0, \tag{2.7b}$$
$$\Pr[\text{explode} \mid \text{bomb}] = \sin^2\theta_1. \tag{2.7c}$$

So we can see that we can increase the rejection probability in the case of a dud at the cost of increasing the explosion probability in the case of a bomb. Unfortunately, this increase is at a rate of 1, which we can verify is no better than probabilistically trying to send a photon (a classical strategy).

Second beam-splitter In order to gain improvement from quantum information, we need to utilize interference patterns. This is where the second beam-splitter comes in. If we allow the second beam-splitter to be a rotation by angle $-\theta_2$, equivalently, in the circuit model, we apply $R_{-\theta_2}$ followed by a standard-basis measurement, as shown in Figure 2.4. We call outcome 0 *accept* and outcome 1 *reject*; explosion is the third possible outcome, determined by if the "Box" is a standard-basis measurement or not. Call this the "one-shot bomb test".

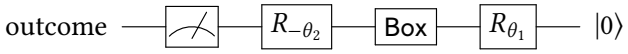


Figure 2.4: The circuit version of the one-shot bomb test, read from right to left. The unknown box is either the identity or the bomb measurement; an explosion terminates the experiment before the final measurement.

Lemma 2.3 (One-shot bomb test)

If the unknown object is a dud, then

$$\Pr[\text{accept} \mid \textit{dud}] = \cos^2(\theta_1 - \theta_2), \tag{2.8a}$$
$$\Pr[\text{reject} \mid \textit{dud}] = \sin^2(\theta_1 - \theta_2), \tag{2.8b}$$
$$\Pr[\text{explode} \mid \textit{dud}] = 0. \tag{2.8c}$$

If the unknown object is a live bomb, then

$$\Pr[\text{accept} \mid \textit{bomb}] = \cos^2\theta_1 \cos^2\theta_2, \tag{2.9a}$$
$$\Pr[\text{reject} \mid \textit{bomb}] = \cos^2\theta_1 \sin^2\theta_2, \tag{2.9b}$$
$$\Pr[\text{explode} \mid \textit{bomb}] = \sin^2\theta_1, \tag{2.9c}$$
$$\tag{2.9d}$$

Proof. Suppose first that the object is a dud. After the first beam-splitter, the state is $R_{\theta_1}|0\rangle$, and the dud does nothing. The second beam splitter therefore yields the state

$$R_{-\theta_2}R_{\theta_1}|0\rangle = R_{\theta_1-\theta_2}|0\rangle = \cos(\theta_1-\theta_2)|0\rangle + \sin(\theta_1-\theta_2)|1\rangle. \tag{2.10}$$

The accept and reject probabilities follow immediately from the Born rule. Now suppose the object is a live bomb. Immediately before reaching the bomb, the state is

$$\cos\theta_1|0\rangle + \sin\theta_1|1\rangle. \tag{2.11}$$

The bomb therefore explodes with probability $\sin^2\theta_1$. Conditioned on survival, the state is $|0\rangle$. The second beam splitter then produces

$$R_{-\theta_2}|0\rangle = \cos\theta_2|0\rangle - \sin\theta_2|1\rangle. \tag{2.12}$$

Thus, conditioned on survival, accept and reject have probabilities $\cos^2\theta_2$ and $\sin^2\theta_2$. Multiplying these by the survival probability $\cos^2\theta_1$ gives the claimed unconditional probabilities. □

Let's pick some choice of parameters for θ_1 and θ_2 that match our error tolerance. Suppose we are

willing to tolerate an explosion probability

$$\varepsilon \stackrel{\text{def}}{=} \sin^2 \theta_1. \tag{2.13}$$

Once θ_1 has been fixed, choosing θ_2 can be viewed as a small statistical distinguishing problem. In the dud case, the rejection probability is $\sin^2(\theta_1 - \theta_2)$, while in the bomb case, conditioned on survival, the rejection probability is $\sin^2 \theta_2$. Note that we care about the conditioned on survival properties, because these are the only probabilities the experimenter will witness.

Equivalently, we are choosing a test for distinguishing two biased coins. Later we will develop more systematic tools for choosing measurements that optimally distinguish two possible quantum states. For now, there is a particularly simple choice:

$$\theta_2 = \theta_1. \tag{2.14}$$

With this choice, a dud always accepts. Therefore, if detector D_1 in Figure 2.3 clicks, the object cannot have been a dud. For a live bomb, the three probabilities become

$$\begin{aligned} \Pr[\text{reject} \mid \text{bomb}] &= (1-\varepsilon)\varepsilon, \\ \Pr[\text{accept} \mid \text{bomb}] &= (1-\varepsilon)^2, \\ \Pr[\text{explode} \mid \text{bomb}] &= \varepsilon. \end{aligned} \tag{2.15}$$

The reject outcome is the phenomenon highlighted by Elitzur and Vaidman: it certifies the presence of a live bomb even though the bomb did not explode [21]. This is usually called an *interaction-free measurement*. The terminology should not be taken too literally—the possibility of interaction is essential to the experiment—but on the successful reject branch the bomb is detected without being triggered.

The drawback is that the one-shot tradeoff is still not especially favorable. For small ε, the probability of detecting the bomb without an explosion is of the same order as the probability of causing an explosion.

2.2.2 Recycling the photon

The original one-shot experiment leaves open whether this tradeoff can be improved. In 1995, Kwiat, Weinfurter, Herzog, Zeilinger, and Kasevich showed that it can: by replacing one substantial interaction with many weak interrogations, the probability of an interaction-free detection can be made arbitrarily large [22]. Their construction is closely related to the quantum Zeno effect, in which repeated measurements inhibit coherent quantum evolution [23].

In our model, the idea is simple. Start with sending a photon through a beam-splitter with a small angle R_θ. After it splits and goes through the safe and dangerous paths, recombine the paths, but pass it again through the original R_θ beam-splitter, and repeat till it has passed through T times. Then, simply measure whether it will strike the "accept" or "reject" detectors.

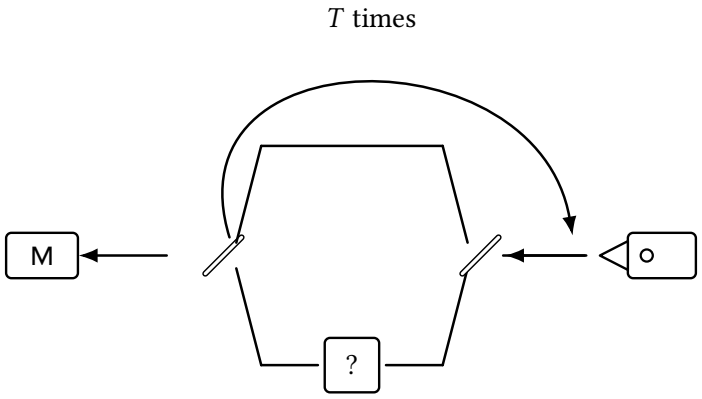


Figure 2.5: The recycled-photon bomb test drawn as a cartoon.

Theorem 2.4 (Recycled-photon bomb test)

If the unknown object is a dud, then the final measurement outputs 1 *with probability* $\sin^2(T\theta)$. *If it is a live bomb, then conditioned on no explosion the final measurement outputs* 0 *with probability* 1, *while*

$$\Pr[\text{explode}] = 1 - \cos^{2T}\theta \le T\sin^2\theta. \tag{2.16}$$

Consequently, for every $\varepsilon > 0$, *one can distinguish the two cases with explosion probability at most* ε *using* $T = O(1/\varepsilon)$ *passes through the box.*

Proof. Suppose first that the object is a dud. The box never does anything, so the small rotations accumulate coherently:

$$R_\theta^T|0\rangle = R_{T\theta}|0\rangle = R_{\pi/2}|0\rangle = |1\rangle. \tag{2.17}$$

The final measurement therefore outputs 1 with probability $\sin^2 T\theta$. If we pick $T = \pi/2\theta$, then this is 1. Now suppose the object is a live bomb. Conditioned on surviving the previous rounds, the photon begins each round in $|0\rangle$. After the small rotation its state is

$$\cos\theta|0\rangle + \sin\theta|1\rangle. \tag{2.18}$$

The bomb explodes in that round with probability $\sin^2\theta$. If it does not explode, the measurement has outcome 0 and the state is returned to $|0\rangle$. The probability of surviving all T rounds is therefore

$$\left(\cos^2\theta\right)^2 \ge 1 - T\sin^2\theta, \tag{2.19}$$

using a union bound. Using the small angle bound, the probability of explosion is at most

$$\Pr[\text{explode} \mid \text{bomb}] \le T\sin^2\theta \le T\theta^2 = \frac{\pi\theta}{2}. \tag{2.20}$$

Thus choosing,

$$\theta \leq \frac{2\varepsilon}{\pi} \tag{2.21}$$

bounds the explosion probability at ε, while

$$T = \frac{\pi}{2\theta} = O(1/\varepsilon). \tag{2.22}$$

□

The two cases of the dud and the live bomb differ because only the dud preserves coherent evolution. With a dud, the small rotations accumulate to a total rotation of $\pi/2$. With a live bomb, every successful passage through the box measures the path and returns the surviving photon to $|0\rangle$, preventing those rotations from accumulating.

This repeated-measurement mechanism is the connection to the quantum Zeno effect [23]. In the bomb-testing setting, it allows the probability of an explosion to approach zero while the two cases remain perfectly distinguishable [22].

2.3 The Deutsch–Jozsa game

The next example is one of the earliest demonstrations that a quantum computer can extract information using fewer queries than a classical computer by leveraging superposition and interference patterns. The one-bit version below was introduced by Deutsch in 1985 [4]; Deutsch and Jozsa later generalized the idea to functions on many input bits [24]. We will study the smallest version because it already contains several ideas that will recur throughout this course.

Imagine a box with two buttons and two lights, as in Figure 2.6. The box is described by a function

$$f : \{0,1\} \to \{0,1\}. \tag{2.23}$$

Pressing button $x \in \{0,1\}$ turns on exactly one light, namely light $f(x)$. It is possible that both buttons turn on the same light, and it is possible that they turn on different lights.

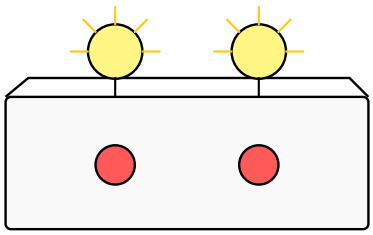


Figure 2.6: A classical query box with two buttons and two lights. Pressing button $x \in \{0,1\}$ turns on exactly one light, namely light $f(x) \in \{0,1\}$.

We are not trying to learn the two values $f(0)$ and $f(1)$ individually. Instead, we only want to know

whether the two buttons have the same action or different actions. Equivalently, our goal is to determine

$$f(0) \oplus f(1). \tag{2.24}$$

Classically, two button presses are necessary if we want the answer with certainty. After pressing only button 0, for example, we know $f(0)$ but have learned nothing about whether $f(1)$ is the same or different. Pressing both buttons determines the answer—and in fact tells us the two values $f(0)$ and $f(1)$ individually.

Quantumly, we can determine whether the two buttons have the same or different actions using only one query. Interestingly, the algorithm does *not* learn which lights the individual buttons turn on. It extracts exactly the one bit: $f(0) \oplus f(1)$—this highlights a powerful property: some "orthogonal" information can be easy to access with a quantum computer that is not easy to access with a classical computer.

2.3.1 Querying the box quantumly

To formulate a quantum version of the box, we first need to decide what it means to query a classical function with a quantum state. The naive transformation

$$|x\rangle \longmapsto |f(x)\rangle \tag{2.25}$$

does not work: if $f(0) = f(1)$, then two orthogonal input states would be mapped to the same output state, so the transformation would not be unitary. Instead, we introduce a second qubit and define a reversible version of the classical query.

Definition 2.5 (Quantum query oracle)

For a function $f : \{0,1\} \to \{0,1\}$, define the two-qubit unitary U_f by

$$U_f |x\rangle |y\rangle = |x\rangle |y \oplus f(x)\rangle \tag{2.26}$$

for every $x, y \in \{0,1\}$, and extend the transformation to all two-qubit states by linearity.

The map is unitary because it simply permutes the computational-basis states. Moreover, if the second qubit begins in $|0\rangle$, then

$$U_f |x\rangle |0\rangle = |x\rangle |f(x)\rangle . \tag{2.27}$$

So, on classical inputs it contains exactly the same information as pressing button x on the box in Figure 2.6. One query to U_f can therefore be simulated classically using one query to f. The difference is that U_f can also be applied to a superposition. For example, linearity gives

$$U_f (\alpha |0\rangle + \beta |1\rangle) |0\rangle = \alpha |0\rangle |f(0)\rangle + \beta |1\rangle |f(1)\rangle . \tag{2.28}$$

This is the natural quantum generalization of pressing a button: the first register can now be in a superposition of the two possible inputs. The exact one-query algorithm we use below is the standard refinement of Deutsch's original algorithm based on phase kickback [25]. Its circuit is shown in Figure 2.7.

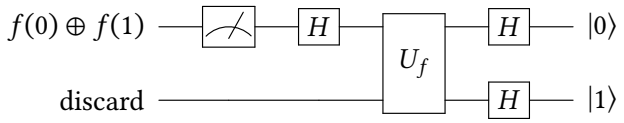


Figure 2.7: The one-query algorithm for Deutsch's problem. The two Hadamards on the right prepare $|+\rangle|-\rangle$, the oracle is queried once, and the final Hadamard converts the relative phase into the bit $f(0) \oplus f(1)$.

Theorem 2.6 (One-query Deutsch algorithm)

The circuit in Figure 2.7 outputs $f(0) \oplus f(1)$ with certainty using one query to U_f.

Proof. The input is $|0\rangle|1\rangle$. After applying $H \otimes H$, the state is

$$|+\rangle|-\rangle = \frac{(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle)}{2} \tag{2.29a}$$

$$= \frac{1}{2} \sum_{x \in \{0,1\}} |x\rangle (|0\rangle - |1\rangle). \tag{2.29b}$$

We now apply the oracle. For a fixed x, observe that

$$|f(x)\rangle - |1 \oplus f(x)\rangle = (-1)^{f(x)}(|0\rangle - |1\rangle). \tag{2.30}$$

Therefore,

$$U_f |+\rangle|-\rangle = U_f \left(\frac{1}{2} \sum_{x \in \{0,1\}} |x\rangle (|0\rangle - |1\rangle) \right) \tag{2.31a}$$

$$= \frac{1}{2} \sum_{x \in \{0,1\}} |x\rangle (|f(x)\rangle - |1 \oplus f(x)\rangle) \tag{2.31b}$$

$$= \frac{1}{2} \sum_{x \in \{0,1\}} (-1)^{f(x)} |x\rangle (|0\rangle - |1\rangle) \tag{2.31c}$$

$$= \left(\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}} \right) \otimes |-\rangle \tag{2.31d}$$

$$= (-1)^{f(0)} \left(\frac{|0\rangle + (-1)^{f(0)\oplus f(1)}|1\rangle}{\sqrt{2}} \right) \otimes |-\rangle . \tag{2.31e}$$

The factor $(-1)^{f(0)}$ is a global phase and can be ignored. If $f(0) \oplus f(1) = 0$, the first qubit is $|+\rangle$; if $f(0) \oplus f(1) = 1$, the first qubit is $|-\rangle$. The final Hadamard therefore produces

$$(-1)^{f(0)} |f(0) \oplus f(1)\rangle \otimes |-\rangle . \tag{2.32}$$

Measuring the first qubit returns $f(0) \oplus f(1)$ with probability 1. □

Intuition 2.7 (Phase kickback)

The second register is prepared in $|-\rangle$, which satisfies

$$X|-\rangle = -|-\rangle. \tag{2.33}$$

Consequently,

$$U_f|x\rangle|-\rangle = (-1)^{f(x)}|x\rangle|-\rangle. \tag{2.34}$$

Instead of leaving a copy of $f(x)$ in the second register, the query writes $f(x)$ into the *phase* of the $|x\rangle$ branch. This phenomenon is called *phase kickback*. The final Hadamard then turns the relative phase between the $|0\rangle$ and $|1\rangle$ branches into an ordinary computational-basis bit.

The significance of this example is not the modest improvement from two classical queries to one quantum query. Rather, it isolates a pattern that will appear repeatedly in quantum algorithms. We query a function on a superposition of inputs, arrange for the function values to appear as relative phases, and then use interference to recover a global property of those values.

2.4 Entanglement and product states

We now move from few-qubit systems to general quantum systems which are the tensor product of two smaller systems, referred to as separated registers. If Alice has an n_1-qubit system and Bob has an n_2-qubit system, then the joint state space is $(\mathbb{C}^2)^{\otimes n_1} \otimes (\mathbb{C}^2)^{\otimes n_2}$.

Definition 2.8 (Product and entangled states)

A pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a product state across the cut $A|B$ if there exist unit vectors $|a\rangle \in \mathcal{H}_A$ and $|b\rangle \in \mathcal{H}_B$ such that $|\psi\rangle = |a\rangle \otimes |b\rangle$. Otherwise, $|\psi\rangle$ is *entangled* across the cut and is called an entangled state.

The central two-qubit example of an entangled state is the EPR state.

Lemma 2.9 (The EPR state is entangled)

The state[a]

$$|\mathrm{EPR}\rangle \stackrel{\text{def}}{=} \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{2.35}$$

cannot be written as a product state between the first and second qubits.

[a]The acronym EPR stands for Einstein, Podelsky, and Rosen who were among the earliest to study the spooky entanglement highlighted by this particular state [26]

Proof. Suppose toward contradiction that $|\mathrm{EPR}\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle)$. Comparing coefficients

gives $\alpha\gamma = 1/\sqrt{2}$ and $\beta\delta = 1/\sqrt{2}$, so all four of $\alpha, \beta, \gamma, \delta$ are nonzero. But the coefficients of $|01\rangle$ and $|10\rangle$ are $\alpha\delta$ and $\beta\gamma$, and in $|\mathrm{EPR}\rangle$ both must equal 0. This is impossible. □

Note 2.10 (Counting intuition)

Entanglement is not a rare corner case. A product state across an $n|n$ cut is described by one n-qubit state on the left and one n-qubit state on the right. A general $2n$-qubit state has amplitudes over a space of dimension 2^{2n}. Thus the number of degrees of freedom in a general state is exponentially larger than the number of degrees of freedom in a pair of n-qubit states. One can make this intuition discrete by considering phase states and large approximately orthogonal subsets: the $2n$-qubit space contains far more well-separated states than can be accounted for by tensor products of two n-qubit states.

What is special about an entangled state? The EPR state gives a first answer: it produces correlations that depend on which basis is measured. In the standard basis, measuring the first qubit gives 0 or 1 with probability $1/2$, and the second qubit becomes the same standard-basis state. The same kind of statement holds in the $\{|+\rangle, |-\rangle\}$ basis as well.

Lemma 2.11 (EPR correlations in the Hadamard basis)

If the first qubit of $|\mathrm{EPR}\rangle$ is measured in the $\{|+\rangle, |-\rangle\}$ basis, then the outcome is uniform. Conditioned on outcome $+$, the second qubit is $|+\rangle$; conditioned on outcome $-$, the second qubit is $|-\rangle$.

Proof. We have that

$$\frac{|{+}{+}\rangle + |{-}{-}\rangle}{\sqrt{2}} = \frac{\frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \frac{|0\rangle-|1\rangle}{\sqrt{2}}\frac{|0\rangle-|1\rangle}{\sqrt{2}}}{\sqrt{2}} \tag{2.36a}$$

$$= \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle + |00\rangle - |01\rangle - |10\rangle + |11\rangle}{2\sqrt{2}} \tag{2.36b}$$

$$= \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{2.36c}$$

$$= |\mathrm{EPR}\rangle \tag{2.36d}$$

The Born rule in the $\{|+\rangle, |-\rangle\}$ basis then gives the two outcomes with probability $1/2$ each, with the corresponding post-measurement state on the second qubit. □

Theorem 2.12 (Remote preparation from an EPR pair)

Let $\{|b_0\rangle, |b_1\rangle\}$ be an orthonormal basis of one qubit. If Alice and Bob share $|\mathrm{EPR}\rangle$ and Alice measures her qubit in this basis, then Alice's outcome i is uniform in $\{0, 1\}$ and Bob's post-measurement state is $|\overline{b_i}\rangle$, where $|\overline{b_i}\rangle$ denotes the entrywise complex conjugate of $|b_i\rangle$ in the computational basis.

Proof. Write $|b_i\rangle = \alpha_i |0\rangle + \beta_i |1\rangle$. The amplitude for Alice to obtain outcome i is obtained by applying $\langle b_i| \otimes I$

to $|\mathrm{EPR}\rangle$, which gives

$$\left(\frac{\overline{\alpha_i}\langle 0|\otimes I+\overline{\beta_i}\langle 1|\otimes I}{\sqrt{2}}\right)\left(\frac{|00\rangle+|11\rangle}{\sqrt{2}}\right)=\frac{\overline{\alpha_i}|0\rangle+\overline{\beta_i}|1\rangle}{\sqrt{2}}=\frac{|\overline{b_i}\rangle}{\sqrt{2}} \tag{2.37}$$

This vector has squared norm $1/2$, so the outcome probability is $1/2$ and the normalized state on Bob's side is $|\overline{b_i}\rangle$. □

It is important that this does not permit faster-than-light communication. Alice can choose a measurement basis, but she cannot choose the random outcome. Without learning Alice's outcome through an ordinary classical message, Bob sees only the same local mixed state. Density matrices will give the clean language for this statement.

2.5 Superdense coding

Superdense coding, introduced by Bennett and Wiesner in 1992 [27], is one of the first examples showing that entanglement can be used as a computational resource. Suppose Alice and Bob are together before Bob leaves on a long trip. While they are together, they are free to prepare and share a quantum state. Later, Alice wants to send Bob a two-bit message. Ordinarily, communicating two classical bits requires sending two bits. Can Alice and Bob use their earlier quantum preparation to reduce the amount of communication?

Remarkably, the answer is yes. If Alice and Bob share an EPR pair in advance, then Alice can later communicate two classical bits by sending Bob only one qubit. In this sense, superdense coding gives us two classical bits for the price of communicating one qubit and one EPR pair. Suppose Alice and Bob initially share

$$|\Phi_{00}\rangle \stackrel{\text{def}}{=} |\mathrm{EPR}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}, \tag{2.38}$$

where Alice holds the first qubit and Bob holds the second. Alice's key observation is that by acting only on her qubit, she can transform the shared state into one of four different and orthogonal states. For $x, z \in \{0,1\}$, define

$$|\Phi_{xz}\rangle \stackrel{\text{def}}{=} (X^x Z^z \otimes I)\,|\Phi_{00}\rangle\,. \tag{2.39}$$

Explicitly,

$$\begin{aligned} |\Phi_{00}\rangle &= \frac{|00\rangle+|11\rangle}{\sqrt{2}}, \quad |\Phi_{10}\rangle = \frac{|10\rangle+|01\rangle}{\sqrt{2}}, \\ |\Phi_{01}\rangle &= \frac{|00\rangle-|11\rangle}{\sqrt{2}}, \quad |\Phi_{11}\rangle = \frac{|10\rangle-|01\rangle}{\sqrt{2}}. \end{aligned} \tag{2.40}$$

These four states are known as the *Bell basis*. They are the four possibilities Alice will use to encode her two-bit message.

Lemma 2.13 (Bell basis)

The four states $\{|\Phi_{xz}\rangle : x, z \in \{0,1\}\}$ *form an orthonormal basis of the two-qubit space.*

Proof. Each state has norm 1. From the explicit expressions above, states supported on $\{|00\rangle, |11\rangle\}$ are orthogonal to those supported on $\{|01\rangle, |10\rangle\}$, while

$$\left\langle \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \frac{|00\rangle - |11\rangle}{\sqrt{2}} \right\rangle = 0, \tag{2.41}$$

and similarly for the remaining pair. Thus the four states are pairwise orthogonal. Since the two-qubit space has dimension 4, they form an orthonormal basis. □

This already suggests the protocol. To send the message (x, z), Alice applies $X^x Z^z$ to her half of the EPR pair. She has now changed the *joint* state of Alice and Bob's two qubits to $|\Phi_{xz}\rangle$. She then sends her one qubit to Bob. Bob now possesses both qubits, and because the four possible states are orthogonal, he can distinguish them perfectly.

The protocol is illustrated in Figure 2.8. It remains to explain how Bob performs the Bell-basis measure-

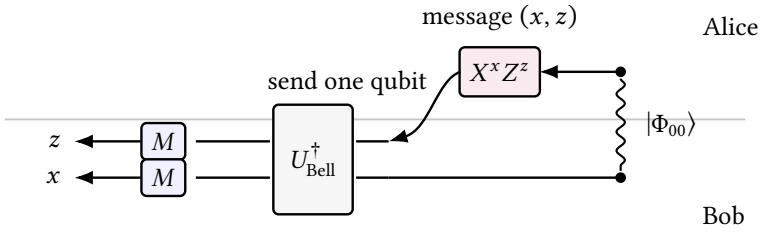


Figure 2.8: Superdense coding, drawn right to left. Alice and Bob begin with a shared EPR pair. Alice encodes the two-bit message (x, z) by applying $X^x Z^z$ to her qubit and sends that qubit to Bob. Bob now holds both qubits and performs a Bell-basis measurement to recover (x, z).

ment. Since the Bell states are an orthonormal basis, some unitary must map them to the computational basis. In this case, there is a particularly simple one: apply a CNOT from Alice's qubit to Bob's qubit, followed by a Hadamard on Alice's qubit. The corresponding circuit is

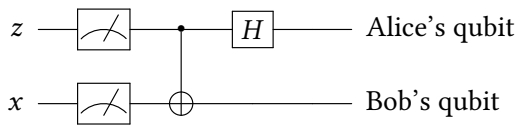


Figure 2.9: A Bell-basis measurement. The circuit maps $|\Phi_{xz}\rangle$ to the computational-basis state $|z\rangle|x\rangle$, after which standard-basis measurements recover the two bits.

Lemma 2.14 (Decoding the Bell basis)

The circuit in Figure 2.9 maps

$$|\Phi_{xz}\rangle \longmapsto |z\rangle\,|x\rangle\,. \tag{2.42}$$

Proof. First consider $|\Phi_{00}\rangle$. Applying the CNOT gives

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \longmapsto \frac{|00\rangle + |10\rangle}{\sqrt{2}} = |+\rangle\,|0\rangle\,, \tag{2.43}$$

and the Hadamard on the first qubit gives $|00\rangle$. Similarly,

$$|\Phi_{10}\rangle \longmapsto |01\rangle\,, \tag{2.44a}$$
$$|\Phi_{01}\rangle \longmapsto |10\rangle\,, \tag{2.44b}$$
$$|\Phi_{11}\rangle \longmapsto |11\rangle\,. \tag{2.44c}$$

Thus $|\Phi_{xz}\rangle$ is mapped to $|z\rangle\,|x\rangle$. □

We can now state the protocol formally.

Theorem 2.15 (Superdense coding)

Suppose Alice and Bob share an EPR pair. Then Alice can transmit two classical bits $(x, z) \in \{0, 1\}^2$ to Bob by sending a single qubit.

Proof. Alice and Bob begin with the state $|\Phi_{00}\rangle$. To send (x, z), Alice applies $X^x Z^z$ to her half of the pair, producing

$$(X^x Z^z \otimes I)\,|\Phi_{00}\rangle = |\Phi_{xz}\rangle\,. \tag{2.45}$$

Alice then sends her qubit to Bob. Bob now holds both qubits and applies the Bell-basis measurement from Figure 2.9. By Lemma 2.14, the state becomes

$$|z\rangle\,|x\rangle\,. \tag{2.46}$$

Measuring both qubits therefore recovers (x, z) with certainty. □

The important point is that Alice is not somehow placing two classical bits inside an isolated qubit. The two parties already possess a shared entangled state before the communication begins. Alice's local operation moves this *global* two-qubit state among four mutually orthogonal possibilities, and after she sends her half of the pair to Bob, he has access to the entire two-qubit state and can determine which possibility she chose.

This is a first example of treating entanglement as a resource rather than merely as an unusual correlation. The EPR pair was prepared before Alice knew which message she wanted to send, but once it is available, it allows two classical bits to be communicated using only one qubit of subsequent communication.

Bibliography

[1] Alonzo Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2):345–363, 1936.

[2] A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(1):230–265, 1936.

[3] Richard P Feynman. Simulating physics with computers. *International journal of theoretical physics*, 21(6/7):467–488, 1982.

[4] David Deutsch. Quantum theory, the church–turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.

[5] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, Bangalore, India, 1984. Reprinted in Theoretical Computer Science 560, 7–11 (2014).

[6] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997.

[7] Daniel R Simon. On the power of quantum computation. *SIAM journal on computing*, 26(5):1474–1483, 1997.

[8] Ronald L. Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[9] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[10] William K. Wootters and Wojciech H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.

[11] John S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195–200, 1964.

[12] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[13] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, Oct 1995.

[14] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate, 1999.

[15] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626:58–65, 2024.

[16] Rajeev Acharya, Laleh Aghababaie-Beni, Igor Aleiner, Trond I. Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, et al. Quantum error correction below the surface code threshold. *Nature*, 638:920–926, 2025.

[17] P. A. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(3):416–418, 1939.

[18] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738–2747, 1995.

[19] Max Born. Zur quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 37:863–867, 1926.

[20] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Physical Review A*, 52(5):3457–3467, 1995.

[21] Avshalom C. Elitzur and Lev Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, 1993.

[22] Paul G. Kwiat, Harald Weinfurter, Thomas Herzog, Anton Zeilinger, and Mark A. Kasevich. Interaction-free measurement. *Physical Review Letters*, 74(24):4763–4766, 1995.

[23] Baidyanath Misra and E. C. George Sudarshan. The zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18(4):756–763, 1977.

[24] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[25] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998.

[26] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can Quantum-Mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, May 1935.

[27] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Physical Review Letters*, 69(20):2881–2884, 1992.