Lecture 1: Quantum information processing basics

Mark M. Wilde*

The simplest quantum system is the physical quantum bit or *qubit*. The qubit is a two-level quantum system—example qubit systems are the spin of an electron, the polarization of a photon, or a two-level atom with a ground state and an excited state. We do not worry too much about physical implementations in this lecture, but instead focus on the mathematical postulates of the quantum theory and operations that we can perform on qubits.

Noise can affect quantum systems, and we must understand methods of modeling noise in the quantum theory because our ultimate aim is to construct schemes for protecting quantum systems against the detrimental effects of noise. In some sense, there are different types of noise that occur in nature. The first, and perhaps more easily comprehensible type of noise, is that which is due to our lack of information about a given scenario. We observe this type of noise in a casino, with every shuffle of cards or toss of dice. These events are random, and the random variables of probability theory model them because the outcomes are unpredictable. This noise is the same as that in all classical information processing systems. We can engineer physical systems to improve their robustness to noise.

On the other hand, the quantum theory features a fundamentally different type of noise. Quantum noise is inherent in nature and is not due to our lack of information, but is due rather to nature itself. An example of this type of noise is the "Heisenberg noise" that results from the uncertainty principle. If we know the momentum of a given particle from performing a precise measurement of it, then we know absolutely nothing about its position—a measurement of its position gives a random result. Similarly, if we know the rectilinear polarization of a photon by precisely measuring it, then a future measurement of its diagonal polarization will give a random result. It is important to keep the distinction clear between these two types of noise.

We explore the postulates of the quantum theory in this lecture, by paying particular attention to qubits. These postulates apply to a closed quantum system that is isolated from everything else in the universe. We call this variant of the quantum theory the "noiseless quantum theory" because closed quantum systems do not interact with their surroundings and are thus not subject to corruption and information loss. Interaction with surrounding systems can lead to loss of information in the sense of the classical noise that we described above. Closed quantum systems do undergo a certain type of quantum noise, such as that from the uncertainty principle and the act of measurement, because they are subject to the postulates of the quantum theory. The name "Noiseless Quantum Theory" thus indicates the closed, ideal nature of the quantum systems discussed in this lecture.

This lecture introduces the four postulates of the quantum theory. The mathematical tools of

*Mark M. Wilde is with the Department of Physics and Astronomy and the Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803. Much of the material is from the book preprint "From Classical to Quantum Shannon Theory" available as `arXiv:1106.1445`.

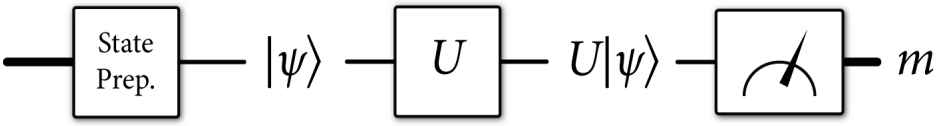

Figure 1: All of the steps in a typical noiseless quantum information processing protocol. A classical control (depicted by the thick black line on the left) initializes the state of a quantum system. The quantum system then evolves according to some unitary operation (described in Section 3). The final step is a measurement that reads out some classical data m from the quantum system.

the quantum theory rely on the fundamentals of linear algebra—vectors and matrices of complex numbers. It may seem strange at first that we need to incorporate the machinery of linear algebra in order to describe a physical system in the quantum theory, but it turns out that this description uses the simplest set of mathematical tools to predict the phenomena that a quantum system exhibits. The hallmark of the quantum theory is that certain operations do not commute with one another, and matrices are the simplest mathematical objects that capture this idea of noncommutativity.

1 Overview

We first briefly overview how information is processed with quantum systems. This usually consists of three steps: state preparation, quantum operations, and measurement. State preparation is where we initialize a quantum system to some beginning state, depending on what operation we would like a quantum system to execute. There could be some classical control device that initializes the state of the quantum system. Observe that the input system for this step is a classical system, and the output system is quantum. After initializing the state of the quantum system, we perform some quantum operations that evolve its state. This stage is where we can take advantage of quantum effects for enhanced information processing abilities. Both the input and output systems of this step are quantum. Finally, we need some way of reading out the result of the computation, and we can do so with a measurement. The input system for this step is quantum, and the output is classical. Figure 1 depicts all of these steps. In a quantum communication protocol, spatially separated parties may execute different parts of these steps, and we are interested in keeping track of the nonlocal resources needed to implement a communication protocol. Section 2 describes quantum states (and thus state preparation), Section 3 describes the noiseless evolution of quantum states, and Section 4 describes "read out" or measurement. For now, we assume that we can perform all of these steps perfectly and later on we discuss how to incorporate the effects of noise.

2 Quantum Bits

The simplest quantum system is a two-state system: a physical qubit. Let $|0\rangle$ denote one possible state of the system. The left vertical bar and the right angle bracket indicate that we are using the Dirac notation to represent this state. The Dirac notation has some advantages for performing calculations in the quantum theory, and we highlight some of these advantages as we progress through our development. Let $|1\rangle$ denote another possible state of the qubit. We can encode a

classical bit or *cbit* into a qubit with the following mapping:

$$0 \rightarrow |0\rangle, \qquad 1 \rightarrow |1\rangle. \tag{1}$$

So far, nothing in our description above distinguishes a classical bit from a qubit, except for the funny vertical bar and angle bracket that we place around the bit values. The quantum theory predicts that the above states are not the only possible states of a qubit. Arbitrary *superpositions* (linear combinations) of the above two states are possible as well because the quantum theory is a linear theory. Suffice it to say that the linearity of the quantum theory results from the linearity of Schrödinger's equation that governs the evolution of quantum systems. A general noiseless qubit can be in the following state:

$$\alpha|0\rangle + \beta|1\rangle, \tag{2}$$

where the coefficients α and β are arbitrary complex numbers with unit norm:

$$|\alpha|^2 + |\beta|^2 = 1. \tag{3}$$

The coefficients α and β are *probability amplitudes*—they are not probabilities themselves but allow us to calculate probabilities. The unit-norm constraint leads to the *Born rule* (the probabilistic interpretation) of the quantum theory, and we speak more on this constraint and probability amplitudes when we introduce the measurement postulate.

The possibility of superposition states indicates that we cannot represent the states $|0\rangle$ and $|1\rangle$ with the Boolean algebra of the respective classical bits 0 and 1 because Boolean algebra does not allow for superposition states. We instead require the mathematics of *linear algebra* to describe these states. It is beneficial at first to define a vector representation of the states $|0\rangle$ and $|1\rangle$:

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{4}$$

$$|1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{5}$$

The $|0\rangle$ and $|1\rangle$ states are called "kets" in the language of the Dirac notation, and it is best at first to think of them merely as column vectors. The superposition state in (2) then has a representation as the following two-dimensional vector:

$$|\psi\rangle \equiv \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \tag{6}$$

The representation of quantum states with vectors is helpful in understanding some of the mathematics that underpins the theory, but it turns out to be much more useful for our purposes to work directly with the Dirac notation. We give the vector representation for now, but later on, we will only employ the Dirac notation.

The *Bloch sphere*, depicted in Figure 2, gives a valuable way to visualize a qubit. Consider any two qubits that are equivalent up to a differing global phase. For example, these two qubits could be

$$|\psi_0\rangle \equiv |\psi\rangle, \qquad |\psi_1\rangle \equiv e^{i\chi}|\psi\rangle, \tag{7}$$

where $0 \leq \chi < 2\pi$. There is a sense in which these two qubits are physically equivalent because they give the same physical results when we measure them (more on this point when we introduce

the measurement postulate in Section 4). Suppose that the probability amplitudes α and β have the following respective representations as complex numbers:

$$\alpha = r_0 e^{i\varphi_0}, \tag{8}$$

$$\beta = r_1 e^{i\varphi_1}. \tag{9}$$

We can factor out the phase $e^{i\varphi_0}$ from both coefficients α and β, and we still have a state that is physically equivalent to the state in (2):

$$|\psi\rangle \equiv r_0|0\rangle + r_1 e^{i(\varphi_1 - \varphi_0)}|1\rangle, \tag{10}$$

where we redefine $|\psi\rangle$ to represent the state because of the equivalence mentioned in (7). Let $\varphi \equiv \varphi_1 - \varphi_0$, where $0 \leq \varphi < 2\pi$. Recall that the unit-norm constraint requires $|r_0|^2 + |r_1|^2 = 1$. We can thus parametrize the values of r_0 and r_1 in terms of one parameter θ so that

$$r_0 = \cos(\theta/2), \tag{11}$$

$$r_1 = \sin(\theta/2). \tag{12}$$

The parameter θ varies between 0 and π. This range of θ and the factor of two give a unique representation of the qubit. One may think to have θ vary between 0 and 2π and omit the factor of two, but this parametrization would not uniquely characterize the qubit in terms of the parameters θ and φ. The parametrization in terms of θ and φ gives the Bloch sphere representation of the qubit in (2):

$$|\psi\rangle \equiv \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i\varphi}|1\rangle. \tag{13}$$

In linear algebra, column vectors are not the only type of vectors—row vectors are useful as well. Is there an equivalent of a row vector in Dirac notation? The Dirac notation provides an entity called a "bra," that has a representation as a row vector. The bras corresponding to the kets $|0\rangle$ and $|1\rangle$ are as follows:

$$\langle 0| \equiv \begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{14}$$

$$\langle 1| \equiv \begin{bmatrix} 0 & 1 \end{bmatrix}, \tag{15}$$

and are the matrix conjugate transpose of the kets $|0\rangle$ and $|1\rangle$:

$$\langle 0| = (|0\rangle)^\dagger, \tag{16}$$

$$\langle 1| = (|1\rangle)^\dagger. \tag{17}$$

We require the conjugate transpose operation (as opposed to just the transpose) because the mathematical representation of a general quantum state can have complex entries.

The bras do not represent quantum states, but are helpful in calculating probability amplitudes. For our example qubit in (2), suppose that we would like to determine the probability amplitude that the state is $|0\rangle$. We can combine the state in (2) with the bra $\langle 0|$ as follows:

$$\langle 0||\psi\rangle = \langle 0|(\alpha|0\rangle + \beta|1\rangle) \tag{18}$$

$$= \alpha\langle 0||0\rangle + \beta\langle 0||1\rangle \tag{19}$$

$$= \alpha \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{20}$$

$$= \alpha \cdot 1 + \beta \cdot 0 \tag{21}$$

$$= \alpha. \tag{22}$$

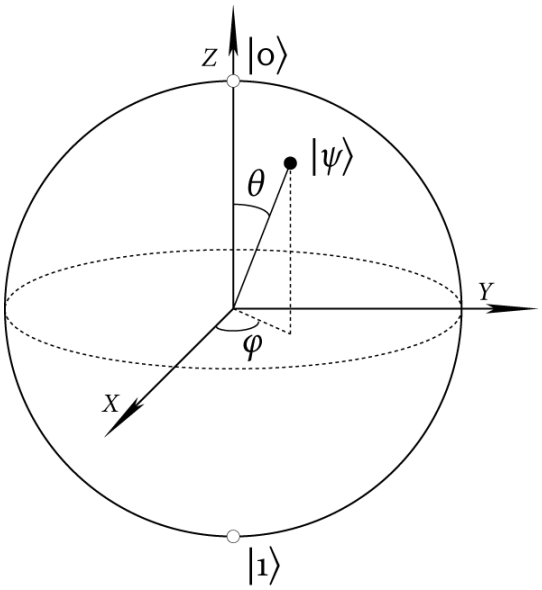

Figure 2: The Bloch sphere representation of a qubit. Any qubit $|\psi\rangle$ admits a representation in terms of two angles θ and φ where $0 \leq \theta \leq \pi$ and $0 \leq \varphi < 2\pi$. The state of any qubit in terms of these angles is $|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$.

The above calculation may seem as if it is merely an exercise in linear algebra, with a "glorified" Dirac notation, but it is a standard calculation in the quantum theory. A quantity like $\langle 0||\psi\rangle$ occurs so often in the quantum theory that we abbreviate it as

$$\langle 0|\psi\rangle \equiv \langle 0||\psi\rangle, \tag{23}$$

and the above notation is known as a "braket."[1] The physical interpretation of the quantity $\langle 0|\psi\rangle$ is that it is the probability amplitude for being in the state $|0\rangle$, and likewise, the quantity $\langle 1|\psi\rangle$ is the probability amplitude for being in the state $|1\rangle$. We can also determine that the amplitude $\langle 1|0\rangle$ (for the state $|0\rangle$ to be in the state $|1\rangle$) and the amplitude $\langle 0|1\rangle$ are both equal to zero. These two states are *orthogonal states* because they have no overlap. The amplitudes $\langle 0|0\rangle$ and $\langle 1|1\rangle$ are both equal to one by following a similar calculation.

Our next task may seem like a frivolous exercise, but we would like to determine the amplitude for any state $|\psi\rangle$ to be in the state $|\psi\rangle$, i.e., to be itself. Following the above method, this amplitude is $\langle\psi|\psi\rangle$ and we calculate it as

$$\begin{aligned}
\langle\psi|\psi\rangle &= (\langle 0|\alpha^* + \langle 1|\beta^*)(\alpha|0\rangle + \beta|1\rangle) && (24)\\
&= \alpha^*\alpha\langle 0|0\rangle + \beta^*\alpha\langle 1|0\rangle + \alpha^*\beta\langle 0|1\rangle + \beta^*\beta\langle 1|1\rangle && (25)\\
&= |\alpha|^2 + |\beta|^2 && (26)\\
&= 1, && (27)
\end{aligned}$$

[1]It is for this (somewhat silly) reason that Dirac decided to use the names "bra" and "ket," because putting them together gives a "braket." The names in the notation may be silly, but the notation itself has persisted over time because this way of representing quantum states turns out to be useful. We will avoid the use of the terms "bra" and "ket" as much as we can, only resorting to these terms if necessary.

where we have used the orthogonality relations of $\langle 0|0\rangle$, $\langle 1|0\rangle$, $\langle 0|1\rangle$, and $\langle 1|1\rangle$, and the unit-norm constraint. We come back to the unit-norm constraint in our discussion of quantum measurement, but for now, we have shown that any quantum state has a unit amplitude for being itself.

The states $|0\rangle$ and $|1\rangle$ are a particular basis for a qubit that we call the *computational basis*. The computational basis is the standard basis that we employ in quantum computation and communication, but other bases are important as well. Consider that the following two vectors form an orthonormal basis:

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{28}$$

The above alternate basis is so important in quantum information theory that we define a Dirac notation shorthand for it, and we can also define the basis in terms of the computational basis:

$$|+\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \tag{29}$$

$$|-\rangle \equiv \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{30}$$

The common names for this alternate basis are the "+/−" basis, the Hadamard basis, or the diagonal basis. It is preferable for us to use the Dirac notation, but we are using the vector representation as an aid for now.

Exercise 1 *Determine the Bloch sphere angles θ and φ for the states $|+\rangle$ and $|-\rangle$.*

What is the amplitude that the state in (2) is in the state $|+\rangle$? What is the amplitude that it is in the state $|-\rangle$? These are questions to which the quantum theory provides simple answers. We employ the bra $\langle +|$ and calculate the amplitude $\langle +|\psi\rangle$ as

$$\langle +|\psi\rangle = \langle +|(\alpha|0\rangle + \beta|1\rangle) \tag{31}$$

$$= \alpha\langle +|0\rangle + \beta\langle +|1\rangle \tag{32}$$

$$= \frac{\alpha + \beta}{\sqrt{2}}. \tag{33}$$

The result follows by employing the definition in (29) and doing similar linear algebraic calculations as the example in (22). We can also calculate the amplitude $\langle -|\psi\rangle$ as

$$\langle -|\psi\rangle = \frac{\alpha - \beta}{\sqrt{2}}. \tag{34}$$

The above calculation follows from similar manipulations.

The +/− basis is a *complete* orthonormal basis, meaning that we can represent any qubit state in terms of the two basis states $|+\rangle$ and $|-\rangle$. Indeed, the above probability amplitude calculations suggest that we can represent the qubit in (2) as the following superposition state:

$$|\psi\rangle = \left(\frac{\alpha + \beta}{\sqrt{2}}\right)|+\rangle + \left(\frac{\alpha - \beta}{\sqrt{2}}\right)|-\rangle. \tag{35}$$

The above representation is an alternate one if we would like to "see" the qubit state represented in the +/− basis. We can substitute the equivalences in (33) and (34) to represent the state $|\psi\rangle$ as

$$|\psi\rangle = \langle +|\psi\rangle|+\rangle + \langle -|\psi\rangle|-\rangle. \tag{36}$$

The amplitudes $\langle +|\psi\rangle$ and $\langle -|\psi\rangle$ are both scalar quantities so that the above quantity is equivalent to the following one:

$$|\psi\rangle = |+\rangle\langle +|\psi\rangle + |-\rangle\langle -|\psi\rangle. \tag{37}$$

The order of the multiplication in the terms $|+\rangle\langle +|\psi\rangle$ and $|-\rangle\langle -|\psi\rangle$ does not matter, i.e., the following equivalence holds

$$|+\rangle(\langle +|\psi\rangle) = (|+\rangle\langle +|)|\psi\rangle, \tag{38}$$

and the same for $|-\rangle\langle -|\psi\rangle$. The quantity on the left is a ket multiplied by an amplitude, whereas the quantity on the right is a linear operator multiplying a ket, but linear algebra tells us that these two quantities are equivalent. The operators $|+\rangle\langle +|$ and $|-\rangle\langle -|$ are special operators—they are rank-one projection operators, meaning that they project onto a one-dimensional subspace. Using linearity, we have the following equivalence:

$$|\psi\rangle = (|+\rangle\langle +| + |-\rangle\langle -|)|\psi\rangle. \tag{39}$$

The above equation indicates a seemingly trivial, but important point—the operator $|+\rangle\langle +|+|-\rangle\langle -|$ is equivalent to the identity operator and we can write

$$I = |+\rangle\langle +| + |-\rangle\langle -|, \tag{40}$$

where I stands for the identity operator. This relation is known as the *completeness relation* or the *resolution of the identity*. Given any orthonormal basis, we can always construct a resolution of the identity by summing over the rank-one projection operators formed from each of the orthonormal basis states. For example, the computational basis states give another way to form a resolution of the identity operator:

$$I = |0\rangle\langle 0| + |1\rangle\langle 1|. \tag{41}$$

This simple trick provides a way to find the representation of a quantum state in any basis.

3 Reversible Evolution

Physical systems evolve as time progresses. The application of a magnetic field to an electron can change its spin and pulsing an atom with a laser can excite one of its electrons from a ground state to an excited state. These are only a couple of ways in which physical systems can change.

The Schrödinger equation governs the evolution of a closed quantum system. In this book, we will not even state the Schrödinger equation, but we will instead focus on its major result. *The evolution of a closed quantum system is reversible if we do not learn anything about the state of the system (that is, if we do not measure it).* Reversibility implies that we can determine the input state of an evolution given the output state and knowledge of the evolution. An example of a single-qubit reversible operation is a NOT gate:

$$|0\rangle \rightarrow |1\rangle, \qquad |1\rangle \rightarrow |0\rangle. \tag{42}$$

In the classical world, we would say that the NOT gate merely flips the value of the input classical bit. In the quantum world, the NOT gate flips the basis states $|0\rangle$ and $|1\rangle$. The NOT gate is reversible because we can simply apply the NOT gate again to recover the original input state—the NOT gate is its own inverse.

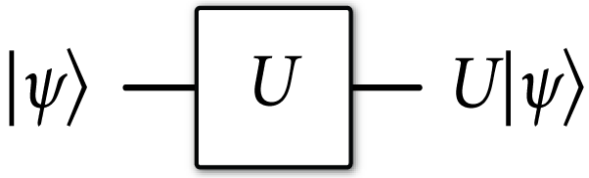

Figure 3: The above figure is a quantum circuit diagram that depicts the evolution of a quantum state $|\psi\rangle$ according to a unitary operator U.

In general, a closed quantum system evolves according to a unitary operator U. Unitary evolution implies reversibility because a unitary operator always possesses an inverse—its inverse is merely $U^\dagger$. This property gives the relations:

$$U^\dagger U = UU^\dagger = I. \tag{43}$$

The unitary property also ensures that evolution preserves the unit-norm constraint (an important requirement for a physical state that we discuss in the section on measurement). Consider applying the unitary operator U to the example qubit state in (2):

$$U|\psi\rangle. \tag{44}$$

Figure 3 depicts a quantum circuit diagram for unitary evolution.

The bra that is dual to the above state is $\langle\psi|U^\dagger$ (we again apply the conjugate transpose operation to get the bra). We showed in (24-27) that every quantum state should have a unit amplitude for being itself. This relation holds for the state $U|\psi\rangle$ because the operator U is unitary:

$$\langle\psi|U^\dagger U|\psi\rangle = \langle\psi|I|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{45}$$

The assumption that a vector always has a unit amplitude for being itself is one of the crucial assumptions of the quantum theory, and the above reasoning demonstrates that unitary evolution complements this assumption.

3.1 Matrix Representations of Operators

We now explore some properties of the NOT gate. Let X denote the operator corresponding to a NOT gate. The action of X on the computational basis states is as follows:

$$X|i\rangle = |i \oplus 1\rangle, \tag{46}$$

where $i = \{0, 1\}$ and $\oplus$ denotes binary addition. Suppose the NOT gate acts on a superposition state:

$$X(\alpha|0\rangle + \beta|1\rangle) \tag{47}$$

By the linearity of the quantum theory, the X operator distributes so that the above expression is equal to the following one:

$$\alpha X|0\rangle + \beta X|1\rangle = \alpha|1\rangle + \beta|0\rangle. \tag{48}$$

Indeed, the NOT gate X merely flips the basis states of any quantum state when represented in the computational basis.

We can determine a *matrix representation* for the operator X by using the bras $\langle 0|$ and $\langle 1|$. Consider the relations in (46). Let us combine the relations with the bra $\langle 0|$:

$$\langle 0|X|0\rangle = \langle 0|1\rangle = 0, \tag{49}$$
$$\langle 0|X|1\rangle = \langle 0|0\rangle = 1. \tag{50}$$

Likewise, we can combine with the bra $\langle 1|$:

$$\langle 1|X|0\rangle = \langle 1|1\rangle = 1, \tag{51}$$
$$\langle 1|X|1\rangle = \langle 1|0\rangle = 0. \tag{52}$$

We can place these entries in a matrix to give a matrix representation of the operator X:

$$\begin{bmatrix} \langle 0|X|0\rangle & \langle 0|X|1\rangle \\ \langle 1|X|0\rangle & \langle 1|X|1\rangle \end{bmatrix}, \tag{53}$$

where we order the rows according to the bras and order the columns according to the kets. We then say that

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{54}$$

and adopt the convention that the symbol X refers to both the operator X and its matrix representation (this is an abuse of notation, but it should be clear from context when X refers to an operator and when it refers to the matrix representation of the operator).

Let us now observe some uniquely quantum behavior. We would like to consider the action of the NOT operator X on the $+/-$ basis. First, let us consider what happens if we operate on the $|+\rangle$ state with the X operator. Recall that the state $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ so that

$$X|+\rangle = X\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \tag{55}$$
$$= \frac{X|0\rangle + X|1\rangle}{\sqrt{2}} \tag{56}$$
$$= \frac{|1\rangle + |0\rangle}{\sqrt{2}} \tag{57}$$
$$= |+\rangle. \tag{58}$$

The above development shows that the state $|+\rangle$ is a special state with respect to the NOT operator X—it is an *eigenstate* of X with *eigenvalue* one. An eigenstate of an operator is one that is invariant under the action of the operator. The coefficient in front of the eigenstate is the *eigenvalue* corresponding to the eigenstate. Under a unitary evolution, the coefficient in front of the eigenstate is just a complex phase, but this global phase has no effect on the observations resulting from a measurement of the state because two quantum states are equivalent up to a differing global phase.

Now, let us consider the action of the NOT operator X on the state $|-\rangle$. Recall that $|-\rangle \equiv 1/\sqrt{2}(|0\rangle - |1\rangle)$. Calculating similarly, we get that

$$X|-\rangle = X\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \tag{59}$$

$$= \frac{X|0\rangle - X|1\rangle}{\sqrt{2}} \tag{60}$$

$$= \frac{|1\rangle - |0\rangle}{\sqrt{2}} \tag{61}$$

$$= -|-\rangle. \tag{62}$$

So the state $|-\rangle$ is also an eigenstate of the operator X, but its eigenvalue is -1.

We can find a matrix representation of the X operator in the $+/-$ basis as well:

$$\begin{bmatrix} \langle +|X|+\rangle & \langle +|X|-\rangle \\ \langle -|X|+\rangle & \langle -|X|-\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{63}$$

This representation demonstrates that the X operator is diagonal with respect to the $+/-$ basis, and therefore, the $+/-$ basis is an *eigenbasis* for the X operator. It is always handy to know the eigenbasis of a unitary operator U because this eigenbasis gives the states that are invariant under an evolution according to U.

Let Z denote the operator that flips states in the $+/-$ basis:

$$Z|+\rangle \rightarrow |-\rangle, \qquad Z|-\rangle \rightarrow |+\rangle. \tag{64}$$

Using an analysis similar to that which we did for the X operator, we can find a matrix representation of the Z operator in the $+/-$ basis:

$$\begin{bmatrix} \langle +|Z|+\rangle & \langle +|Z|-\rangle \\ \langle -|Z|+\rangle & \langle -|Z|-\rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{65}$$

Interestingly, the matrix representation for the Z operator in the $+/-$ basis is the same as that for the X operator in the computational basis. For this reason, we call the Z operator the *phase flip* operator.[2]

We expect the following steps to hold because the quantum theory is a linear theory:

$$Z\left(\frac{|+\rangle + |-\rangle}{\sqrt{2}}\right) = \frac{Z|+\rangle + Z|-\rangle}{\sqrt{2}} = \frac{|-\rangle + |+\rangle}{\sqrt{2}} = \frac{|+\rangle + |-\rangle}{\sqrt{2}}, \tag{66}$$

$$Z\left(\frac{|+\rangle - |-\rangle}{\sqrt{2}}\right) = \frac{Z|+\rangle - Z|-\rangle}{\sqrt{2}} = \frac{|-\rangle - |+\rangle}{\sqrt{2}} = -\left(\frac{|+\rangle - |-\rangle}{\sqrt{2}}\right). \tag{67}$$

The above steps demonstrate that the states $1/\sqrt{2}(|+\rangle + |-\rangle)$ and $1/\sqrt{2}(|+\rangle - |-\rangle)$ are both eigenstates of the Z operators. These states are none other than the respective computational basis states $|0\rangle$ and $|1\rangle$, by inspecting the definitions in (29-30) of the $+/-$ basis. Thus, a matrix representation of the Z operator in the computational basis is

$$\begin{bmatrix} \langle 0|Z|0\rangle & \langle 0|Z|1\rangle \\ \langle 1|Z|0\rangle & \langle 1|Z|1\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{68}$$

[2]A more appropriate name might be the "bit flip in the $+/-$ basis operator," but this name is too long, so we stick with the term "phase flip."

and is a diagonalization of the operator Z. So, the behavior of the Z operator in the computational basis is the same as the behavior of the X operator in the $+/-$ basis.

3.2 The Pauli Matrices

The convention in quantum theory is to take the computational basis as the *standard basis* for representing physical qubits. The standard matrix representation for the above two operators is as follows when we choose the computational basis as the standard basis:

$$X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{69}$$

The identity operator I has the following representation in any basis:

$$I \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{70}$$

Another operator, the Y operator, is a useful one to consider as well. The Y operator has the following matrix representation in the computational basis:

$$Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{71}$$

It is easy to check that $Y = iXZ$, and for this reason, we can think of the Y operator as a combined bit and phase flip. The four matrices I, X, Y, and Z are special for the manipulation of physical qubits and are known as the *Pauli matrices*.

Exercise 2 *Show that the Pauli matrices are all Hermitian, unitary, they square to the identity, and their eigenvalues are* ± 1.

Exercise 3 *Represent the eigenstates of the* Y *operator in the computational basis.*

Exercise 4 *Show that the Pauli matrices either commute or anticommute.*

Exercise 5 *Let us label the Pauli matrices as* $\sigma_0 \equiv I$, $\sigma_1 \equiv X$, $\sigma_2 \equiv Y$, *and* $\sigma_3 \equiv Z$, *Show that* $Tr\{\sigma_i \sigma_j\} = 2\delta_{ij}$ *for all* $i, j \in \{0, \ldots, 3\}$.

4 Measurement

Measurement is another type of evolution that a quantum system can undergo. It is an evolution that allows us to retrieve classical information from a quantum state and thus is the way that we can "read out" information. Suppose that we would like to learn something about the quantum state $|\psi\rangle$ in (2). Nature prevents us from learning anything about the probability amplitudes α and β if we have only one quantum measurement that we can perform. Nature only allows us to measure *observables*. Observables are physical variables such as the position or momentum of a particle. In the quantum theory, we represent observables as Hermitian operators in part because their eigenvalues are real numbers and every measuring device outputs a real number. Examples of qubit observables that we can measure are the Pauli operators X, Y, and Z.

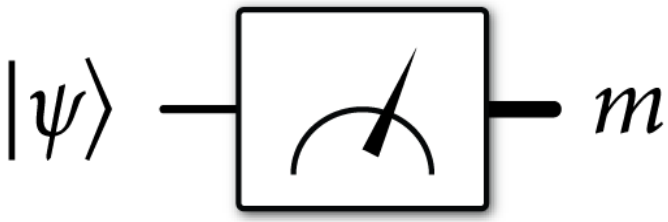

Figure 4: The above figure depicts our diagram of a quantum measurement. Thin lines denote quantum information and thick lines denote classical information. The result of the measurement is to output a classical variable m according to a probability distribution governed by the Born rule of the quantum theory.

Suppose we measure the Z operator. This measurement is called a "measurement in the computational basis" or a "measurement of the Z observable" because we are measuring the eigenvalues of the Z operator. The measurement postulate of the quantum theory, also known as the *Born rule*, states that the system "collapses" into the state $|0\rangle$ with probability $|\alpha|^2$ and collapses into the state $|1\rangle$ with probability $|\beta|^2$. That is, the resulting probabilities are the squares of the probability amplitudes. After the measurement, our measuring apparatus tells us whether the state collapsed into $|0\rangle$ or $|1\rangle$—it returns $+1$ if the resulting state is $|0\rangle$ and returns -1 if the resulting state is $|1\rangle$. These returned values are the eigenvalues of the Z operator. The measurement postulate is the aspect of the quantum theory that makes it probabilistic or "jumpy" and is part of the "strangeness" of the quantum theory. Figure 4 depicts the notation for a measurement that we will use in diagrams throughout this book.

What is the result if we measure the state $|\psi\rangle$ in the $+/-$ basis? Consider that we can represent $|\psi\rangle$ as a superposition of the $|+\rangle$ and $|-\rangle$ states, as given in (35). The measurement postulate then states that a measurement of the X operator gives the state $|+\rangle$ with probability $|\alpha+\beta|^2/2$ and the state $|-\rangle$ with probability $|\alpha-\beta|^2/2$. Quantum interference is now playing a role because the amplitudes α and β interfere with each other. So this effect plays an important role in quantum information theory.

In some cases, the basis states $|0\rangle$ and $|1\rangle$ may not represent the spin states of an electron, but may represent the *location* of an electron. So, a way to interpret this measurement postulate is that the electron "jumps into" one location or another depending on the outcome of the measurement. But what is the state of the electron before the measurement? We will just say in this book that it is in a superposed, indefinite, or unsharp state, rather than trying to pin down a philosophical interpretation. Some might say that the electron is in "two different locations at the same time."

Also, we should stress that we cannot interpret this measurement postulate as meaning that the state is in $|0\rangle$ or $|1\rangle$ with respective probabilities $|\alpha|^2$ and $|\beta|^2$ before the measurement occurs, because this latter scenario is completely classical. The superposition state $\alpha|0\rangle+\beta|1\rangle$ gives fundamentally different behavior from the probabilistic description of a state that is in $|0\rangle$ or $|1\rangle$ with respective probabilities $|\alpha|^2$ and $|\beta|^2$. Suppose that we have the two different descriptions of a state (superposition and probabilistic) and measure the Z operator. We get the same result for both cases—the resulting state is $|0\rangle$ or $|1\rangle$ with respective probabilities $|\alpha|^2$ and $|\beta|^2$.

But now suppose that we measure the X operator. The superposed state gives the result

Quantum State	**Probability of $\|+\rangle$**	**Probability of $\|-\rangle$**
Superposition state	$\|\alpha+\beta\|^2/2$	$\|\alpha-\beta\|^2/2$
Probabilistic description	$1/2$	$1/2$

Table 1: The above table summarizes the differences in probabilities for a quantum state in a superposition $\alpha|0\rangle + \beta|1\rangle$ and a classical state that is a probabilistic mixture of $|0\rangle$ and $|1\rangle$.

from before—we get the state $|+\rangle$ with probability $|\alpha+\beta|^2/2$ and the state $|-\rangle$ with probability $|\alpha-\beta|^2/2$. The probabilistic description gives a much different result. Suppose that the state is $|0\rangle$. We know that $|0\rangle$ is a uniform superposition of $|+\rangle$ and $|-\rangle$:

$$|0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}. \tag{72}$$

So the state collapses to $|+\rangle$ or $|-\rangle$ with equal probability in this case. If the state is $|1\rangle$, then it collapses again to $|+\rangle$ or $|-\rangle$ with equal probabilities. Summing up these probabilities, it follows that a measurement of the X operator gives the state $|+\rangle$ with probability $\left(|\alpha|^2 + |\beta|^2\right)/2 = 1/2$ and gives the state $|-\rangle$ with the same probability. These results are fundamentally different from those where the state is the superposition state $|\psi\rangle$, and experiment after experiment supports the predictions of the quantum theory. Table 1 summarizes the results described in the above paragraph.

Now we consider a "Stern-Gerlach" like argument to illustrate another example of fundamental quantum behavior. The Stern-Gerlach experiment was a crucial one for determining the "strange" behavior of quantum spin states. Suppose we prepare the state $|0\rangle$. If we measure this state in the Z basis, the result is that we always obtain the state $|0\rangle$ because it is a definite Z eigenstate. Suppose now that we measure the X operator. The state $|0\rangle$ is equivalent to a uniform superposition of $|+\rangle$ and $|-\rangle$. The measurement postulate then states that we get the state $|+\rangle$ or $|-\rangle$ with equal probability after performing this measurement. If we then measure the Z operator again, the result is completely random. The Z measurement result is $|0\rangle$ or $|1\rangle$ with equal probability if the result of the X measurement is $|+\rangle$ and the same distribution holds if the result of the X measurement is $|-\rangle$. This argument demonstrates that the measurement of the X operator throws off the measurement of the Z operator. The Stern-Gerlach experiment was one of the earliest to validate the predictions of the quantum theory.

4.1 Probability, Expectation, and Variance of an Operator

We have an alternate, more formal way of stating the measurement postulate that turns out to be more useful for a general quantum system. Suppose that we are measuring the Z operator. The diagonal representation of this operator is

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{73}$$

Consider the operator

$$\Pi_0 \equiv |0\rangle\langle 0|. \tag{74}$$

It is a projection operator because applying it twice has the same effect as applying it once: $\Pi_0^2 = \Pi_0$. It projects onto the subspace spanned by the single vector $|0\rangle$. A similar line of analysis applies to

the projection operator

$$\Pi_1 \equiv |1\rangle\langle 1|. \tag{75}$$

So we can represent the Z operator as $\Pi_0 - \Pi_1$. Performing a measurement of the Z operator is equivalent to asking the question: Is the state $|0\rangle$ or $|1\rangle$? Consider the quantity $\langle\psi|\Pi_0|\psi\rangle$:

$$\langle\psi|\Pi_0|\psi\rangle = \langle\psi|0\rangle\langle 0|\psi\rangle = \alpha^*\alpha = |\alpha|^2. \tag{76}$$

A similar analysis demonstrates that

$$\langle\psi|\Pi_1|\psi\rangle = |\beta|^2. \tag{77}$$

These two quantities then give the probability that the state collapses to $|0\rangle$ or $|1\rangle$.

A more general way of expressing a measurement of the Z basis is to say that we have a set $\{\Pi_i\}_{i\in\{0,1\}}$ of measurement operators that determine the outcome probabilities. These measurement operators also determine the state that results after the measurement. If the measurement result is $+1$, then the resulting state is

$$\frac{\Pi_0|\psi\rangle}{\sqrt{\langle\psi|\Pi_0|\psi\rangle}} = |0\rangle, \tag{78}$$

where we implicitly ignore the irrelevant global phase factor $\frac{\alpha}{|\alpha|}$. If the measurement result is -1, then the resulting state is

$$\frac{\Pi_1|\psi\rangle}{\sqrt{\langle\psi|\Pi_1|\psi\rangle}} = |1\rangle, \tag{79}$$

where we again implicitly ignore the irrelevant global phase factor $\frac{\beta}{|\beta|}$. Dividing by $\sqrt{\langle\psi|\Pi_i|\psi\rangle}$ for $i = 0, 1$ ensures that the state resulting after measurement corresponds to a physical state that has unit norm.

We can also measure any orthonormal basis in this way—this type of projective measurement is called a *von Neumann measurement*. For any orthonormal basis $\{|\phi_i\rangle\}_{i\in\{0,1\}}$, the measurement operators are $\{|\phi_i\rangle\langle\phi_i|\}_{i\in\{0,1\}}$, and the state collapses to $|\phi_i\rangle\langle\phi_i|\psi\rangle/|\langle\phi_i|\psi\rangle|$ with probability $\langle\psi|\phi_i\rangle\langle\phi_i|\psi\rangle = |\langle\phi_i|\psi\rangle|^2$.

Exercise 6 *Determine the set of measurement operators corresponding to a measurement of the X observable.*

We might want to determine the expected measurement result when measuring the Z operator. The probability of getting the $+1$ value corresponding to the $|0\rangle$ state is $|\alpha|^2$ and the probability of getting the -1 value corresponding to the -1 eigenstate is $|\beta|^2$. Standard probability theory then gives us a way to calculate the expected value of a measurement of the Z operator when the state is $|\psi\rangle$:

$$\mathbb{E}[Z] = |\alpha|^2(1) + |\beta|^2(-1) \tag{80}$$

$$= |\alpha|^2 - |\beta|^2. \tag{81}$$

We can formulate an alternate way to write this expectation, by making use of the Dirac notation:

$$\begin{aligned}\mathbb{E}[Z] &= |\alpha|^2(1) + |\beta|^2(-1) && (82)\\ &= \langle\psi|\Pi_0|\psi\rangle + \langle\psi|\Pi_1|\psi\rangle(-1) && (83)\\ &= \langle\psi|\Pi_0 - \Pi_1|\psi\rangle && (84)\\ &= \langle\psi|Z|\psi\rangle && (85)\end{aligned}$$

It is common for physicists to denote the expectation as

$$\langle Z\rangle \equiv \langle\psi|Z|\psi\rangle, \tag{86}$$

when it is understood that the expectation is with respect to the state $|\psi\rangle$. This type of expression is a general one and the next exercise asks you to show that it works for the X and Y operators as well.

Exercise 7 *Show that the expressions $\langle\psi|X|\psi\rangle$ and $\langle\psi|Y|\psi\rangle$ give the respective expectations $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ when measuring the state $|\psi\rangle$ in the respective X and Y basis.*

We also might want to determine the variance of the measurement of the Z operator. Standard probability theory again gives that

$$\text{Var}[Z] = \mathbb{E}\left[Z^2\right] - \mathbb{E}[Z]^2. \tag{87}$$

Physicists denote the standard deviation of the measurement of the Z operator as

$$\Delta Z \equiv \left\langle \left(Z - \langle Z\rangle\right)^2 \right\rangle^{1/2}, \tag{88}$$

and thus the variance is equal to $(\Delta Z)^2$. Physicists often refer to ΔZ as the uncertainty of the observable Z when the state is $|\psi\rangle$.

In order to calculate the variance $\text{Var}[Z]$, we really just need the second moment $\mathbb{E}\left[Z^2\right]$ because we already have the expecation $\mathbb{E}[Z]$:

$$\begin{aligned}\mathbb{E}\left[Z^2\right] &= |\alpha|^2(1)^2 + |\beta|^2(-1)^2 && (89)\\ &= |\alpha|^2 + |\beta|^2. && (90)\end{aligned}$$

We can again calculate this quantity with the Dirac notation. The quantity $\langle\psi|Z^2|\psi\rangle$ is the same as $\mathbb{E}\left[Z^2\right]$ and the next exercise asks you for a proof.

Exercise 8 *Show that $\mathbb{E}\left[X^2\right] = \langle\psi|X^2|\psi\rangle$, $\mathbb{E}\left[Y^2\right] = \langle\psi|Y^2|\psi\rangle$, and $\mathbb{E}\left[Z^2\right] = \langle\psi|Z^2|\psi\rangle$.*

5 Composite Quantum Systems

A single physical qubit is an interesting physical system that exhibits uniquely quantum phenomena, but it is not particularly useful on its own (just as a single classical bit is not particularly useful for classical communication or computation). We can only perform interesting quantum information processing tasks when we combine qubits together. Therefore, we should have a way for describing their behavior when they combine to form a composite quantum system.

Consider two classical bits c_0 and c_1. In order to describe bit operations on the pair of cbits, we write them as an ordered pair (c_1, c_0). The space of all possible bit values is the Cartesian product $\mathbb{Z}_2 \times \mathbb{Z}_2$ of two copies of the set $\mathbb{Z}_2 \equiv \{0, 1\}$:

$$\mathbb{Z}_2 \times \mathbb{Z}_2 \equiv \{(0,0), (0,1), (1,0), (1,1)\}. \tag{91}$$

Typically, we make the abbreviation $c_1 c_0 \equiv (c_1, c_0)$ when representing cbit states.

We can represent the state of two cbits with particular states of qubits. For example, we can represent the two-cbit state 00 with the following mapping:

$$00 \rightarrow |0\rangle|0\rangle. \tag{92}$$

Many times, we make the abbreviation $|00\rangle \equiv |0\rangle|0\rangle$ when representing two-cbit states with qubits. In general, any two-cbit state $c_1 c_0$ has the following representation as a two-qubit state:

$$c_1 c_0 \rightarrow |c_1 c_0\rangle. \tag{93}$$

The above qubit states are not the only possible states that can occur in the quantum theory. By the superposition principle, any possible linear combination of the set of two-cbit states is a possible two-qubit state:

$$|\xi\rangle \equiv \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle. \tag{94}$$

The unit-norm condition $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$ again must hold for the two-qubit state to correspond to a physical quantum state. It is now clear that the Cartesian product is not sufficient for representing two-qubit quantum states because it does not allow for linear combinations of states (just as the mathematics of Boolean algebra is not sufficient to represent single-qubit states).

We again turn to linear algebra to determine a representation that suffices. The *tensor product* is the mathematical operation that gives a sufficient representation of two-qubit quantum states. Suppose we have two two-dimensional vectors:

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \qquad \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}. \tag{95}$$

The tensor product of these two vectors is

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \otimes \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \equiv \begin{bmatrix} a_1 \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \\ b_1 \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_1 a_2 \\ a_1 b_2 \\ b_1 a_2 \\ b_1 b_2 \end{bmatrix}. \tag{96}$$

Recall, from (4-5), the vector representation of the single-qubit states $|0\rangle$ and $|1\rangle$. Using these vector representations and the above definition of the tensor product, the two-qubit basis states have the following vector representations:

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{97}$$

A simple way to remember these representations is that the bits inside the ket index the element equal to one in the vector. For example, the vector representation of $|01\rangle$ has a one as its second element because 01 is the second index for the two-bit strings. The vector representation of the superposition state in (94) is

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}. \tag{98}$$

There are actually many different ways that we can write two-qubit states, and we go through all of these right now. Physicists have developed many shorthands, and it is important to know each of these because they often appear in the literature (we even use different notations depending on the context). We may use any of the following two-qubit notations if the two qubits are local to one party and only one party is involved in a protocol:

$$\alpha|0\rangle \otimes |0\rangle + \beta|0\rangle \otimes |1\rangle + \gamma|1\rangle \otimes |0\rangle + \delta|1\rangle \otimes |1\rangle, \tag{99}$$

$$\alpha|0\rangle|0\rangle + \beta|0\rangle|1\rangle + \gamma|1\rangle|0\rangle + \delta|1\rangle|1\rangle, \tag{100}$$

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle. \tag{101}$$

We can put labels on the qubits if two or more parties are involved in the protocol:

$$\alpha|0\rangle^A \otimes |0\rangle^B + \beta|0\rangle^A \otimes |1\rangle^B + \gamma|1\rangle^A \otimes |0\rangle^B + \delta|1\rangle^A \otimes |1\rangle^B, \tag{102}$$

$$\alpha|0\rangle^A|0\rangle^B + \beta|0\rangle^A|1\rangle^B + \gamma|1\rangle^A|0\rangle^B + \delta|1\rangle^A|1\rangle^B, \tag{103}$$

$$\alpha|00\rangle^{AB} + \beta|01\rangle^{AB} + \gamma|10\rangle^{AB} + \delta|11\rangle^{AB}. \tag{104}$$

This second scenario is different from the first scenario because two spatially separated parties share the two-qubit state. If the state has quantum correlations, then it can be valuable as a communication resource. We go into more detail on this topic in Section 5.6 on *entanglement*.

5.1 Evolution of Composite Systems

The postulate on unitary evolution extends to the two-qubit scenario as well. First, let us establish that the tensor product $A \otimes B$ of two operators A and B is

$$A \otimes B \equiv \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{105}$$

$$\equiv \begin{bmatrix} a_{11}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} & a_{12}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ a_{21}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} & a_{22}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \end{bmatrix} \tag{106}$$

$$= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix}. \tag{107}$$

Consider the two-qubit state in (94). We can perform a NOT gate on the first qubit so that it changes to

$$\alpha|10\rangle + \beta|11\rangle + \gamma|00\rangle + \delta|01\rangle. \tag{108}$$

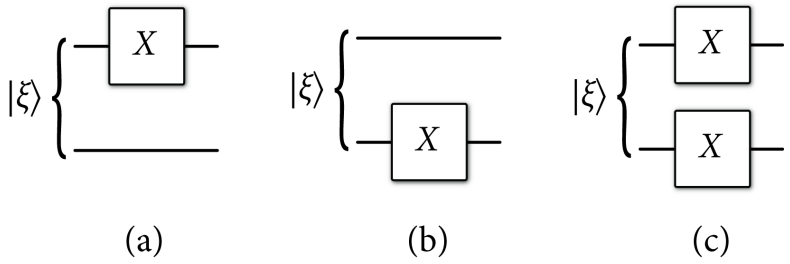

Figure 5: The above figure depicts circuits for the example two-qubit unitaries $X_1 I_2$, $I_1 X_2$, and $X_1 X_2$.

We can likewise flip its second qubit:

$$\alpha|01\rangle + \beta|00\rangle + \gamma|11\rangle + \delta|10\rangle, \tag{109}$$

or flip both at the same time:

$$\alpha|11\rangle + \beta|10\rangle + \gamma|01\rangle + \delta|00\rangle. \tag{110}$$

Figure 5 depicts quantum circuit representations of these operations. These are all reversible operations because applying them again gives the original state in (94). In the first case, we did nothing to the second qubit, and in the second case, we did nothing to the first qubit. The identity operator acts on the qubits that have nothing happen to them.

Let us label the first qubit as "1" and the second qubit as "2." We can then label the operator for the first operation as $X_1 I_2$ because this operator flips the first qubit and does nothing (applies the identity) to the second qubit. We can also label the operators for the second and third operations respectively as $I_1 X_2$ and $X_1 X_2$. The matrix representation of the operator $X_1 I_2$ is the tensor product of the matrix representation of X with the matrix representation of I—this relation similarly holds for the operators $I_1 X_2$ and $X_1 X_2$. We show that it holds for the operator $X_1 I_2$ and ask you to verify the other two cases. We can use the two-qubit computational basis to get a matrix representation for the two-qubit operator $X_1 I_2$:

$$\begin{bmatrix} \langle 00|X_1 I_2|00\rangle & \langle 00|X_1 I_2|01\rangle & \langle 00|X_1 I_2|10\rangle & \langle 00|X_1 I_2|11\rangle \\ \langle 01|X_1 I_2|00\rangle & \langle 01|X_1 I_2|01\rangle & \langle 01|X_1 I_2|10\rangle & \langle 01|X_1 I_2|11\rangle \\ \langle 10|X_1 I_2|00\rangle & \langle 10|X_1 I_2|01\rangle & \langle 10|X_1 I_2|10\rangle & \langle 10|X_1 I_2|11\rangle \\ \langle 11|X_1 I_2|00\rangle & \langle 11|X_1 I_2|01\rangle & \langle 11|X_1 I_2|10\rangle & \langle 11|X_1 I_2|11\rangle \end{bmatrix}$$

$$= \begin{bmatrix} \langle 00|10\rangle & \langle 00|11\rangle & \langle 00|00\rangle & \langle 00|01\rangle \\ \langle 01|10\rangle & \langle 01|11\rangle & \langle 01|00\rangle & \langle 01|01\rangle \\ \langle 10|10\rangle & \langle 10|11\rangle & \langle 10|00\rangle & \langle 10|01\rangle \\ \langle 11|10\rangle & \langle 11|11\rangle & \langle 11|00\rangle & \langle 11|01\rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{111}$$

This last matrix is equal to the tensor product $X \otimes I$ by inspecting the definition of the tensor product for matrices in (105).

Exercise 9 *Show that the matrix representation of the operator $I_1 X_2$ is equal to the tensor product $I \otimes X$. Show the same for $X_1 X_2$ and $X \otimes X$.*

5.2 Probability Amplitudes for Composite Systems

We relied on the orthogonality of the two-qubit computational basis states for evaluating amplitudes such as $\langle 00|10\rangle$ or $\langle 00|00\rangle$ in the above matrix representation. It turns out that there is another way to evaluate these amplitudes that relies only on the orthogonality of the single-qubit computational basis states. Suppose that we have four single-qubit states $|\phi_0\rangle$, $|\phi_1\rangle$, $|\psi_0\rangle$, $|\psi_1\rangle$, and we make the following two-qubit states from them:

$$|\phi_0\rangle \otimes |\psi_0\rangle, \tag{112}$$

$$|\phi_1\rangle \otimes |\psi_1\rangle. \tag{113}$$

We may represent these states equally well as follows:

$$|\phi_0, \psi_0\rangle, \tag{114}$$

$$|\phi_1, \psi_1\rangle. \tag{115}$$

because the Dirac notation is versatile (virtually anything can go inside a ket as long as its meaning is not ambiguous). The bra $\langle \phi_1, \psi_1|$ is dual to the ket $|\phi_1, \psi_1\rangle$, and we can use it to calculate the following amplitude:

$$\langle \phi_1, \psi_1 | \phi_0, \psi_0 \rangle. \tag{116}$$

This amplitude is equivalent to the multiplication of the single-qubit amplitudes:

$$\langle \phi_1, \psi_1 | \phi_0, \psi_0 \rangle = \langle \phi_1 | \phi_0 \rangle \langle \psi_1 | \psi_0 \rangle. \tag{117}$$

Exercise 10 *Verify that the amplitudes* $\{\langle ij|kl\rangle\}_{i,j,k,l\in\{0,1\}}$ *are respectively equal to the amplitudes* $\{\langle i|k\rangle\langle j|l\rangle\}_{i,j,k,l\in\{0,1\}}$. *By linearity, this exercise justifies the relation in (117) (at least for two-qubit states).*

5.3 Controlled Gates

An important two-qubit unitary evolution is the controlled-NOT (CNOT) gate. We consider its classical version first. The classical gate acts on two cbits. It does nothing if the first bit is equal to zero, and flips the second bit if the first bit is equal to one:

$$00 \rightarrow 00, \qquad 01 \rightarrow 01, \qquad 10 \rightarrow 11, \qquad 11 \rightarrow 10. \tag{118}$$

We turn this gate into a quantum gate[3] by demanding that it act in the same way on the two-qubit computational basis states:

$$|00\rangle \rightarrow |00\rangle, \qquad |01\rangle \rightarrow |01\rangle, \qquad |10\rangle \rightarrow |11\rangle, \qquad |11\rangle \rightarrow |10\rangle. \tag{119}$$

This behavior carries over to superposition states as well:

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle \quad \underrightarrow{\text{CNOT}} \quad \alpha|00\rangle + \beta|01\rangle + \gamma|11\rangle + \delta|10\rangle. \tag{120}$$

[3]There are other terms for the action of turning a classical operation into a quantum one. Some examples are "making it coherent," "coherifying," or the quantum gate is a "coherification" of the classical one. The term "coherify" is not a proper English word, but we will use it regardless at certain points.

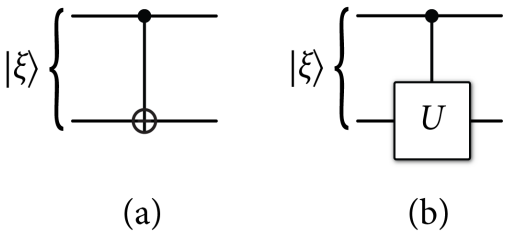

Figure 6: The above figure depicts the circuit diagrams that we use for (a) a CNOT gate and (b) a controlled-U gate.

A useful operator representation of the CNOT gate is

$$\text{CNOT} \equiv |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X. \tag{121}$$

The above representation truly captures the coherent quantum nature of the CNOT gate. In the classical CNOT gate, we can say that it is a conditional gate, in the sense that the gate applies to the second bit conditional on the value of the first bit. In the quantum CNOT gate, the second operation is *controlled* on the basis state of the first qubit (hence the choice of the name "controlled-NOT"). That is, the gate always applies the second operation regardless of the actual qubit state on which it acts.

A controlled-U gate is similar to the CNOT gate in (121). It simply applies the unitary U (assumed to be a single-qubit unitary) to the second qubit, controlled on the first qubit:

$$\text{Controlled-}U \equiv |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U. \tag{122}$$

The control qubit can be controlled with respect to any orthonormal basis $\{|\phi_0\rangle, |\phi_1\rangle\}$:

$$|\phi_0\rangle\langle\phi_0| \otimes I + |\phi_1\rangle\langle\phi_1| \otimes U. \tag{123}$$

Figure 6 depicts the circuit diagrams for a controlled-NOT and controlled-U operation.

Exercise 11 *Verify that the matrix representation of the CNOT gate in the computational basis is*

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{124}$$

Exercise 12 *Consider applying Hadamards to the first and second qubits before and after a CNOT acts on them. Show that this gate is equivalent to a CNOT in the +/− basis (recall that the Z operator flips the +/− basis):*

$$H_1 H_2 \; CNOT \; H_1 H_2 = |+\rangle\langle +| \otimes I + |-\rangle\langle -| \otimes Z. \tag{125}$$

Example 13 *Show that two CNOT gates with the same control qubit commute.*

Exercise 14 *Show that two CNOT gates with the same target qubit commute.*

5.4 The No Cloning Theorem

The no cloning theorem is one of the simplest results in the quantum theory, yet it has some of the most profound consequences. It states that it is impossible to build a *universal copier* of quantum states. A universal copier would be a device that could copy any arbitrary quantum state that is input to it. It may be surprising at first to hear that copying quantum information is impossible because copying classical information is ubiquitous.

We give a simple proof for the no-cloning theorem. Suppose for a contradiction that there is a two-qubit unitary operator U acting as a universal copier of quantum information. That is, if we input an arbitrary state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ as the first qubit and input an ancilla qubit $|0\rangle$ as the second qubit, it "writes" the first qubit to the second qubit slot as follows:

$$\begin{aligned} U|\psi\rangle|0\rangle &= |\psi\rangle|\psi\rangle & (126)\\ &= (\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle) & (127)\\ &= \alpha^2|0\rangle|0\rangle + \alpha\beta|0\rangle|1\rangle + \alpha\beta|1\rangle|0\rangle + \beta^2|1\rangle|1\rangle. & (128) \end{aligned}$$

The copier is universal, meaning that it copies an arbitrary state. In particular, it also copies the states $|0\rangle$ and $|1\rangle$:

$$U|0\rangle|0\rangle = |0\rangle|0\rangle, \quad (129)$$

$$U|1\rangle|0\rangle = |1\rangle|1\rangle. \quad (130)$$

Linearity of the quantum theory then implies that the unitary operator acts on a superposition $\alpha|0\rangle + \beta|1\rangle$ as follows:

$$U(\alpha|0\rangle + \beta|1\rangle)|0\rangle = \alpha|0\rangle|0\rangle + \beta|1\rangle|1\rangle. \quad (131)$$

The result in (128) contradicts the result in (131) because these two expressions do not have to be equal for all α and β:

$$\exists \alpha, \beta : \alpha^2|0\rangle|0\rangle + \alpha\beta|0\rangle|1\rangle + \alpha\beta|1\rangle|0\rangle + \beta^2|1\rangle|1\rangle \neq \alpha|0\rangle|0\rangle + \beta|1\rangle|1\rangle. \quad (132)$$

Thus, unitarity in the quantum theory contradicts the existence of a universal quantum copier.

We would like to stress that this proof does not mean that it is impossible to copy certain quantum states—it only implies the impossibility of a *universal* copier. Another proof of the no-cloning theorem gives insight into the type of states that we can copy. Let us again suppose that a universal copier U exists. Consider two arbitrary states $|\psi\rangle$ and $|\phi\rangle$. If a universal copier U exists, then it performs the following copying operation for both states:

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle, \quad (133)$$

$$U|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle. \quad (134)$$

Consider the probability amplitude $\langle\psi|\langle\psi||\phi\rangle|\phi\rangle$:

$$\langle\psi|\langle\psi||\phi\rangle|\phi\rangle = \langle\psi|\phi\rangle\langle\psi|\phi\rangle = \langle\psi|\phi\rangle^2. \quad (135)$$

The following relation for $\langle\psi|\langle\psi||\phi\rangle|\phi\rangle$ holds as well by using the results in (133) and the unitarity property $U^\dagger U = I$:

$$\begin{aligned} \langle\psi|\langle\psi||\phi\rangle|\phi\rangle &= \langle\psi|\langle 0|U^\dagger U|\phi\rangle|0\rangle & (136)\\ &= \langle\psi|\langle 0||\phi\rangle|0\rangle & (137)\\ &= \langle\psi|\phi\rangle\langle 0|0\rangle & (138)\\ &= \langle\psi|\phi\rangle. & (139) \end{aligned}$$

It then holds that

$$\langle\psi|\langle\psi||\phi\rangle|\phi\rangle = \langle\psi|\phi\rangle^2 = \langle\psi|\phi\rangle, \tag{140}$$

by employing the above two results. The relation $\langle\psi|\phi\rangle^2 = \langle\psi|\phi\rangle$ holds for exactly two cases, $\langle\psi|\phi\rangle = 1$ and $\langle\psi|\phi\rangle = 0$. The first case holds only when the two states are the same state and the second case holds when the two states are orthogonal to each other. Thus, it is impossible to copy quantum information in any other case because we would again contradict unitarity.

The no-cloning theorem has several applications in quantum information processing. First, it underlies the security of the quantum key distribution protocol because it ensures that an attacker cannot copy the quantum states that two parties use to establish a secret key. It finds application in quantum Shannon theory because we can use it to reason about the quantum capacity of a certain quantum channel known as the erasure channel.

5.5 Measurement of Composite Systems

The measurement postulate also extends to composite quantum systems. Suppose again that we have the two-qubit quantum state in (94). By a straightforward analogy with the single-qubit case, we can determine the following amplitudes:

$$\langle 00|\xi\rangle = \alpha, \quad \langle 01|\xi\rangle = \beta, \quad \langle 10|\xi\rangle = \gamma, \quad \langle 11|\xi\rangle = \delta. \tag{141}$$

We can also define the following projection operators

$$\Pi_{00} \equiv |00\rangle\langle 00|, \quad \Pi_{01} \equiv |01\rangle\langle 01|, \quad \Pi_{10} \equiv |10\rangle\langle 10|, \quad \Pi_{00} \equiv |11\rangle\langle 11|, \tag{142}$$

and apply the Born rule to determine the probabilities for each result:

$$\langle\xi|\Pi_{00}|\xi\rangle = |\alpha|^2, \quad \langle\xi|\Pi_{01}|\xi\rangle = |\beta|^2, \quad \langle\xi|\Pi_{10}|\xi\rangle = |\gamma|^2, \quad \langle\xi|\Pi_{11}|\xi\rangle = |\delta|^2. \tag{143}$$

Suppose that we wish to perform a measurement of the Z operator on the first qubit only. What is the set of projection operators that describes this measurement? The answer is similar to what we found for the evolution of a composite system. We apply the identity operator to the second qubit because no measurement occurs on it. Thus, the set of measurement operators is

$$\{\Pi_0 \otimes I, \Pi_1 \otimes I\}, \tag{144}$$

where the definition of Π_0 and Π_1 is in (74-75). The state collapses to

$$\frac{(\Pi_0 \otimes I)|\xi\rangle}{\sqrt{\langle\xi|(\Pi_0 \otimes I)|\xi\rangle}} = \frac{\alpha|00\rangle + \beta|01\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}}, \tag{145}$$

with probability $\langle\xi|(\Pi_0 \otimes I)|\xi\rangle = |\alpha|^2 + |\beta|^2$, and collapses to

$$\frac{(\Pi_1 \otimes I)|\xi\rangle}{\sqrt{\langle\xi|(\Pi_1 \otimes I)|\xi\rangle}} = \frac{\gamma|10\rangle + \delta|11\rangle}{\sqrt{|\gamma|^2 + |\delta|^2}}, \tag{146}$$

with probability $\langle\xi|(\Pi_1 \otimes I)|\xi\rangle = |\gamma|^2 + |\delta|^2$. The divisions by $\sqrt{\langle\xi|(\Pi_0 \otimes I)|\xi\rangle}$ and $\sqrt{\langle\xi|(\Pi_1 \otimes I)|\xi\rangle}$ again ensure that the resulting state is a normalized, physical state.

5.6 Entanglement

Composite quantum systems give rise to the most uniquely quantum phenomenon: *entanglement*. Schrödinger first observed that two or more quantum systems can be entangled and coined the term after noticing some of the bizarre consequences of this phenomenon.[4]

We first consider a simple, unentangled state that two parties, Alice and Bob, may share, in order to see how an unentangled state contrasts with an entangled state. Suppose that they share the state

$$|0\rangle^A|0\rangle^B, \tag{147}$$

where Alice has the qubit in system A and Bob has the qubit in system B. Alice can definitely say that her qubit is in the state $|0\rangle^A$ and Bob can definitely say that his qubit is in the state $|0\rangle^B$. There is nothing really too strange about this scenario.

Now, consider the composite quantum state $|\Phi^+\rangle^{AB}$:

$$\left|\Phi^+\right\rangle^{AB} \equiv \frac{|0\rangle^A|0\rangle^B + |1\rangle^A|1\rangle^B}{\sqrt{2}}. \tag{148}$$

Alice again has possession of the first qubit in system A and Bob has possession of the second qubit in system B. But now, it is not clear from the above description how to determine the individual state of Alice or the individual state of Bob. The above state is really a uniform superposition of the joint state $|0\rangle^A|0\rangle^B$ and the joint state $|1\rangle^A|1\rangle^B$, and it is not possible to describe either Alice's or Bob's individual state in the noiseless quantum theory. We also cannot describe the entangled state $|\Phi^+\rangle^{AB}$ as a product state of the form $|\phi\rangle^A|\psi\rangle^B$.

Exercise 15 *Show that the entangled state* $|\Phi^+\rangle^{AB}$ *has the following representation in the* $+/-$ *basis:*

$$\left|\Phi^+\right\rangle^{AB} = \frac{|+\rangle^A|+\rangle^B + |-\rangle^A|-\rangle^B}{\sqrt{2}}. \tag{149}$$

Figure 7 gives a graphical depiction of entanglement. We use this depiction often throughout this book. Alice and Bob must receive the entanglement in some way, and the diagram indicates that some source distributes the entangled pair to them. It indicates that Alice and Bob are spatially separated and they possess the entangled state after some time. If they share the entangled state in (148), we say that they share one bit of entanglement, or one *ebit*. The term "ebit" implies that there is some way to quantify entanglement.

5.6.1 Entanglement as a Resource

In this book, we are interested in the use of entanglement as a resource. Much of this book concerns the theory of quantum information processing resources and we have a standard notation for the theory of resources. Let us represent the resource of a shared ebit as

$$[qq], \tag{150}$$

[4]Schrodinger actually used the German word "Verschränkung" to describe the phenomenon, which literally translates as "little parts that, though far from one another, always keep the exact same distance from each other." The one-word English translation is "entanglement." Einstein described the "Verschränkung" as a "spukhafte Fernwirkung," most closely translated as "long-distance ghostly effect" or the more commonly stated "spooky action at a distance."

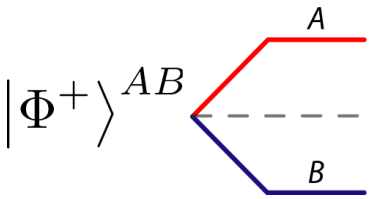

Figure 7: We use the above diagram to depict entanglement shared between two parties A and B. The diagram indicates that a source location creates the entanglement and distributes one system (the red system) to A and the other system (the blue system) to B. The standard unit of entanglement is the ebit in a Bell state $|\Phi^+\rangle \equiv (|00\rangle^{AB} + |11\rangle^{AB})/\sqrt{2}$.

meaning that the ebit is a noiseless, quantum resource shared between two parties. Square brackets indicate a noiseless resource, the letter q indicates a quantum resource, and the two copies of the letter q indicate a two-party resource.

Our first example of the use of entanglement is its role in generating *common randomness*. We define one bit of common randomness as the following probability distribution for two binary random variables X_A and X_B:

$$p_{X_A,X_B}(x_A, x_B) = \frac{1}{2}\delta(x_A, x_B), \tag{151}$$

where δ is the Kronecker delta function. Suppose Alice possesses random variable X_A and Bob possesses random variable X_B. Thus, with probability $1/2$, they either both have a zero or they both have a one. We represent the resource of one bit of common randomness as

$$[cc], \tag{152}$$

indicating that a bit of common randomness is a noiseless, classical resource shared between two parties.

Now suppose that Alice and Bob share an ebit and they decide that they will each measure their qubits in the computational basis. Without loss of generality, suppose that Alice performs a measurement first. Thus, Alice performs a measurement of the Z^A operator, meaning that she measures $Z^A \otimes I^B$ (she cannot perform anything on Bob's qubit because they are spatially separated). The projection operators for this measurement are the same from (144), and they project the joint state. Just before Alice looks at her measurement result, she does not know the outcome, and we can describe the system as being in the following ensemble of states:

$$|0\rangle^A|0\rangle^B \text{ with probability } \frac{1}{2}, \tag{153}$$

$$|1\rangle^A|1\rangle^B \text{ with probability } \frac{1}{2}. \tag{154}$$

The interesting thing about the above ensemble is that Bob's result is already determined even before he measures, just after Alice's measurement occurs. Suppose that Alice knows the result of her measurement is $|0\rangle^A$. When Bob measures his system, he obtains the state $|0\rangle^B$ with probability one and *Alice knows that he has measured this result*. Additionally, Bob knows that Alice's state is $|0\rangle^A$ if he obtains $|0\rangle^B$. The same results hold if Alice knows that the result of her measurement

is $|1\rangle^A$. Thus, this protocol is a method for them to generate one bit of common randomness as defined in (151).

We can phrase the above protocol as the following *resource inequality*:

$$[qq] \geq [cc]. \tag{155}$$

The interpretation of the above resource inequality is *that there exists a protocol which generates the resource on the right by consuming the resource on the left and using only local operations*, and for this reason, the resource on the left is a stronger resource than the one on the right. The theory of resource inequalities plays a prominent role in this book and is a useful shorthand for expressing quantum protocols.

A natural question is to wonder if there exists a protocol to generate entanglement from common randomness. It is not possible to do so and one reason for this inequivalence of resources is another type of inequality (different from the resource inequality mentioned above), called a Bell's inequality. In short, Bell's theorem places an upper bound on the correlations present in any two classical systems. Entanglement violates this inequality, showing that it has no known classical equivalent. Thus, entanglement is a strictly stronger resource than common randomness and the resource inequality in (155) only holds in the given direction.

Common randomness is a resource in classical information theory, and may be useful in some scenarios, but it is actually a rather weak resource. Surely, generating common randomness is not the only use of entanglement. It turns out that we can construct far more exotic protocols such as the teleportation protocol or the super-dense coding protocol by combining the resource of entanglement with other resources. We discuss these protocols later on.

Exercise 16 *Use the representation of the ebit in Exercise 15 to show that Alice and Bob can measure the X operator to generate common randomness. This ability to obtain common randomness by both parties measuring in either the Z or X basis is the basis for an entanglement-based secret key distribution protocol.*

Exercise 17 (Cloning implies signaling) *Prove that if a universal quantum cloner were to exist, then it would be possible for Alice to signal to Bob faster than the speed of light by exploiting only the ebit state $|\Phi^+\rangle^{AB}$ shared between them and no communication. That is, show the existence of a protocol that would allow for this. (Hint: One possibility is for Alice to measure the X or Z Pauli operator locally on her share of the ebit, and then for Bob to exploit the universal quantum cloner. Consider the representation of the ebit in (148) and (149).)*

5.6.2 Entanglement in the CHSH Game

One of the simplest means for demonstrating the power of entanglement is with a two-player game known as the CHSH game (after Clauser, Horne, Shimony, and Holt). We first present the rules of the game, and then we find an upper bound on the probability that players sharing classical correlations can win. We finally leave it as an exercise to show that players sharing a maximally entangled Bell state $|\Phi^+\rangle$ can have an approximately 10% higher chance of winning the game with a quantum strategy.

The players of the game are Alice and Bob. The game begins with a referee selecting two bits x and y uniformly at random. He then sends x to Alice and y to Bob. Alice and Bob are not allowed to communicate in any way at this point. Alice sends back to the referee a bit a, and Bob sends

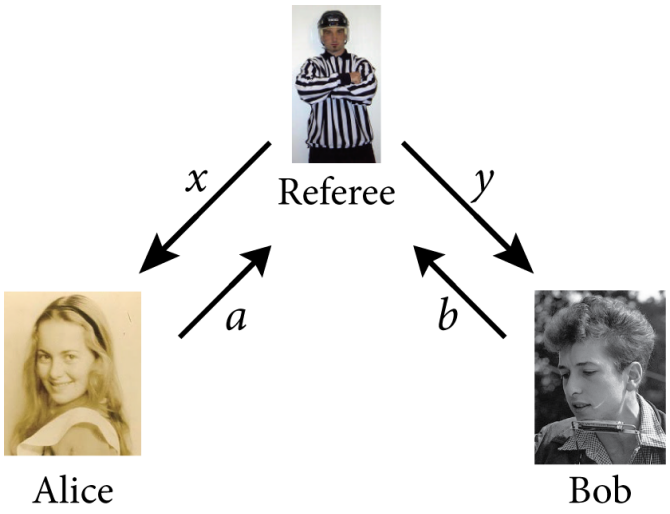

Figure 8: A depiction of the CHSH game. The referee distributes the bits x and y to Alice and Bob in the first round. In the second round, Alice and Bob return the bits a and b to the referee.

back a bit b. Since they are spatially separated, Alice's bit a can only depend on x, and similarly, Bob's bit b can only depend on y. The referee then determines if the AND of x and y is equal to the exclusive OR of a and b. If so, then Alice and Bob win the game. That is, the winning condition is

$$x \wedge y = a \oplus b. \tag{156}$$

Figure 8 depicts the CHSH game.

Before the game begins, Alice and Bob are allowed to coordinate on a strategy. A deterministic strategy would have Alice select a bit a_x conditioned on the bit x that she receives, and similarly, Bob would select a bit b_y conditioned on y. The following table presents the winning conditions for the four different values of x and y with this deterministic strategy:

x	y	$x \wedge y$	$= a_x \oplus b_y$
0	0	0	$= a_0 \oplus b_0$
0	1	0	$= a_0 \oplus b_1$
1	0	0	$= a_1 \oplus b_0$
1	1	1	$= a_1 \oplus b_1$

. (157)

Though, we can observe that it is impossible for them to always win. If we add the entries in the column $x \wedge y$, the binary sum is equal to one, while if we add the entries in the column $= a_x \oplus b_y$, the binary sum is equal to zero. Thus, it is impossible for all of these equations to be satisfied. At most, only three out of four of them can be satisfied, so that the maximal winning probability with a classical deterministic strategy is at most $3/4$. This upper bound also serves as an upper bound on the winning probability for the case in which they employ a randomized strategy coordinated by shared randomness—any such strategy would just be a convex combination of deterministic strategies. We can then see that a strategy for them to achieve this upper bound is for Alice and Bob always to return $a = 0$ and $b = 0$ no matter the values of x and y.

Interestingly, if Alice and Bob share a maximally entangled state, they can achieve a higher winning probability than if they share classical correlations only. This is one demonstration of the power of entanglement, and we leave it as an exercise to prove that the following quantum strategy achieves a winning probability of $\cos^2(\pi/8) \approx 0.85$ in the CHSH game.

Exercise 18 *Suppose that Alice and Bob share a maximally entangled state $|\Phi^+\rangle$. Show that the following strategy has a winning probability of $\cos^2(\pi/8)$. If Alice receives $x = 0$ from the referee, then she performs a measurement of Pauli Z on her system and returns the outcome as a. If she receives $x = 1$, then she performs a measurement of Pauli X and returns the outcome as a. If Bob receives $y = 0$ from the referee, then he performs a measurement of $(X + Z)/\sqrt{2}$ on his system and returns the outcome as b. If Bob receives $y = 1$ from the referee, then he performs a measurement of $(Z - X)/\sqrt{2}$ and returns the outcome as b.*

5.6.3 The Bell States

There are other useful entangled states besides the standard ebit. Suppose that Alice performs a Z^A operation on her half of the ebit $|\Phi^+\rangle^{AB}$. Then the resulting state is

$$\big|\Phi^-\big\rangle^{AB} \equiv \frac{1}{\sqrt{2}}\Big(|00\rangle^{AB} - |11\rangle^{AB}\Big). \tag{158}$$

Similarly, if Alice performs an X operator or a Y operator, the global state transforms to the following respective states (up to a global phase):

$$\big|\Psi^+\big\rangle^{AB} \equiv \frac{1}{\sqrt{2}}\Big(|01\rangle^{AB} + |10\rangle^{AB}\Big), \tag{159}$$

$$\big|\Psi^-\big\rangle^{AB} \equiv \frac{1}{\sqrt{2}}\Big(|01\rangle^{AB} - |10\rangle^{AB}\Big). \tag{160}$$

The states $|\Phi^+\rangle^{AB}$, $|\Phi^-\rangle^{AB}$, $|\Psi^+\rangle^{AB}$, and $|\Psi^-\rangle^{AB}$ are known as the *Bell states* and are the most important entangled states for a two-qubit system. They form an orthonormal basis, called the *Bell basis*, for a two-qubit space. We can also label the Bell states as

$$|\Phi_{zx}\rangle^{AB} \equiv \big(Z^A\big)^z \big(X^A\big)^x \big|\Phi^+\big\rangle^{AB}, \tag{161}$$

where the two-bit binary number zx indicates whether Alice applies I^A, Z^A, X^A, or ZX^A. Then the states $|\Phi_{00}\rangle^{AB}$, $|\Phi_{01}\rangle^{AB}$, $|\Phi_{10}\rangle^{AB}$, and $|\Phi_{11}\rangle^{AB}$ are in correspondence with the respective states $|\Phi^+\rangle^{AB}$, $|\Psi^+\rangle^{AB}$, $|\Phi^-\rangle^{AB}$, and $|\Psi^-\rangle^{AB}$.

Exercise 19 *Show that the Bell states form an orthonormal basis:*

$$\langle \Phi_{zx}|\Phi_{z'x'}\rangle = \delta\big(z,z'\big)\delta\big(x,x'\big). \tag{162}$$

Exercise 20 *Show that the following identities hold:*

$$|00\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(\big|\Phi^+\big\rangle^{AB} + \big|\Phi^-\big\rangle^{AB}\Big), \tag{163}$$

$$|01\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(\big|\Psi^+\big\rangle^{AB} + \big|\Psi^-\big\rangle^{AB}\Big), \tag{164}$$

$$|10\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(\big|\Psi^+\big\rangle^{AB} - \big|\Psi^-\big\rangle^{AB}\Big), \tag{165}$$

$$|11\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(\big|\Phi^+\big\rangle^{AB} - \big|\Phi^-\big\rangle^{AB}\Big). \tag{166}$$

Exercise 21 *Show that the following identities hold by using the relation in (161):*

$$\big|\Phi^+\big\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(|{+}{+}\rangle^{AB} + |{-}{-}\rangle^{AB}\Big), \tag{167}$$

$$\big|\Phi^-\big\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(|{-}{+}\rangle^{AB} + |{+}{-}\rangle^{AB}\Big), \tag{168}$$

$$\big|\Psi^+\big\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(|{+}{+}\rangle^{AB} - |{-}{-}\rangle^{AB}\Big), \tag{169}$$

$$\big|\Psi^-\big\rangle^{AB} = \frac{1}{\sqrt{2}}\Big(|{-}{+}\rangle^{AB} - |{+}{-}\rangle^{AB}\Big). \tag{170}$$

Entanglement is one of the most useful resources in quantum computing, quantum communication, and in the setting of quantum Shannon theory that we explore in this book. Our goal in this book is merely to study entanglement as a resource, but there are many other aspects of entanglement that one can study, such as measures of entanglement, multiparty entanglement, and generalized Bell's inequalities.

6 Summary and Extensions to Qudit States

We now end our overview of the noiseless quantum theory by summarizing its main postulates in terms of quantum states that are on d-dimensional systems. Such states are called *qudit states*, in analogy with the name "qubit" for two-dimensional quantum systems.

6.1 Qudits

A qudit state $|\psi\rangle$ is an arbitrary superposition of some set of orthonormal basis states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$ for a d-dimensional quantum system:

$$|\psi\rangle \equiv \sum_{j=0}^{d-1} \alpha_j |j\rangle. \tag{171}$$

The amplitudes α_j obey the normalization condition $\sum_{j=0}^{d-1} |\alpha_j|^2 = 1$.

6.2 Unitary Evolution

The first postulate of the quantum theory is that we can perform a unitary (reversible) evolution U on this state. The resulting state is

$$U|\psi\rangle, \tag{172}$$

meaning that we apply the operator U to the state $|\psi\rangle$.

One example of a unitary evolution is the cyclic shift operator $X(x)$ that acts on the orthonormal states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$ as follows:

$$X(x)|j\rangle = |x \oplus j\rangle, \tag{173}$$

where $\oplus$ is a cyclic addition operator, meaning that the result of the addition is $(x + j) \bmod(d)$. Notice that the X Pauli operator has a similar behavior on the qubit computational basis states because

$$X|i\rangle = |i \oplus 1\rangle, \tag{174}$$

for $i \in \{0, 1\}$. Therefore, the operator $X(x)$ is one qudit analog of the X Pauli operator.

Exercise 22 *Show that the inverse of $X(x)$ is $X(-x)$.*

Exercise 23 *Show that the matrix representation $X(x)$ of the $X(x)$ operator is a matrix with elements*

$$[X(x)]_{i,j} = \delta_{i,j\oplus x}. \tag{175}$$

Another example of a unitary evolution is the *phase operator* $Z(z)$. It applies a state-dependent phase to a basis state. It acts as follows on the qudit computational basis states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$:

$$Z(z)|j\rangle = \exp\{i2\pi zj/d\}|j\rangle. \tag{176}$$

This operator is the qudit analog of the Pauli Z operator. The d^2 operators $\{X(x)Z(z)\}_{x,z\in\{0,\ldots,d-1\}}$ are known as the *Heisenberg-Weyl operators*.

Exercise 24 *Show that $Z(1)$ is equivalent to the Pauli Z operator for the case that the dimension $d = 2$.*

Exercise 25 *Show that the inverse of $Z(z)$ is $Z(-z)$.*

Exercise 26 *Show that the matrix representation of the phase operator $Z(z)$ is*

$$[Z(z)]_{j,k} = \exp\{i2\pi zj/d\}\delta_{j,k}. \tag{177}$$

In particular, this result implies that the $Z(z)$ operator has a diagonal matrix representation with respect to the qudit computational basis states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$. Thus, the qudit computational basis states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$ are eigenstates of the phase operator $Z(z)$ (similar to the qubit computational basis states being eigenstates of the Pauli Z operator). The eigenvalue corresponding to the eigenstate $|j\rangle$ is $\exp\{i2\pi zj/d\}$.

Exercise 27 *Show that the eigenstates $|l\rangle_X$ of the cyclic shift operator $X(1)$ are the Fourier-transformed states $|l\rangle_X$ where*

$$|l\rangle_X \equiv \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\{i2\pi lj/d\}|j\rangle, \tag{178}$$

l is an integer in the set $\{0, \ldots, d-1\}$, and the subscript X for the state $|l\rangle_X$ indicates that it is an X eigenstate. Show that the eigenvalue corresponding to the state $|l\rangle_X$ is $\exp\{-i2\pi l/d\}$. Conclude that these states are also eigenstates of the operator $X(x)$, but the corresponding eigenvalues are $\exp\{-i2\pi lx/d\}$.

Exercise 28 *Show that the $+/-$ basis states are a special case of the states in (178) when $d = 2$.*

Exercise 29 *The Fourier transform operator F is the qudit analog of the Hadamard H. We define it to take Z eigenstates to X eigenstates.*

$$F \equiv \sum_{j=0}^{d-1} |j\rangle_X \langle j|_Z, \tag{179}$$

where the subscript Z indicates a Z eigenstate. It performs the following transformation on the qudit computational basis states:

$$|j\rangle \rightarrow \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \exp\{i2\pi jk/d\}|k\rangle. \tag{180}$$

Show that the following relations hold for the Fourier transform operator F:

$$FX(x)F^\dagger = Z(x), \tag{181}$$

$$FZ(z)F^\dagger = X(-z). \tag{182}$$

Exercise 30 *Show that the commutation relations of the cyclic shift operator $X(x)$ and the phase operator $Z(z)$ are as follows:*

$$X(x_1)Z(z_1)X(x_2)Z(z_2) = \exp\{2\pi i(z_1x_2 - x_1z_2)/d\}X(x_2)Z(z_2)X(x_1)Z(z_1). \tag{183}$$

You can get this result by first showing that

$$X(x)Z(z) = \exp\{-2\pi izx/d\}Z(z)X(x). \tag{184}$$

6.3 Measurement of Qudits

Measurement of qudits is similar to measurement of qubits. Suppose that we have some state $|\psi\rangle$. Suppose further that we would like to measure some Hermitian operator A with the following diagonalization:

$$A = \sum_j f(j)\Pi_j, \tag{185}$$

where $\Pi_j \Pi_k = \Pi_j \delta_{j,k}$, and $\sum_j \Pi_j = I$. A measurement of the operator A then returns the result j with the following probability:

$$p(j) = \langle\psi|\Pi_j|\psi\rangle, \tag{186}$$

and the resulting state is

$$\frac{\Pi_j|\psi\rangle}{\sqrt{p(j)}}. \tag{187}$$

The calculation of the expectation of the operator A is similar to how we calculate in the qubit case:

$$\mathbb{E}[A] = \sum_j f(j)\langle\psi|\Pi_j|\psi\rangle \tag{188}$$

$$= \langle\psi| \sum_j f(j)\Pi_j|\psi\rangle \tag{189}$$

$$= \langle\psi|A|\psi\rangle. \tag{190}$$

We give two quick examples of qudit operators that we might like to measure. The operators $X(1)$ and $Z(1)$ are not completely analogous to the respective Pauli X and Pauli Z operators because $X(1)$ and $Z(1)$ are not Hermitian. Thus, we cannot directly measure these operators. Instead, we construct operators that are essentially equivalent to "measuring the operators" $X(1)$ and $Z(1)$. Let us first consider the $Z(1)$ operator. Its eigenstates are the qudit computational basis states $\{|j\rangle\}_{j\in\{0,\ldots,d-1\}}$. We can form the operator $M_{Z(1)}$ as

$$M_{Z(1)} \equiv \sum_{j=0}^{d-1} j|j\rangle\langle j|. \tag{191}$$

Measuring this operator is equivalent to measuring in the qudit computational basis. The expectation of this operator for a qudit $|\psi\rangle$ in the state in (171) is

$$\mathbb{E}\big[M_{Z(1)}\big] = \langle\psi|M_{Z(1)}|\psi\rangle \tag{192}$$

$$= \sum_{j'=0}^{d-1} \big\langle j'\big|\alpha_{j'}^* \sum_{j=0}^{d-1} j|j\rangle\langle j| \sum_{j''=0}^{d-1} \alpha_{j''}\big|j''\big\rangle \tag{193}$$

$$= \sum_{j',j,j''=0}^{d-1} j\alpha_{j'}^*\alpha_{j''}\big\langle j'|j\big\rangle\big\langle j|j''\big\rangle \tag{194}$$

$$= \sum_{j=0}^{d-1} j|\alpha_j|^2. \tag{195}$$

Similarly, we can construct an operator $M_{X(1)}$ for "measuring the operator $X(1)$" by using the eigenstates $|j\rangle_X$ of the $X(1)$ operator:

$$M_{X(1)} \equiv \sum_{j=0}^{d-1} j|j\rangle_X\langle j|_X. \tag{196}$$

We leave it as an exercise to determine the expectation when measuring the $M_{X(1)}$ operator.

Exercise 31 *Suppose the qudit is in the state $|\psi\rangle$ in (171). Show that the expectation of the $M_{X(1)}$ operator is*

$$\mathbb{E}\big[M_{X(1)}\big] = \frac{1}{d}\sum_{j=0}^{d-1} j \left| \sum_{j'=0}^{d-1} \alpha_{j'} \exp\big\{-i2\pi j' j/d\big\} \right|^2 . \tag{197}$$

Hint: First show that we can represent the state $|\psi\rangle$ in the $X(1)$ eigenbasis as follows:

$$|\psi\rangle = \sum_{l=0}^{d-1} \frac{1}{\sqrt{d}} \left(\sum_{j=0}^{d-1} \alpha_j \exp\{-i2\pi lj/d\} \right) |l\rangle_X . \tag{198}$$

6.4 Composite Systems of Qudits

We can define a system of multiple qudits again by employing the tensor product. A general two-qudit state on systems A and B has the following form:

$$|\xi\rangle^{AB} \equiv \sum_{j,k=0}^{d-1} \alpha_{j,k} |j\rangle^A |k\rangle^B . \tag{199}$$

Evolution of two-qudit states is similar as before. Suppose Alice applies a unitary U^A to her qudit. The result is as follows:

$$\big(U^A \otimes I^B\big)|\xi\rangle^{AB} = \big(U^A \otimes I^B\big) \sum_{j,k=0}^{d-1} \alpha_{j,k} |j\rangle^A |k\rangle^B \tag{200}$$

$$= \sum_{j,k=0}^{d-1} \alpha_{j,k} \Big(U^A |j\rangle^A\Big) |k\rangle^B , \tag{201}$$

which follows by linearity. Bob applying a local unitary U^B has a similar form. The application of some global unitary U^{AB} is as follows:

$$U^{AB}|\xi\rangle^{AB} . \tag{202}$$

6.4.1 The Qudit Bell States

Two-qudit states can be entangled as well. The maximally-entangled qudit state is as follows:

$$|\Phi\rangle^{AB} \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle^A |i\rangle^B . \tag{203}$$

When Alice possesses the first qudit and Bob possesses the second qudit and they are also separated in space, the above state is a resource known as an *edit* (pronounced "ee · dit"). It is useful in the qudit versions of the teleportation protocol and the super-dense coding protocol discussed later on.

Consider applying the operator $X(x)Z(z)$ to Alice's side of the maximally entangled state $|\Phi\rangle^{AB}$. We use the following notation:

$$|\Phi_{x,z}\rangle^{AB} \equiv \big(X^A(x) Z^A(z) \otimes I^B\big)|\Phi\rangle^{AB} . \tag{204}$$

The d^2 states $\left\{|\Phi_{x,z}\rangle^{AB}\right\}_{x,z=0}^{d-1}$ are known as the qudit Bell states and are important in qudit quantum protocols and in quantum Shannon theory. Exercise 32 asks you to verify that these states form a complete, orthonormal basis. Thus, one can measure two qudits in the qudit Bell basis.

Similar to the qubit case, it is straightforward to see that the qudit state can generate a *dit* of common randomness by extending the arguments in Section 5.6.1. We end our review of the noiseless quantum theory with some exercises. The transpose trick is one of our most important tools for manipulating maximally entangled states.

Exercise 32 *Show that the set of states* $\left\{|\Phi_{x,z}\rangle^{AB}\right\}_{x,z=0}^{d-1}$ *form a complete, orthonormal basis:*

$$\langle\Phi_{x',z'}|\Phi_{x,z}\rangle = \delta_{x,x'}\delta_{z,z'}, \tag{205}$$

$$\sum_{x,z=0}^{d-1} |\Phi_{x,z}\rangle\langle\Phi_{x,z}| = I^{AB}. \tag{206}$$

Exercise 33 (Transpose Trick) *Show that the following "transpose trick" or "ricochet" property holds for a maximally entangled state* $|\Phi\rangle^{AB}$ *and any matrix* M*:*

$$\left(M^A \otimes I^B\right)|\Phi\rangle^{AB} = \left(I^A \otimes \left(M^T\right)^B\right)|\Phi\rangle^{AB}. \tag{207}$$

The implication is that some local action of Alice on $|\Phi\rangle^{AB}$ *is equivalent to Bob performing the transpose of this action on his half of the state.*