2

Quantum mechanics: Hilbert space formalism

Classical mechanics can describe physical properties of macroscopic objects, whereas quantum mechanics can describe physical properties at the microscopic scale. Many quantum mechanical phenomena are counter-intuitive, and researchers have developed mathematical models to explain these phenomena. In particular, the Hilbert space model has been used successfully to explain and predict behaviors of quantum systems. Quantum information and quantum computation concern the use of quantum properties to store, transmit, and process information. It is important to understand the basic rules governing quantum systems. These basic rules are summarized as the postulates of quantum mechanics, which cannot be proven theoretically but they are justified through empirical facts and by numerous experiments. By the Hilbert space model, one can use vectors and linear maps to give interpretation to these postulates, and explain the counter-intuitive phenomena. Furthermore, one can develop theory and design algorithms in quantum information and quantum computing using mathematical tools if one has a good understanding of the mathematical model of quantum mechanics. There are different sets of postulates for quantum mechanics. Here we give one that turns out to be most convenient in the study of quantum information and quantum computation. We will use the matrix concepts presented in Chapter 1 to illustrate the postulates of quantum mechanics and explain how quantum properties are used in quantum information and quantum computation in our subsequent discussions. We recommend [1, 2, 3, 4, 5, 6] for a general introduction to quantum mechanics.

2.1 Postulates of Quantum Mechanics

Quantum mechanics was discovered in early 20th century. In spite of its long history, the interpretation of the wave function, which describes the state of a quantum system, remains an open question. Here we adopt the most popular one, called the **Copenhagen interpretation**.

A1 A pure state in quantum mechanics is represented in terms of a normalized vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ (a complex vector space with an

inner product): $\langle\psi|\psi\rangle = 1$. Suppose two states $|\psi_1\rangle$ and $|\psi_2\rangle$ are physical states of the system. Then their linear superposition $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ ($c_k \in \mathbb{C}$) is also a possible state of the same system. This is called the **superposition principle**.

A2 For any physical quantity (i.e., **observable**) a, such as energy, spin, position and momentum, there exists a corresponding Hermitian operator A acting on the Hilbert space $\mathcal{H}$. When we measure a in the state $|\psi\rangle$, we obtain one of the eigenvalues λ_j of the operator A and the state undergoes an abrupt change to the corresponding eigenvector $|\lambda_j\rangle$. This phenomenon is called the **collapse of wave function**. If we prepare an ensemble of many identical state $|\psi\rangle$ and measure a, the probability of obtaining the outcome $|\lambda_j\rangle$ is $p_j = |\langle\lambda_j|\psi\rangle|^2$.

Let us expand $|\psi\rangle$ in terms of a complete orthonormal basis $\{|\lambda_j\rangle\}$ as $|\psi\rangle = \sum_k c_k|\lambda_k\rangle$, where $c_k = \langle\lambda_k|\psi\rangle \in \mathbb{C}$ is called the **probability amplitude**, and consider $\langle\psi|A|\psi\rangle$. We find

$$\langle\psi|A|\psi\rangle = \sum_{j,k} c_j^* c_k \langle\lambda_j|A|\lambda_k\rangle = \sum_{j,k} c_j^* c_k \lambda_k \delta_{jk} = \sum_k \lambda_k |c_k|^2.$$

Since $|c_k|^2 = |\langle\lambda_k|\psi\rangle|^2$ is the probability of observing λ_k, the above identity shows that the expectation value of measurement outcome of a is

$$\langle a\rangle = \langle\psi|A|\psi\rangle. \tag{2.1}$$

Note that the average is over many measurements of a with respect to many copies of $|\psi\rangle$ and only one outcome λ_j is obtained for each measurement. Note also that $\sum_k p_k = \sum_k |c_k|^2 = 1$ is guaranteed because of the normalization condition.

A3 The time evolution of a state is governed by the **Schrödinger equation**

$$i\hbar\frac{d|\psi\rangle}{dt} = H|\psi\rangle, \tag{2.2}$$

where $\hbar$ is a physical constant known as the **Planck constant** and H is a Hermitian operator (matrix) corresponding to the energy of the system and is called the **Hamiltonian**.

Several comments are in order.

- In Axiom A1, the phase of the vector may be chosen arbitrarily; $|\psi\rangle$ in fact represents the "ray" $\{e^{i\alpha}|\psi\rangle \ |\alpha \in \mathbb{R}\}$. This is called the **ray representation**. In other words, we can totally ignore the overall phase of a vector since it has no observable consequence. For example, neither the probability $|\langle\lambda_k|\psi\rangle|^2$ nor the average $\langle\psi|A|\psi\rangle$ depends on the phase. Note, however, that the *relative* phase of two different states is meaningful. Although $|\langle\phi|e^{i\alpha}\psi\rangle|^2$ is independent of α, $|\langle\phi|\psi_1 + e^{i\alpha}\psi_2\rangle|^2$ does depend on α.

- Axiom A2 is the most counter-intuitive quantum phenomenon which puzzles people. Here, we explain its meaning in mathematical terms. First, an expectation value of a can be viewed as a numerical quantity assigned to a quantum system represented by a state $|\psi\rangle$. Since $|\psi\rangle$ and $e^{i\alpha}|\psi\rangle$ represent the same state, we want to assign this value to $|\psi\rangle\langle\psi|$, which is independent of the phase α. (As we shall see, this is indeed to the density operator representation of a quantum state.) It is believed this process is linear on the linear space spanned by all density operators. By Hilbert space theory, specifically, the Riesz representation theorem of linear functional, there is a Hermitian operator A such that the measured quantity is $\langle\psi|A|\psi\rangle$. Suppose $\mathcal{H}$ has dimension n, then we can identify it with $\mathbb{C}^n$. The Hermitian operator A has a spectral decomposition $A = \sum_{j=1}^n \lambda_j|\lambda_j\rangle\langle\lambda_j|$. One can expand $|\psi\rangle$ in terms of $|\lambda_j\rangle$ as $|\psi\rangle = \sum_{j=1}^n c_j|\lambda_j\rangle$, where $c_j = \langle\lambda_j|\psi\rangle$ for $j = 1, \ldots, n$. Then the probability of observing λ_j upon measurement of a is $|c_i|^2$, and therefore the expectation value after many measurements is

$$\langle\psi|A|\psi\rangle = \sum_{r,s} c_r^* c_s \langle\lambda_r|A|\lambda_s\rangle = \sum_{r,s} c_r^* c_s \lambda_s \delta_{rs} = \sum_r \lambda_r |c_r|^2.$$

  This measurement is called the **projective measurement**. Any particular outcome λ_j will be found with the probability

$$|c_j|^2 = \langle\psi|P_j|\psi\rangle, \tag{2.3}$$

  where $P_j = |\lambda_j\rangle\langle\lambda_j|$ is the projection operator, and the state immediately after the measurement is $|\lambda_j\rangle$ or equivalently

$$\frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}, \tag{2.4}$$

  where the overall phase has been ignored.

  The discussion is similar but more involved if $\mathcal{H}$ has an infinite dimension.

- The Schrödinger equation (2.2) in Axiom A3 is formally solved to yield

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle, \tag{2.5}$$

  if the Hamiltonian H is time-independent, while

$$|\psi(t)\rangle = \mathcal{T}\exp\left[-\frac{i}{\hbar}\int_0^t H(t)dt\right]|\psi(0)\rangle \tag{2.6}$$

  if H depends on t, where $\mathcal{T}$ is the time-ordering operator defined by

$$\mathcal{T}[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2), & t_1 > t_2 \\ B(t_2)A(t_1), & t_2 \geq t_1 \end{cases},$$

for a product of two operators. Note that $\mathcal{T}$ is required since $H(t_1)$ and $H(t_2)$ do not commute with each other in general if $t_1 \neq t_2$. Generalization to products of more than two operators should be obvious. We write Eqs. (2.5) and (2.6) as

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle, \qquad \text{where } U(t) = \mathcal{T}\exp\left[-\frac{i}{\hbar}\int_0^t H(t)dt\right].$$

The operator $U(t) : |\psi(0)\rangle \mapsto |\psi(t)\rangle$, which we call the **time-evolution operator**, is unitary. Unitarity of $U(t)$ guarantees that the norm of $|\psi(t)\rangle$ is conserved: $\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|U^\dagger(t)U(t)|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle = 1$.

Extensive use of unitary matrices is made in quantum information and quantum computing. These matrices are implemented by manipulating the Hamiltonian by making use of (2.6). Efficient implementation of unitary matrices is an active area of quantum control theory.

It might be instructive to think about the scalar case and compare the situation with the one variable differential equation $x'(t) = (-ig(t))x(t)$ for a real-valued function $g(t)$, for which time-ordering is not necessary.

2.2 Multipartite System

So far, we have assumed implicitly that the system is made of a single component. Suppose a system is made of two components; one lives in a Hilbert space $\mathcal{H}_1$ and the other in another Hilbert space $\mathcal{H}_2$. A system composed of two separate components is called **bipartite**. Then the system as a whole lives in a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, whose general vector is written as

$$|\psi\rangle = \sum_{i,j} c_{ij}|e_{1,i}\rangle \otimes |e_{2,j}\rangle, \tag{2.7}$$

where $\{|e_{a,i}\rangle\}$ $(a = 1, 2)$ is an orthonormal basis in $\mathcal{H}_a$ and $\sum_{i,j}|c_{ij}|^2 = 1$.

A state $|\psi\rangle \in \mathcal{H}$ written as a tensor product of two vectors as $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, $(|\psi_a\rangle \in \mathcal{H}_a)$ is called a **separable state** or a **tensor product state**. A separable state admits a classical interpretation such as "The first system is in the state $|\psi_1\rangle$, while the second system is in $|\psi_2\rangle$, and viewed as a vector of the form $\begin{pmatrix}|\psi_1\rangle\\|\psi_2\rangle\end{pmatrix}$". It is clear such a set of vectors spans a linear space of dimension $\dim\mathcal{H}_1 + \dim\mathcal{H}_2$. Note however that the total space $\mathcal{H}$ has different dimensions since we find, by counting the number of coefficients in (2.7), that $\dim\mathcal{H} = \dim\mathcal{H}_1\dim\mathcal{H}_2$. This number is considerably larger than the dimension of the separable states when $\dim\mathcal{H}_a$ $(a = 1, 2)$ are large. What

are the missing states then? Suppose the spin states or a particle are denoted and represented by $|\uparrow\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|\downarrow\rangle = \begin{pmatrix}0\\1\end{pmatrix}$. Then the spin states for two particles are

$$|\uparrow\rangle\otimes|\uparrow\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ |\uparrow\rangle\otimes|\downarrow\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ |\downarrow\rangle\otimes|\uparrow\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ |\downarrow\rangle\otimes|\downarrow\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

Consider a spin state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle\otimes|\uparrow\rangle + |\downarrow\rangle\otimes|\downarrow\rangle\right) \tag{2.8}$$

of two separated electrons. Suppose $|\psi\rangle$ may be decomposed as

$$\begin{aligned}|\psi\rangle &= (c_1|\uparrow\rangle + c_2|\downarrow\rangle)\otimes(d_1|\uparrow\rangle + d_2|\downarrow\rangle)\\ &= c_1d_1|\uparrow\rangle\otimes|\uparrow\rangle + c_1d_2|\uparrow\rangle\otimes|\downarrow\rangle + c_2d_1|\downarrow\rangle\otimes|\uparrow\rangle + c_2d_2|\downarrow\rangle\otimes|\downarrow\rangle.\end{aligned}$$

However this decomposition is not possible since we must have

$$c_1d_2 = c_2d_1 = 0,\ c_1d_1 = c_2d_2 = \frac{1}{\sqrt{2}}$$

simultaneously, and it is clear that the above equations have no common solution. Therefore the state $|\psi\rangle$ is not separable.

Such non-separable states are called **entangled** in quantum theory [9]. The fact

$$\dim\mathcal{H}_1\dim\mathcal{H}_2 \gg \dim\mathcal{H}_1 + \dim\mathcal{H}_2$$

tells us that most states in a Hilbert space of a bipartite system are entangled when the constituent Hilbert spaces are higher dimensional. These entangled states refuse classical descriptions. Entanglement will be used extensively as a powerful computational resource in quantum information processing and quantum computation.

Suppose a bipartite state (2.7) is given. We are interested in when the state is separable and when entangled.

In case $\mathcal{H}_1 = \mathbb{C}^m$ and $\mathcal{H}_2 = \mathbb{C}^n$, we may choose $\mathcal{B}_1 = \{|e_{1,1}\rangle,\ldots,|e_{1,m}\rangle\}$ and $\mathcal{B}_2 = \{|e_{2,1}\rangle,\ldots,|e_{2,n}\rangle\}$ to be the standard bases for $\mathbb{C}^m$ and $\mathbb{C}^n$, i.e., $|e_{1,1}\rangle,\ldots,|e_{1,m}\rangle$ are the columns of I_m and $|e_{2,1}\rangle,\ldots,|e_{2,n}\rangle$ are the columns of I_n. Then

$$\mathcal{B} = \{|e_{1,r}\rangle\otimes|e_{2,s}\rangle : 1\le r\le m, 1\le s\le n\}$$

is the standard basis of $\mathbb{C}^{mn} = \mathbb{C}^m\otimes\mathbb{C}^n$. This can be seen by checking that $\mathcal{B}$ is an orthonormal set with mn vectors.

We can use the SVD to express $|\psi\rangle = \sum_{r,s} c_{rs}|e_{1,r}\rangle \otimes |e_{2,s}\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n$ in a simpler form and detect whether $|\psi\rangle$ is entangled or not as follows. For the SVD of

$$(c_{rs}) = U\Sigma V^\dagger = \sum_{j=1}^{k} s_j |u_j\rangle\langle v_j|,$$

we can set $|f_{1,j}\rangle = |u_j\rangle$ and $|f_{2,j}\rangle = |v_j\rangle^*$, the conjugate of $|v_j\rangle$. Then

$$|\psi\rangle = \sum_{r,s} c_{rs}|e_{1,r}\rangle \otimes |e_{2,s}\rangle = \sum_{j=1}^{k} s_j |f_{1,j}\rangle \otimes |f_{2,j}\rangle.$$

Clearly, $|\psi\rangle$ is a product state if and only if (c_{rs}) has rank one, i.e., $k = 1$ and $s_1 = 1$.

EXAMPLE 2.2.1. *Let $\{|e_{1,1}\rangle, |e_{1,2}\rangle, |e_{1,3}\rangle\}$ and $\{|e_{2,1}\rangle, |e_{2,2}\rangle\rangle\}$ be the standard bases of $\mathbb{C}^3$ and $\mathbb{C}^2$, and consider the bipartite state*

$$|\psi\rangle = \frac{1}{2}(|e_{1,1}\rangle|e_{2,1}\rangle + |e_{1,1}\rangle|e_{2,2}\rangle + i|e_{1,3}\rangle|e_{2,1}\rangle + i|e_{1,3}\rangle|e_{2,2}\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^3$$

whose coefficients form a matrix

$$C = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{pmatrix}.$$

It is immediate that the matrix is rank one, and $\sum_{i,j} |c_{ij}|^2 = 1$. So, $|\psi\rangle$ is a product state. We can find a simple representation of $|\psi\rangle$ using the SVD of C obtained in Example 1.6.2. We have

$$|\psi\rangle = |f_{1,1}\rangle|f_{2,1}\rangle,$$

where

$$|f_{1,1}\rangle = \sum_{i=1}^{3} U_{i1}|e_{1,i}\rangle = \frac{1}{\sqrt{2}}(|e_{1,1}\rangle + i|e_{1,3}\rangle)$$

and

$$|f_{2,1}\rangle = \sum_{j=1}^{2} V_{j1}^*|e_{2,j}\rangle = \frac{1}{\sqrt{2}}(|e_{2,1}\rangle + |e_{2,2}\rangle).$$

In general, we have the following.

PROPOSITION 2.2.2. *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ be the Hilbert space of a bipartite system. Then a vector $|\psi\rangle \in \mathcal{H}$ admits the* **Schmidt decomposition**

$$|\psi\rangle = \sum_{i=1}^{r} s_i |f_{1,i}\rangle \otimes |f_{2,i}\rangle \text{ with } \sum_i s_j^2 = 1, \tag{2.9}$$

where $s_i > 0$ are called the **Schmidt coefficients** *and $\{|f_{a,i}\rangle\}$ is an orthonormal set of $\mathcal{H}_a$. The number $r \in \mathbb{N}$ is called the* **Schmidt number** *of $|\psi\rangle$.*

Proof. This is a direct consequence of SVD introduced in §1.6. Let $|\psi\rangle$ be expanded as in Eq. (2.7). Note that the coefficients c_{ij} form a $\dim\mathcal{H}_1 \times \dim\mathcal{H}_2$ matrix C. We apply the SVD to obtain $C = U\Sigma V^\dagger$, where U and V are unitary matrices and Σ is a matrix whose (j,j) elements are nonnegative real numbers while all the other elements vanish. Now $|\psi\rangle$ of Eq. (2.7) is put in the form

$$|\psi\rangle = \sum_{i,j,k,\ell} U_{ik}\Sigma_{kl}V^*_{jl}|e_{1,i}\rangle \otimes |e_{2,j}\rangle.$$

Now define $|f_{1,k}\rangle = \sum_i U_{ik}|e_{1,i}\rangle$ and $|f_{2,k}\rangle = \sum_j V^*_{jk}|e_{2,j}\rangle$. Unitarity of U and V guarantees that they are orthonormal vectors of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. By noting that the (k,l) entry of Σ is $s_k\delta_{kl}$, we obtain

$$|\psi\rangle = \sum_{j=1}^{r} s_j|f_{1,j}\rangle \otimes |f_{2,j}\rangle,$$

where r is the number of positive (diagonal) elements in Σ. Since $\langle\psi|\psi\rangle = 1 = \sum_{r,s}|c_{rs}|^2$, which is the trace of $C^\dagger C$, and therefore equal to the sum of the eigenvalues of $C^\dagger C$, i.e., $\sum_j s_j^2$.

It follows from the above proposition that a bipartite state $|\psi\rangle$ is separable if and only if its Schmidt number r is 1.

Generalization to a system with more components, i.e., a **multipartite system**, should be obvious. A system composed of N components has a Hilbert space

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_N, \tag{2.10}$$

where $\mathcal{H}_a$ is the Hilbert space to which the ath component belongs. Classification of entanglement in a multipartite system is far from obvious, and an analogue of the Schmidt decomposition is not known to date for $N \geq 3$.*

2.3 No-Cloning Theorem

We copy classical data almost every day. In fact, this is amongst the most common functions with digital media. (Of course we should not copy media that are copyright protected.) However, the situation is different in quantum mechanics. The no-cloning theorem below asserts that it is impossible to create an independent and identical copy of an arbitrary unknown quantum

*See, however, [10, 11].

state. As shown in the proof, one cannot duplicate an unknown quantum system with a quantum operation, which is a unitary transformation by (A3).

THEOREM 2.3.1. *(Wootters and Zurek [19], Dieks [20]) An unknown quantum system cannot be cloned by unitary transformations.*

Proof. Suppose there would exist a unitary transformation U that makes a clone of a quantum system. Namely, suppose U acts, for any state $|\varphi\rangle$, as

$$U : |\varphi 0\rangle \to |\varphi\varphi\rangle. \tag{2.11}$$

Let $|\varphi\rangle$ and $|\phi\rangle$ be two states that are linearly independent. Then we should have $U|\varphi 0\rangle = |\varphi\varphi\rangle$ and $U|\phi 0\rangle = |\phi\phi\rangle$ by definition. Then the action of U on $|\psi\rangle = \frac{1}{\sqrt{2}}(|\varphi\rangle + |\phi\rangle)$ yields

$$U|\psi 0\rangle = \frac{1}{\sqrt{2}}(U|\varphi 0\rangle + U|\phi 0\rangle) = \frac{1}{\sqrt{2}}(|\varphi\varphi\rangle + |\phi\phi\rangle).$$

If U were a cloning transformation, we must also have

$$U|\psi 0\rangle = |\psi\psi\rangle = \frac{1}{2}(|\varphi\varphi\rangle + |\varphi\phi\rangle + |\phi\varphi\rangle + |\phi\phi\rangle),$$

which contradicts the previous result. Therefore, there does not exist a unitary cloning transformation. ∎

The no-cloning theorem may be proved by using the special theory of relativity, which assumes no information can propagate faster than the speed of light. [27]

Suppose Alice and Bob share a Bell state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle) = \frac{1}{\sqrt{2}}(|-\rangle|+\rangle - |+\rangle|-\rangle). \tag{2.12}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Readers are encouraged to verify the second equality. Alice keeps the first qubit while Bob keeps the second. If Alice wants to send Bob a bit "0", she measures her qubit in $\{|0\rangle, |1\rangle\}$ basis while if she wants to send "1", she employs $\{|+\rangle, |-\rangle\}$ basis for her measurement. Bob always measures his qubit in $\{|0\rangle, |1\rangle\}$ basis.

After Alice's measurment and before Bob's measurment, Bob's qubit is $|0\rangle$ or $|1\rangle$ if Alice sent "0" while it is $|+\rangle$ or $|-\rangle$ if Alice sent "1".

Suppose Bob is able to clone his qubit. He makes many copies of his qubit and measures them in $\{|0\rangle, |1\rangle\}$ basis. If Alice sent "0", Bob will obtain $0, 0, 0, \ldots$ or $1, 1, 1, \ldots$ while if she sent "1", Bob will obtain approximately 50% of 0's and 50% of 1's. Suppose Bob received $|\pm\rangle$ and made N clones, then the probability of obtaining the same outcome is $1/2^{N-1}$, which is negligible if N is sufficiently large. Note that Bob obtains the bit Alice wanted to send immediately after Alice's measurement assuming it does not take long to clone his qubit. This could happen even if Alice and Bob are separated many light years apart, thus in contradiction with the special theory of relativity. ∎

2.4 Qubits

A (Boolean) **bit** assumes two distinct values, 0 and 1. Bits constitute the building blocks of the classical information theory founded by C. Shannon. Quantum information theory, on the other hand, is based on **qubits**.

2.4.1 Qubit and Bloch Sphere

A qubit is a (unit) vector in the vector space $\mathbb{C}^2$, whose basis vectors are denoted as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.13}$$

What these vectors physically mean depends on the physical realization employed for quantum-information processing.

- In some cases, $|0\rangle$ stands for a horizontally polarized photon $|\leftrightarrow\rangle$, while $|1\rangle$ represents a vertically polarized photon $|\updownarrow\rangle$. Alternatively they might correspond to photons polarized in different directions. For example, $|0\rangle$ may represent a polarization state

  $$|\nearrow\rangle = \frac{1}{\sqrt{2}}(|\leftrightarrow\rangle + |\updownarrow\rangle),$$

  while $|1\rangle$ represents a state

  $$|\nwarrow\rangle = \frac{1}{\sqrt{2}}(|\leftrightarrow\rangle - |\updownarrow\rangle).$$

  Note that if $|\leftrightarrow\rangle$ ($|\updownarrow\rangle$) corresponds to an eigenstate of σ_z with the eigenvalue $+1$ (-1), respectively, then $|\nearrow\rangle$ ($|\nwarrow\rangle$) corresponds to an eigenstate of σ_x with the eigenvalue $+1$ (-1), respectively.

  Similarly, the states

  $$|\sigma^+\rangle = \frac{1}{\sqrt{2}}(|\leftrightarrow\rangle + i|\updownarrow\rangle), \quad |\sigma^-\rangle = \frac{1}{\sqrt{2}}(|\leftrightarrow\rangle - i|\updownarrow\rangle)$$

  correspond to the eigenstates of σ_y with the eigenvalues ± 1 and represent circularly polarized photons.

- They may represent spin states of an electron, $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$. Electrons are replaced by nuclei with spin-1/2 in NMR quantum computing.

- Truncated two states from many levels may also be employed as a qubit. Take the ground state and the first excited state of ionic energy levels or atomic energy levels, for example. We may assign $|0\rangle$ to the ground state and $|1\rangle$ to the first excited state.

In any case, we have to fix a set of basis vectors when we carry out quantum information processing. All the physics should be described with respect to this basis. In the following, the basis is written in an abstract form as $\{|0\rangle, |1\rangle\}$, unless otherwise stated.

Note that the third example of a qubit above suggests that a quantum system with more than two states may be employed for information storage and information processing. If a quantum system admits three different states, it is called a **qutrit**, while if it takes d different states, it is called a **qudit**. A spin S particle, for example, takes $d = 2S + 1$ spin states and works as a qudit. The significance of qutrits and qudits in information processing is still to be explored.

Qubits can be used as a **Hardware random number generator** or **true random number generator** to generate cryptographic keys. (See [28] for background and references for random number generation.) Based on the quantum postulates, one can generate a zero-one sequence by measuring a sequence of quibts in a state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ one by one.

It is convenient to assume the vector $|0\rangle$ corresponds to the classical value 0, while $|1\rangle$ to 1 in quantum computation. Moreover it is possible for a qubit to be in a superposition state:

$$|\psi\rangle = a|0\rangle + b|1\rangle \text{ with } \ a, b \in \mathbb{C}, \ |a|^2 + |b|^2 = 1. \tag{2.14}$$

The fundamental requirement of quantum mechanics is that if we make measurement on $|\psi\rangle$ to see whether it is in $|0\rangle$ or $|1\rangle$, the outcome will be 0 (1) with the probability $|a|^2$ ($|b|^2$), and the state immediately after the measurement is $|0\rangle$ ($|1\rangle$).

Although a qubit may take infinitely many different states, it should be kept in mind that we can extract from it as the same amount of information as that of a classical bit. Information can be extracted only through measurements. When we make measurement on a qubit, the state vector "collapses" to the eigenvector that corresponds to the eigenvalue observed. Suppose that a spin is in a state $a|0\rangle + b|1\rangle$. If we observe that the z-component of the spin is $+1/2$, the system immediately after the measurement is *definitely* in the state $|0\rangle$. This happens with probability $\langle\psi|0\rangle\langle 0|\psi\rangle = |a|^2$. The outcome of a measurement on a qubit is always one of the eigenvalues, which we call abstractly 0 and 1, just like for a classical bit. As a result, if we can make measurements of a large number of copies of unknown state, we may estimate $|a|$ and $|b|$, but we cannot estimate the coefficients a and b.

It is useful, for many purposes, to express a state of a single qubit graphically. Note that every qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ corresponds to the rank one orthogonal projection

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{pmatrix} = \frac{1}{2}(I_2 + n_x\sigma_x + n_y\sigma_y + n_z\sigma_z),$$

where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices, $(n_x - in_y)/2 = a\bar{b}, n_z = 2|a|^2 - 1$. As we will see in the next chapter, the matrix ρ is a density matrix corre-

sponding to a pure state. We use the notation in Proposition 1.5.8, and let $\hat{\boldsymbol{n}} = (n_x, n_y, n_z) \in \mathbb{R}^3$ be a unit vector, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and

$$A = \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma} = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}.$$

Then

$$A = P_1 - P_2 \quad \text{with} \quad P_1 = (I + A)/2 \quad \text{and} \quad P_2 = (I - A)/2,$$

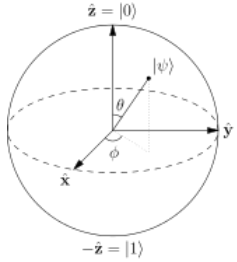
where $|\psi\rangle\langle\psi| = P_1$ is precisely the eigenprojection of the eigenvalue 1 of $|\psi\rangle\langle\psi|$. In particular, we may parametrize a one-qubit state $|\psi\rangle$ with θ and ϕ as

$$|\psi\rangle = |\psi(\theta, \phi)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle, \quad \theta \in [0, \pi],\ \phi \in [0, 2\pi), \tag{2.15}$$

where we take advantage of the freedom of the phase to make the coefficient of $|0\rangle$ real. Then

$$|\psi\rangle\langle\psi| = \frac{1}{2}(I + \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma})$$

with $\hat{\boldsymbol{n}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ corresponding to the unique point $\hat{\boldsymbol{n}}$ on the unit sphere in $\mathbb{R}^3$ using the polar co-ordinates. So, there is a natural correspondence between a unit vector $\hat{\boldsymbol{n}}(\theta, \phi)$ and a state vector $|\psi(\theta, \phi)\rangle$. Namely, a state $|\psi(\theta, \phi)\rangle$ is expressed as a unit vector $\hat{\boldsymbol{n}}(\theta, \phi)$ on the surface of the unit sphere, called the **Bloch sphere**.



The Bloch sphere

A more general density matrix ρ can be constructed as $\rho = \sum_{j=1}^{r} p_r |\psi_j\rangle\langle\psi_j|$, where $|\psi_1\rangle\langle\psi_1|, \ldots, |\psi_r\rangle\langle\psi_r|$ are rank one orthogonal projections and $p_1, \ldots, p_r$ are positive numbers summing up to one. Then ρ has the form

$$\rho = \frac{1}{2}\left(I + \sum_{j=x,y,z} u_j \sigma_i\right), \tag{2.16}$$

where u_i are components of a real vector $\boldsymbol{u}$ satisfying $|\boldsymbol{u}| \leq 1$. The Hermitian matrix $u_x\sigma_x + u_y\sigma_y + u_z\sigma_z$ has trace 0 and determinant $-|\boldsymbol{u}|^2 = -(u_x^2 + u_y^2 + u_z^2)$. So, it has eigenvalues $\pm|\boldsymbol{u}|$, and hence the eigenvalues of ρ are

$$\lambda_+ = \frac{1}{2}(1 + |\boldsymbol{u}|),\ \lambda_- = \frac{1}{2}(1 - |\boldsymbol{u}|) \tag{2.17}$$

and therefore non-negative. In case $|\boldsymbol{u}| = 1$, the eigenvalue λ_- vanishes and rank $\rho = 1$. Therefore the surface of the unit ball corresponds to pure states. The converse is also shown easily. The ball is called the **Bloch ball**, and its boundary called the Bloch sphere. The vector $\boldsymbol{u}$ is also called the Bloch vector. The normalized vector $\hat{\boldsymbol{n}}$ of the Bloch sphere is a special case of $\boldsymbol{u}$ restricted in pure states.

2.4.2 Multi-Qubit Systems and Entangled States

Let us consider a group of many (n) qubits next. Such a system behaves quite differently from a classical one, and this difference gives a distinguishing aspect to quantum information theory. An n-qubit system is often called a (quantum) **register** in the context of quantum computing.

Consider a classical system made of several components. The state of this system is completely determined by specifying the state of each component. This is *not* the case for a quantum system. A quantum system made of many components is not necessarily described by specifying the state of each component as we have learned in §2.2.

As an example, let us consider an n-qubit register. Suppose we specify the state of each qubit separately in analogy with a classical case. Each of the qubits is then described by a two-dimensional complex vector of the form $a_i|0\rangle + b_i|1\rangle$, and we need $2n$ complex numbers $\{a_i, b_i\}_{1\leq i\leq n}$ to specify the state. This corresponds the the tensor product state

$$(a_1|0\rangle + b_1|1\rangle) \otimes (a_2|0\rangle + b_2|1\rangle) \otimes \ldots \otimes (a_n|0\rangle + b_n|1\rangle)$$

introduced in §2.2. If the system is treated in a fully quantum-mechanical way, however, we have to include superposition of such tensor product states, which is not necessarily decomposable into a tensor product form. Such a state is **entangled** (see §2.2). A general state vector of the register is represented as

$$|\psi\rangle = \sum_{i_k=0,1} c_{i_1 i_2 \ldots i_n} |i_1\rangle \otimes |i_2\rangle \otimes \ldots \otimes |i_n\rangle$$

and lives in a 2^n-dimensional complex vector space. Note that $2^n \gg 2n$ for a large number n. The ratio $2^n/2n$ is $\sim 6.3 \times 10^{27}$ for $n = 100$ and $\sim 5.4 \times 10^{297}$ for $n = 1000$. These astronomical numbers tell us that most quantum states in a Hilbert space with large n are entangled, i.e., they do not have classical analogy which tensor product states have. Entangled states that have no classical counterparts are extremely powerful resources for quantum computation and quantum communication as we will show later.

Let us consider a system of two qubits for definiteness. The combined system has a basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. More generally, a basis for a system of n qubits may be taken to be $\{|b_{n-1}b_{n-2}\ldots b_0\rangle\}$, where $b_{n-1}, b_{n-2}, \ldots, b_0 \in \{0, 1\}$. It is also possible to express the basis in terms of the decimal system. We write $|x\rangle$, instead of $|b_{n-1}b_{n-2}\ldots b_0\rangle$, where $x = b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \ldots + b_0$ is the decimal expression of the binary number $b_{n-1}b_{n-2}\ldots b_0$. Thus the basis for a two-qubit system may be written also as $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ with this decimal notation. Whether the binary system or the decimal system is employed should be clear from the context. An n-qubit system has $2^n = \exp(n \ln 2)$ basis vectors.

The set

$$\begin{aligned}\{|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\\ |\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\}\end{aligned} \tag{2.18}$$

is an orthonormal basis of a two-qubit system and is called the **Bell basis**. Each vector is called the **Bell state** or the **Bell vector**. Note that all the Bell states are entangled.

Among three-qubit entangled states, the following two states are important for various reasons and hence deserve special names. The state

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{2.19}$$

is called the **Greenberger-Horne-Zeilinger state** and is often abbreviated as the **GHZ state** [21]. Another important three-qubit state is the **W state** [22],

$$|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle). \tag{2.20}$$

The GHZ state has a different entanglement pattern as the W state. If any qubit in the GHZ state is measured, the resulting state is a separable state while if any qubit in the W state is measured, there is a probability of $2/3$ that the resulting state is entangled. Accordingly. it is impossible to transform the GHZ state to the W state by local operations and *vice versa*.

2.5 Measurement

Classical information theory is formulated independently of measurements of the system under consideration. This is because the readout of the result is always the same for anyone and at any time, provided that the system processes the same information without error. This is completely different in quantum information processing. Measurement is an essential part of the theory as we see below.

By making a measurement on a system, we *project* the state vector to one of the basis vectors that the measurement equipment defines.[†] Suppose we have a state vector $|\psi\rangle = a|0\rangle + b|1\rangle$ and measure it to see if it is in the state $|0\rangle$ or $|1\rangle$. Depending on the system, this means if a spin points up or down or a photon is polarized horizontally or vertically, for example. The result is either

[†]This is called a projective measurement as was noted in Section 2.1.

0 or 1. In the first case, the state "collapses" to $|0\rangle$ while in the second case, to $|1\rangle$. We find, after many measurements, the probability of obtaining outcome 0 (1) is $|a|^2$ ($|b|^2$). Here, we may assume that the observable is associate with the Hermitian operator $M = \lambda_0|0\rangle\langle 0| + \lambda_1|1\rangle\langle 1|$. Note that the eigenvalues λ_0 and λ_1 represent the readings produced by the apparatus if the states $|0\rangle$ and $|1\rangle$ are measured, respectively. They may be determined or limited by the set up of the measuring device.

To be more formal, we construct a Hermitian matrix corresponding to the observable $M = \sum_m \lambda_m P_m$ such that the eigenprojection P_m will serve as the **measurement operator** P_m such that the probability of obtaining the outcome λ_m in the state $|\psi\rangle$ is

$$p(m) = \langle\psi|P_m|\psi\rangle, \tag{2.21}$$

and the state immediately after the measurement is

$$|m\rangle = \frac{P_m|\psi\rangle}{\sqrt{p(m)}}. \tag{2.22}$$

In the above example, the measurement operators are nothing but projection operators; $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$. In fact, we have

$$p(0) = \langle\psi|P_0|\psi\rangle = \langle\psi|0\rangle\langle 0|\psi\rangle = |a|^2,$$

and

$$\frac{P_0|\psi\rangle}{\sqrt{p(0)}} = \frac{a}{|a|}|0\rangle \simeq |0\rangle,$$

and similarly for the other case P_1. It should be noted that a quantum state is defined up to a phase and hence $a/|a|$ does not play any role. [‡]

Suppose we are given many copies of a particular state $|\psi\rangle$. If we measure an observable $M = \sum_m \lambda_m P_m$ in each of the copies, the expectation value of M is given, in terms of the projection operators, by

$$\begin{aligned} E(M) &= \sum_m \lambda_m\, p(m) = \sum_m \lambda_m \langle\psi|P_m|\psi\rangle \\ &= \langle\psi|\sum_m \lambda_m P_m|\psi\rangle = \langle\psi|M|\psi\rangle, \end{aligned} \tag{2.23}$$

where use has been made of the spectral decomposition $M = \sum_m \lambda_m P_m$. The standard deviation of the measurement outcomes of M is given by

$$\Delta(M) = \sqrt{\langle(M - \langle M\rangle)^2\rangle} = \sqrt{\langle M^2\rangle - \langle M\rangle^2}. \tag{2.24}$$

[‡]We will consider some more general measurement operators. One may see [18] if desired.

Let us analyze measurements in a two-qubit system in some detail. An arbitrary state is written as

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle, \quad |a|^2 + |b|^2 + |c|^2 + |d|^2 = 1,$$

where $a, b, c, d \in \mathbb{C}$. We make a measurement of the first qubit with respect to the basis $\{|0\rangle, |1\rangle\}$. To this end, we rewrite the state as

$$\begin{aligned} &a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \\ &= |0\rangle \otimes (a|0\rangle + b|1\rangle) + |1\rangle \otimes (c|0\rangle + d|1\rangle) \\ &= u|0\rangle \otimes \left(\frac{a}{u}|0\rangle + \frac{b}{u}|1\rangle\right) + v|1\rangle \otimes \left(\frac{c}{v}|0\rangle + \frac{d}{v}|1\rangle\right), \end{aligned}$$

where $u = \sqrt{|a|^2 + |b|^2}$ and $v = \sqrt{|c|^2 + |d|^2}$. The measurement operators acting on the first qubit are

$$M_0 = |0\rangle\langle 0| \otimes I, \quad M_1 = |1\rangle\langle 1| \otimes I. \tag{2.25}$$

Note that we need to specify $\otimes I$ explicitly since we are working in a two-qubit Hilbert space $\mathbb{C}^4$. Here we may assume that the Hermitian operator corresponding to the observable is $M = M_0 - M_1$, say. Upon a measurement of the first qubit, we obtain 0 with the probability

$$\langle\psi|M_0|\psi\rangle = u^2 = |a|^2 + |b|^2,$$

projecting the state to

$$\frac{M_0|\psi\rangle}{\sqrt{p(0)}} = |0\rangle \otimes \left(\frac{a}{u}|0\rangle + \frac{b}{u}|1\rangle\right),$$

while we obtain $|1\rangle$ with the probability $v^2 = |c|^2 + |d|^2$, projecting the state to $|1\rangle \otimes \left(\frac{c}{v}|0\rangle + \frac{d}{v}|1\rangle\right)$. Note that the state after the measurement has unit norm in both cases. The measurement of the second qubit can be carried out similarly. Measurements on an n-qubit system can be carried out by repeating one-qubit measurement n times.

In the two-qubit example above, the Hilbert space for the system is separated into a direct sum of $\mathcal{H}_0$, where the first qubit is in the state $|0\rangle$, and $\mathcal{H}_1$, where it is in $|1\rangle$: $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$. An arbitrary two-qubit state $|\psi\rangle$ is uniquely decomposed into two vectors, each of which belongs to $\mathcal{H}_0$ or $\mathcal{H}_1$ as

$$(|0\rangle\langle 0| \otimes I)|\psi\rangle \in \mathcal{H}_0, \quad (|1\rangle\langle 1| \otimes I)|\psi\rangle \in \mathcal{H}_1,$$

where normalization has been ignored. More generally, a measurement of k qubits in an n-qubit system yields 2^k possible outcomes m_i $(1 \le i \le 2^k)$. Accordingly, the 2^n-dimensional Hilbert space of the system is separated into the direct sum of mutually orthogonal subspaces $\mathcal{H}_{m_1}, \mathcal{H}_{m_2}, \ldots, \mathcal{H}_{m_{2^k}}$ as $\mathcal{H} =$

$\mathcal{H}_{m_1} \oplus \mathcal{H}_{m_2} \oplus \ldots \oplus \mathcal{H}_{m_{2^k}}$. When the result of the measurement of the k qubits is m_i, the state after the measurement is projected to the subspace $\mathcal{H}_{m_i}$. It should be clear from the construction that each subspace $\mathcal{H}_{m_i}$ has dimension $2^n/2^k = 2^{n-k}$. The measurement device projects the state before the measurement

$$|\psi\rangle = c_{m_1}|\psi_{m_1}\rangle + c_{m_2}|\psi_{m_2}\rangle + \ldots + c_{m_{2^k}}|\psi_{m_{2^k}}\rangle, \quad (|\psi_{m_i}\rangle \in \mathcal{H}_{m_i})$$

into one of the subspaces $\mathcal{H}_{m_i}$ randomly with the probability $|c_{m_i}|^2$.

Measurement gives an alternative viewpoint to entangled states. A state is not entangled if a measurement of a qubit does not affect the state of the other qubits. Suppose the first qubit of the state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

was measured to be 0 (1). Then the outcome of the measurement of the second qubit is *definitely* 0 (1). Therefore the measurement of the first qubit affects the outcome of the measurement on the second qubit, which shows that the initial state is an entangled state. In other words, there exists a strong correlation between the two qubits. This correlation may be used for information processing as will be shown later. In contrast with this, the state $\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$ is not entangled since it can be written as

$$\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

Regardless of the measurement of the second qubit, the measurement of the first qubit definitely yields 0. Moreover, the second qubit is measured to be 0 (1) with the probability 1/2, independently of whether the first qubit is measured or not.

2.5.1 Einstein-Podolsky-Rosen (EPR) Phenomenon

Einstein, Podolsky and Rosen (EPR) proposed a Gedanken experiment which, at first glance, shows that an entangled state violates an axiom of the special theory of relativity [23]. Suppose a particle source produces the so-called EPR pair in the state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

and it sends the first qubit to Alice and the second to Bob, who may be separated far away (see Fig. 2.1).[§] Alice measures her qubit and obtains her reading $|0\rangle$ or $|1\rangle$. Depending on her reading, the EPR state is projected to

[§]Alice and Bob are names frequently used in information theory.

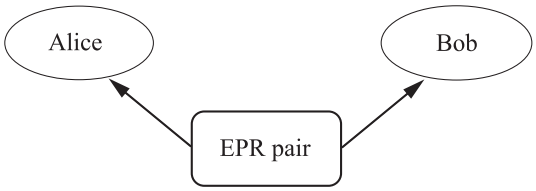


FIGURE 2.1
EPR pair produced by a source in the middle. One qubit is sent to Alice and the other to Bob.

$|01\rangle$ ($|10\rangle$), and Bob will *definitely* observe $|1\rangle$ ($|0\rangle$) in his measurement. The change of the state

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \;\rightarrow\; |01\rangle \quad \text{or} \quad |10\rangle \tag{2.26}$$

takes place instantaneously even when they are separated by a large distance. It seems that Alice's measurement propageted to Bob's qubit instantaneously, and it violates the special theory of relativity. This is the very point EPR proposed to defeat quantum mechanics, which claims that nothing can propagate faster than the speed of light.

Note, however, that nothing has propagated from Alice to Bob and *vice versa*, upon Alice's measurement. Clearly no energy has propagated. What about information? It is impossible for Alice to control her and hence Bob's readings. Therefore it is impossible to use EPR pairs to send a sensible message from Alice to Bob. If they could, the message would be sent instantaneously, which certainly violates the special theory of relativity. If a large number of EPR pairs are sent to Alice and Bob and they independently measure their qubits, they will observe random sequences of 0 and 1. They notice that their readings are strongly correlated only after they exchange their sequences by means of classical communication, which can be done at most with the speed of light.

2.5.2 Bell inequality

About 30 years after the EPR paper was published, an experiment test was proposed to check whether the measurement of entangled pairs follow a certain predetermined rule imposed by Nature, or the postulate of quantum mechanics.

Here is the proposed experiments. Suppose Charlie prepares an entangled pair of qubits (photons or particles) and sends the first one to Alice and the second one to Bob. Alice will apply one of her two measurement schemes, say, Q and R, each will produce a measured value in $\{1, -1\}$. Bob will also

apply one of his two measurement schemes, say, S and T, each will produce a measured value in $\{1, -1\}$.

Let us consider

$$QS + RS + RT - QT = (Q + R)S + (R - Q)T.$$

Because $R, Q \in \{1, -1\}$, it follows that either $(Q+R)S = 0$ or $(R-Q)T = 0$. As a result, $QS + RS + RT - QT \in \{2, -2\}$.

Suppose there is a hidden rule governing the measurement outcomes, and $p(q, r, s, t)$ is the probability that, *before* the measurements are performed, the system is in the state $(Q, R, S, T) = (q, r, s, t)$. Then the expectation value $E(QS + RS + RT - QT) = E(QS) + E(RS) + E(RT) - E(QT)$ satisfies

$$\begin{aligned} |E(QS + RS + RT - QT)| &= \sum_{(q,r,s,t)} p(q, r, s, t)|qs + rs + rt - qt| \\ &\leq \sum_{(q,r,s,t)} p(q, r, s, t) \cdot 2 = 2. \end{aligned}$$

So, we get the **Bell inequality**

$$|E(QS) + E(RS) + E(RT) - E(QT)| \leq 2. \tag{2.27}$$

Suppose Charlie prepares an entangled state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

and gives Alice the first qubit, and Bob the second one. Alice uses the measurement operators $Q = \sigma_z$ and $R = \sigma_x$, and Bob uses the measurement operators $S = \frac{-1}{\sqrt{2}}(\sigma_z + \sigma_x)$ and $T = \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x)$. Then

$$E(QS) = \langle\Psi^-|Q \otimes S|\Psi^-\rangle = \frac{1}{\sqrt{2}}, \quad E(RS) = \langle\Psi^-|R \otimes S|\Psi^-\rangle = \frac{1}{\sqrt{2}},$$

$$E(RT) = \langle\Psi^-|R \otimes T|\Psi^-\rangle = \frac{1}{\sqrt{2}}, \quad E(QT) = \langle\Psi^-|Q \otimes T|\Psi^-\rangle = \frac{-1}{\sqrt{2}},$$

and hence

$$E(QS + RS + RT - QT) = 4/\sqrt{2} = 2\sqrt{2}. \tag{2.28}$$

This equality clearly violates the Bell inequality.

To determine whether (2.27) or (2.28) is valid, Alice and Bob can estimate $E(QS)$ by performing measurements on many copies of $|\Psi^-\rangle$, and record their results. After the experiments, they can multiply their measurements when they used the measurement schemes Q and S, respectively. Similarly, they can estimate $E(RS), E(RT), E(QT)$, so as to obtain an estimate of $E(QS + RS + RT - QT)$.

Experimental results showed strong support to (2.28). Hence, the EPR proposal that there is a hidden rule governing the measurement results of entangled pair was ruled out.

Note that measurement of a qubit in $|\Psi^-\rangle$ disentangles the pair. As a result, a tensor product state $|0\rangle|1\rangle$ or $|1\rangle|0\rangle$ satisfies the Bell inequality. This fact will be used to detect eavesdroppers in quantum key distribution protocols.

2.6 Additional Topics

2.6.1 The uncertainty principle

Historically, a deep quantum mechanical property that differs from classical mechanics is the **Uncertainty Principle**. We can use matrix theory to explain the uncertainly principle as follows. Recall that the expectation value of an observable associated with a Hermitian matrix A with respect to a quantum state $|\psi\rangle$ is $\langle A\rangle = \langle\psi|A|\psi\rangle$. Let A and B be Hermitian operators and $|\psi\rangle$ be some quantum state on which A and B operate. As mentioned before, if A corresponds to an observable, then the standard deviation of the observable is

$$\Delta(A) = \sqrt{\langle\psi|(A-\alpha I)^2|\psi\rangle} = \sqrt{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2},$$

where $\alpha = \langle\psi|A|\psi\rangle$. We have the following.

THEOREM 2.6.1. *Let $A, B \in \mathbf{M}_n$ be Hermitian, $|\psi\rangle \in \mathbb{C}^n$ be a unit vector. Then*

$$\Delta(A)\Delta(B) \geq |\langle\psi|(AB-BA)|\psi\rangle|/2. \tag{2.29}$$

Suppose $\alpha = \langle\psi|A|\psi\rangle$ and $\beta = \langle\psi|B|\psi\rangle$. The equality holds in (2.29) *if and only if there is $\theta \in \mathbb{R}$ such that*

$$\cos\theta(A-\alpha I)|\psi\rangle + i\sin\theta(B-\beta I)|\psi\rangle = |\mathbf{0}\rangle.$$

Proof. Let $\hat{A} = A - \alpha I$ and $\hat{B} = B - \beta I$. Note first that $\Delta(A)\Delta(B) = \sqrt{\langle\psi|\hat{A}^2|\psi\rangle}\sqrt{\langle\psi|\hat{B}^2|\psi\rangle}$ and $\langle\psi|[A,B]|\psi\rangle = \langle\psi|[\hat{A},\hat{B}]|\psi\rangle$. So, we only need to show that $4\langle\psi|\hat{A}^2|\psi\rangle\langle\psi|\hat{B}^2|\psi\rangle \geq |\langle\psi|[\hat{A},\hat{B}]|\psi\rangle|^2$. Note that the matrices

$$C_1 = \begin{pmatrix} \langle\psi|\hat{A}^2|\psi\rangle & \langle\psi|\hat{A}\hat{B}|\psi\rangle \\ \langle\psi|\hat{B}\hat{A}|\psi\rangle & \langle\psi|\hat{B}^2|\psi\rangle \end{pmatrix} \quad \text{and} \quad C_2 = \begin{pmatrix} \langle\psi|\hat{A}^2|\psi\rangle & -\langle\psi|\hat{B}\hat{A}|\psi\rangle \\ -\langle\psi|\hat{A}\hat{B}|\psi\rangle & \langle\psi|\hat{B}^2|\psi\rangle \end{pmatrix}$$

are positive semi-definite as proved by checking that all their principal minors are nonnegative using the Cauchy-Schwartz inequality. Thus, $C = C_1 + C_2$ is positive semi-definite and

$$4\langle\psi|\hat{A}^2|\psi\rangle\langle\psi|\hat{B}^2|\psi\rangle - |\langle\psi|[\hat{A},\hat{B}]|\psi\rangle|^2 = \det(C) \geq 0.$$

The equality $\det(C) = 0$ holds if and only if C is singular, equivalently, the positive semi-definite matrices C_1 and C_2 are singular and share a common null vector. Since C_1 and C_2 have the same trace, we see that

(1) $C_1 = C_2 = (\text{tr } C_1)|u\rangle\langle u|$ for some unit vector $|u\rangle \in \mathbb{C}^n$, and

(2) $\langle\psi|\hat{A}\hat{B}|\psi\rangle = -\langle\psi|\hat{B}\hat{A}|\psi\rangle$, i.e., $\langle\psi|\{\hat{A}, \hat{B}\}|\psi\rangle = 0$.

Condition (1) implies $\det(C_1) = 0$, namely $\langle\psi|\hat{A}^2|\psi\rangle\langle\psi|\hat{B}^2|\psi\rangle = |\langle\psi|\hat{A}\hat{B}|\psi\rangle|^2$. By the Cauchy-Schwartz inequality, $\hat{A}|\psi\rangle$ and $\hat{B}|\psi\rangle$ are linearly dependent. Condition (2) implies that $\langle\psi|\hat{A}\hat{B}|\psi\rangle \in i\mathbb{R}$. So, $\hat{A}|\psi\rangle$ and $i\hat{B}|\psi\rangle$ are linearly dependent over $\mathbb{R}$. Thus, there is $\theta \in [0, 2\pi)$ such that $\cos\theta\hat{A}|\psi\rangle + i\sin\theta\hat{B}|\psi\rangle$ is the zero vector. Conversely, if $\cos\theta\hat{A}|\psi\rangle + i\sin\theta\hat{B}|\psi\rangle$ is the zero vector, one readily checks that $C_1 = C_2$ and $\det(C_1 + C_2) = 0$. ∎

For example, if $A = \sigma_x$, $B = \sigma_y$, and $|\psi\rangle = (1, 0)^t$, we have the equality.

Now, suppose $A = Q$ and $B = P \equiv \frac{\hbar}{i}\frac{d}{dQ}$; see [26, Example 2.20]. Deduce from the above arguments that

$$\Delta(Q)\Delta(P) \geq \frac{\hbar}{2}.$$

Note that $AB - BA = \gamma I$ with $\gamma \neq 0$ can only happen for infinite dimensional operators as $\text{Tr}\,(AB - BA) = 0$ for finite dimensional operators. The Uncertainty principle in terms of standard deviation has been formulated first in [7] and [8].

2.6.2 Schrödinger Picture and Heisenberg Picture

In the previous section, we have introduced the Schrödinger equation

$$i\hbar\frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle$$

where $|\psi(t)\rangle = U(t)|\psi(0)\rangle$. The expectation value of an observable a at t is given by $\langle\psi(t)|A|\psi(t)\rangle$. This description with time-dependent state $|\psi(t)\rangle$ and time-independent operator A is called the **Schrödinger picture**.

There is another equivalent description in which the state is time-independent while the operator depends on time, called the **Heisenberg picture**. Let

$$A(t) = U(t)^\dagger A U(t) \tag{2.30}$$

and consider $\langle\psi(0)|A(t)|\psi(0)\rangle$. We find

$$\langle\psi(0)|A(t)|\psi(0)\rangle = \langle\psi(0)|U(t)^\dagger A U(t)|\psi(0)\rangle = \langle\psi(t)|A|\psi(t)\rangle,$$

which shows the two pictures give the same expectation value at any t. It is also shown that any matrix element satisfies

$$\langle\phi(t)|A|\psi(t)\rangle = \langle\phi(0)|A(t)|\psi(0)\rangle$$

and hence two pictures are equivalent.

Now the dynamics of the system in the Heisenberg picture is carried by operators. We find

$$\frac{d}{dt}A(t) = \frac{1}{\hbar}(iHA(t) - A(t)iH) = \frac{i}{\hbar}[H, A(t)]. \tag{2.31}$$

If, moreover, A has explicit time-dependence, we obtain

$$\frac{d}{dt}A(t) = \left(\frac{\partial A}{\partial t}\right)(t) + \frac{i}{\hbar}[H, A(t)], \tag{2.32}$$

where $(\partial A/\partial t)(t) = e^{iHt/\hbar}(\partial A/\partial t)e^{-iHt/\hbar}$.

In the Schrödinger picture, the state evolves with time. In contrast, observables, such as position and momentum, evolve with time in the Heisenberg picture. In this sense, the Heisenberg picture depicts situation similar to the classical dynamics. Let H be a Hamiltonian of a classical system and $a(t)$ be any physical quantity such as the coordinates or the momentum. Then the time evolution of $a(t)$ is described by the Hamilton's equation of motion

$$\frac{da(t)}{dt} = \frac{\partial a(t)}{\partial t} + [a(t), H]_{\rm PB}, \tag{2.33}$$

where $[x, y]_{\rm PB}$ is the Poisson bracket defined by

$$[x, y]_{\rm PB} = \sum_k \left(\frac{\partial x}{\partial q_k}\frac{\partial y}{\partial p_k} - \frac{\partial y}{\partial q_k}\frac{\partial x}{\partial p_k}\right), \tag{2.34}$$

where $\{q_k, p_k\}$ is the set of coordinates and corresponding momenta. Notice the similarity between Eqs. (2.33) and (2.32).

2.6.3 Some Examples

We now give some examples to clarify the axioms introduced §2.1. They turn out to have relevance to certain physical realizations of a quantum computer.

EXAMPLE 2.6.2. *Let us consider a time-independent Hamiltonian*

$$H = -\frac{\hbar}{2}\omega\sigma_x. \tag{2.35}$$

Suppose the system is in the eigenstate of σ_z with the eigenvalue $+1$ at time $t = 0$;

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

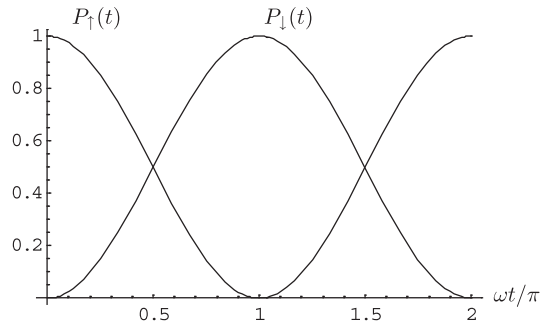


FIGURE 2.2
Probability $P_\uparrow(t)$ with which a spin is observed in the $\uparrow$-state and $P_\downarrow(t)$ observed in the $\downarrow$-state.

The wave function $|\psi(t)\rangle$ $(t > 0)$ *is then found from Eq. (2.5) to be*

$$|\psi(t)\rangle = \exp\left(i\frac{\omega}{2}\sigma_x t\right)|\psi(0)\rangle. \tag{2.36}$$

The matrix exponential function in this equation is evaluated with the help of Eq. (1.23) and we find

$$|\psi(t)\rangle = \begin{pmatrix} \cos\omega t/2 & i\sin\omega t/2 \\ i\sin\omega t/2 & \cos\omega t/2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\omega t/2 \\ i\sin\omega t/2 \end{pmatrix}. \tag{2.37}$$

Suppose we measure the observable σ_z*. Note that* $|\psi(t)\rangle$ *is expanded in terms of the eigenvectors of* σ_z *as*

$$|\psi(t)\rangle = \cos\frac{\omega}{2}t|\sigma_z = +1\rangle + i\sin\frac{\omega}{2}t|\sigma_z = -1\rangle.$$

Therefore we find the spin is in the spin-up state with the probability $P_\uparrow(t) = \cos^2(\omega t/2)$ *and in the spin-down state with the probability* $P_\downarrow(t) = \sin^2(\omega t/2)$ *as depicted in Fig. 2.2. Of course, the total probability is independent of time since* $\cos^2(\omega t/2) + \sin^2(\omega t/2) = 1$*. This result is consistent with classical spin dynamics. The Hamiltonian (2.35) depicts a spin under a magnetic field along the x-axis. Our initial condition signifies that the spin points the z-direction at* $t = 0$*. Then the spin starts precession around the x-axis, and the z-component of the spin oscillates sinusoidally as is shown above.*

Next let us take the initial state

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which is an eigenvector of σ_x (and hence of the Hamiltonian) with the eigenvalue $+1$. We find $|\psi(t)\rangle$ in this case as

$$|\psi(t)\rangle = \begin{pmatrix} \cos\omega t/2 & i\sin\omega t/2 \\ i\sin\omega t/2 & \cos\omega t/2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{e^{i\omega t/2}}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{2.38}$$

Therefore the state remains in its initial state at an arbitrary $t > 0$. This is an expected result since the system at $t = 0$ is an eigenstate of the Hamiltonian.

Now let us formulate Example 2.6.2 (and also Exercise 2.1) in the most general form. Consider a Hamiltonian

$$H = -\frac{\hbar}{2}\omega \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}, \tag{2.39}$$

where $\hat{\boldsymbol{n}}$ is a unit vector in $\mathbb{R}^3$. The time-evolution operator is readily obtained, by making use of the result of Proposition 1.5.8, as

$$U(t) = \exp(-iHt/\hbar) = \cos\frac{\omega}{2}t\ I + i(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma})\sin\frac{\omega}{2}t. \tag{2.40}$$

Suppose the initial state is

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

for example. Then we find

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = \begin{pmatrix} \cos(\omega t/2) + in_z \sin(\omega t/2) \\ i(n_x + in_y)\sin(\omega t/2) \end{pmatrix}. \tag{2.41}$$

The reader should verify that $|\psi(t)\rangle$ is normalized at any instant of time $t > 0$.

EXAMPLE 2.6.3. (Rabi oscillation) *This example is often employed for a quantum gate implementation. We will take the natural unit $\hbar = 1$ to simplify our notation throughout this example. Let us consider a spin-1/2 particle in a magnetic field along the z-axis, whose Hamiltonian is given by*

$$H_0 = -\frac{\omega_0}{2}\sigma_z. \tag{2.42}$$

Suppose the particle is irradiated by an oscillating magnetic field of angular frequency ω, which introduces transitions between two energy eigenstates of H_0. Then the perturbed Hamiltonian is modeled as

$$H = -\frac{\omega_0}{2}\sigma_z + \frac{\omega_1}{2}\begin{pmatrix} 0 & e^{i\omega t} \\ e^{-i\omega t} & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\omega_0 & \omega_1 e^{i\omega t} \\ \omega_1 e^{-i\omega t} & \omega_0 \end{pmatrix}, \tag{2.43}$$

where $\omega_1 > 0$ is a parameter proportional to the amplitude of the oscillating field. Let us evaluate the wave function $|\psi(t)\rangle$ at time $t > 0$ assuming that the system is in the ground state of the unperturbed Hamiltonian

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.44}$$

at $t = 0$. Note that we cannot simply exponentiate the Hamiltonian since it is time-dependent. Surprisingly, however, the following trick makes it time-independent. Let us consider the following unitary transformation:

$$|\phi(t)\rangle = e^{-i\omega\sigma_z t/2}|\psi(t)\rangle. \tag{2.45}$$

A straightforward calculation shows that $|\phi(t)\rangle$ satisfies

$$i\frac{d}{dt}|\phi(t)\rangle = \tilde{H}|\phi(t)\rangle, \tag{2.46}$$

where

$$\begin{aligned}\tilde{H} &= e^{-i\omega\sigma_z t/2}He^{i\omega\sigma_z t/2} - ie^{-i\omega\sigma_z t/2}\frac{d}{dt}e^{i\omega\sigma_z t/2} = \frac{1}{2}\begin{pmatrix} -\omega_0+\omega & \omega_1 \\ \omega_1 & \omega_0-\omega \end{pmatrix} \\ &= -\frac{\delta}{2}\sigma_z + \frac{\omega_1}{2}\sigma_x \end{aligned} \tag{2.47}$$

is, in fact, time-independent. Here $\delta = \omega_0 - \omega$ stands for the "detuning" between ω and ω_0. Note that the Hamiltonian $\tilde{H}$ can be put into the form (2.39) as

$$\tilde{H} = \frac{\Delta}{2}\left(\frac{\omega_1}{\Delta}\sigma_x - \frac{\delta}{\Delta}\sigma_z\right), \quad \Delta \equiv \sqrt{\delta^2+\omega_1^2}. \tag{2.48}$$

Now it is easy to solve Eq. (2.46). The time evolution operator is obtained using Eq. (2.40) as

$$\begin{aligned}\tilde{U}(t) &= \cos\frac{\Delta t}{2}I - i\left(\frac{\omega_1}{\Delta}\sigma_x - \frac{\delta}{\Delta}\sigma_z\right)\sin\frac{\Delta t}{2} \\ &= \begin{pmatrix} \cos\dfrac{\Delta t}{2} + i\dfrac{\delta}{\Delta}\sin\dfrac{\Delta t}{2} & -i\dfrac{\omega_1}{\Delta}\sin\dfrac{\Delta t}{2} \\ -i\dfrac{\omega_1}{\Delta}\sin\dfrac{\Delta t}{2} & \cos\dfrac{\Delta t}{2} - i\dfrac{\delta}{\Delta}\sin\dfrac{\Delta t}{2} \end{pmatrix}. \end{aligned} \tag{2.49}$$

The wave function $|\phi(t)\rangle$ with the initial condition $|\phi(0)\rangle = (1,0)^t$ is

$$|\phi(t)\rangle = \tilde{U}(t)|\phi(0)\rangle = \begin{pmatrix} \cos\dfrac{\Delta t}{2} + i\dfrac{\delta}{\Delta}\sin\dfrac{\Delta t}{2} \\ -i\dfrac{\omega_1}{\Delta}\sin\dfrac{\Delta t}{2} \end{pmatrix}. \tag{2.50}$$

We find $|\psi(t)\rangle$ from Eq. (2.45) as

$$|\psi(t)\rangle = e^{i\omega\sigma_z t/2}|\phi(t)\rangle = \begin{pmatrix} e^{i\omega t/2}\left(\cos\dfrac{\Delta t}{2} + i\dfrac{\delta}{\Delta}\sin\dfrac{\Delta t}{2}\right) \\ -ie^{-i\omega t/2}\dfrac{\omega_1}{\Delta}\sin\dfrac{\Delta t}{2} \end{pmatrix}. \tag{2.51}$$

Suppose the applied field is in resonance with the energy difference of two levels, namely $\omega = \omega_0$. We obtain $\delta = 0$ and $\Delta = \omega_1$ in this case. The wave function $|\psi(t)\rangle$ at later time $t > 0$ is

$$|\psi(t)\rangle = e^{i\omega\sigma_z t/2}|\phi(t)\rangle = \begin{pmatrix} e^{i\omega_0 t/2}\cos\dfrac{\omega_1 t}{2} \\ -ie^{-i\omega_0 t/2}\sin\dfrac{\omega_1 t}{2} \end{pmatrix}. \tag{2.52}$$

The probability with which the system is found in the ground (excited) state of H_0 is given by

$$P_0 = \cos^2\omega_1 t/2 \quad (P_1 = \sin^2\omega_1 t/2). \tag{2.53}$$

This oscillatory behavior is called the **Rabi oscillation**. *The frequency ω_1 is called the* **Rabi frequency**, *while Δ in Eq. (2.48) is called the* **generalized Rabi frequency**

2.7 Notes and open problems

The EPR states and GHZ states are useful in the verification of the Bell's inequality and its generalization; see [26, Theorems 4.5 and 4.7]. These results show that the entanglement of quantum states cannot be explained by the hidden variable theory suggested in the EPR paper, and the Copenhagen interpretation for entanglement is valid. Experiments to verify the Bell's inequality and its generalizations have been performed. See https://arxiv.org/pdf/quant-ph/0504166.pdf

Here are some open problems on quantum states.

1. Let $\{|x_1\rangle, \ldots, |x_k\rangle\}, \{|y_1\rangle, \ldots, |y_k\rangle\} \subseteq \mathbb{C}^n$. It is known that there is a unitary $U \in \mathbf{M}_n$ such that $U|x_j\rangle = |y_j\rangle$ for all j if and only if $(\langle x_r|x_s\rangle) = (\langle y_r|y_s\rangle)$.

   If $n = n_1 n_2$, determine when there will be unitary $U = U_1 \otimes U_2$ such that $U|x_j\rangle = |y_j\rangle$ for all j.

2. Suppose many identical copied of pure state $\rho = |\psi\rangle\langle\psi| \in \mathbf{M}_n$ is given. Using the orthogonal projection $P_j = |e_j\rangle\langle e_j|$ for $j = 1, \ldots, n$, corresponding to the standard basis $\{|e_1\rangle, \ldots, |e_n\rangle\}$ for $\mathbb{C}^n$, the measurement can provide estimate for the diagonal entries of ρ. Suppose $n = 2^m$ so that $|\psi\rangle$ is an m qubit state. A unitary matrix of the form $U_1 \otimes \cdots \otimes U_m$ with $U_1, \ldots, U_m \in \mathrm{U}(2)$ is called local unitary. In state tomography problem, it is known that there are $M = 3^m$ local unitary matrices

$V_1, \ldots, V_M \in \mathbf{M}_n$ such that the diagonal entries of $V_j \rho V_j^\dagger$ can be used to determine/reconstruct ρ.

Can we determine ρ using fewer than 3^m local unitary matrices?

If we only want to distinguish pure states, find the smallest number of local unitary matrices $W_1, \ldots, W_K$ such that no two pure states ρ_1, ρ_2 will have zero diagonal entries for all $W_j(\rho_1 - \rho_2)W_j^\dagger$.

Exercises for Chapter 2

EXERCISE 2.1. *Let us consider a Hamiltonian*

$$H = -\frac{\hbar}{2}\omega\sigma_y. \tag{2.54}$$

Suppose the initial state of the system is

$$|\psi(0)\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.55}$$

(1) Find the wave function $|\psi(t)\rangle$ at later time $t > 0$.
(2) Find the probability for the system to have the outcome $+1$ upon measurement of σ_z at $t > 0$.
(3) Find the probability for the system to have the outcome $+1$ upon measurement of σ_x at $t > 0$.

EXERCISE 2.2. *Prove that if $\mathcal{B}_1$ and $\mathcal{B}_2$ are orthonormal bases for $\mathbb{C}^m$ and $\mathbb{C}^n$, then $\mathcal{B} = \{|u\rangle \otimes |v\rangle : |u\rangle \in \mathcal{B}_1, |v\rangle \in \mathcal{B}_2\}$ is an orthonormal basis for $\mathbb{C}^{mn} = \mathbb{C}^m \otimes \mathbb{C}^n$.*

EXERCISE 2.3. *Suppose U is a cloning unitary transformation, such that*

$$|\Psi\rangle \equiv U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$$
$$|\Phi\rangle \equiv U|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle$$

for arbitrary $|\psi\rangle$ and $|\phi\rangle$.
(1) Write down $\langle\Psi|\Phi\rangle$ in all possible ways.
(2) Show, by inspecting the result of (1), that such U does not exist.

EXERCISE 2.4. *Let $|\psi(\theta,\phi)\rangle$ be the state given by Eq. (2.15). Show that*

$$\langle\psi(\theta,\phi)|\boldsymbol{\sigma}|\psi(\theta,\phi)\rangle = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{2.56}$$

where the left hand side is defined as

$$(\langle\psi(\theta,\phi)|\sigma_x|\psi(\theta,\phi)\rangle, \langle\psi(\theta,\phi)|\sigma_y|\psi(\theta,\phi)\rangle, \langle\psi(\theta,\phi)|\sigma_z|\psi(\theta,\phi)\rangle).$$

EXERCISE 2.5. *Find the density matrix of a pure state (2.15) and write it in the form of Eq. (2.16).*

EXERCISE 2.6. *Let ρ be given by Eq. (2.16). Show that*

$$\langle \boldsymbol{\sigma} \rangle = \mathrm{Tr}\,(\rho \boldsymbol{\sigma}) = \boldsymbol{u}. \tag{2.57}$$

EXERCISE 2.7. *The Bell basis is obtained from the binary basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ by a unitary transformation. Write down the unitary transformation explicitly.*

EXERCISE 2.8. *Find the expectation value of $\sigma_x \otimes \sigma_z$ measured in each of the Bell states.*

EXERCISE 2.9. *In many quantum algorithms, the result of an action of a function f on x is encoded into the form*

$$U_f : \mathcal{N} \sum_x |x\rangle|0\rangle \mapsto \mathcal{N} \sum_x |x\rangle|f(x)\rangle,$$

where $|x\rangle$ stands for the tensor product state $|b_{n-1}b_{n-2}\ldots b_0\rangle$ with $x = b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \ldots + b_0$ and $\mathcal{N}$ is the normalization constant. The first register is for the input x, while the second one is for the corresponding output $f(x)$. Note that U_f acts on all possible states simultaneously.

Let $f(x) = a^x \bmod N$, where a and N are coprime, and consider the state

$$U_f \left[\frac{1}{\sqrt{512}} \sum_{x=0}^{511} |x\rangle|0\rangle \right] = \frac{1}{\sqrt{512}} \sum_{x=0}^{511} |x\rangle|a^x \bmod N\rangle$$

with $a = 6$ and $N = 91$. Suppose the measurement of the first register results in (1) $x = 11$, (2) $x = 23$ and (3) $x = 35$. What is the state immediately after each measurement?

References

[1] P. A. M. Dirac, *Principles of Quantum Mechanics* (4th ed.), Clarendon Press (1981).

[2] L. I. Shiff, *Quantum Mechanics* (3rd ed.), McGraw-Hill (1968).

[3] A. Messiah, *Quantum Mechanics*, Dover (2000).

[4] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (3rd Edition), Cambridge University Press, Cambridge (2020).

[5] L. E. Ballentine, *Quantum Mechanics*, World Scientific, Singapore (1998).

[6] A. Peres, *Quantum Theory: Concepts and Methods*, Springer (2006).

[7] E. H. Kennard, Z. Phys. **44**, 326 (1927).

[8] H. P. Robertson, Phys. Rev. **34** 163, (1929).

[9] R. Horodecki *et al.*, eprint, quant-ph/0702225 (2007).

[10] A. Acín *et al.*, Phys. Rev. Lett. **85**, 1560 (2000).

[11] A. Acín *et al.*, Phys. Rev. Lett. **87**, 040401 (2001).

[12] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[13] M. Horodecki *et al.*, Phys. Lett. A **223**, 1 (1996).

[14] G. Vidal, J. Mod. Opt. **47**, 355 (2000).

[15] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[16] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[17] E. Riefell and W. Polak, ACM Computing Surveys (CSUR) **32**, 300 (2000) and E. G. Rieffel and W. H. Polak, Quantum Computing: A Gentle Introduction, The MIT Press (2011).

[18] M. A. Neilsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).

[19] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).

[20] D. Dieks, Phys. Lett. A **92**, 271 (1982).

[21] D. M. Greenberger, M. A. Horne and A. Zeilinger, in *'Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, ed. M. Kafatos, Kluwer, Dordrecht (1989). Also avilable as arXiv:0712.0921 [quant-ph].

[22] W. Dür, G. Vidal and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[23] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. **41**, 777 (1935).

[24] C. H. Bennett *et al.*, Phys. Rev. A **59**, 1070 (1999) and D. P. DiVincenzo, D. W. Leung and B. M. Terhal, IEEE Trans. Info. Theory **48**, 580 (2002). See also A. SaiToh, R. Rahimi and M. Nakahara, Phys. Rev. A **77**, 052101 (2008).

[25] C. H. Bennett and G. Brassard, in Proc. IEEE Int. Conf. Comp., Systems and Signal Processing **175** (1984).

[26] W. Scherer, Mathematics of Quantum Computing, An Introduction, Springer (2019).

[27] https://en.wikipedia.org/wiki/Speed_of_light

[28] https://en.wikipedia.org/wiki/Random_number_generation