Quantum Information and Computing-Measurement Postulates

Dipan Kumar Ghosh
Department of Physics
Indian Institute of Technology Bombay
Powai, Mumbai 400076

March 16, 2017

1 Introduction

In our discussion on postulates of quantum mechanics we described the salient points of the Copenhagen interpretation to comprise of

1. description of the state of a system as a ray in the Hilbert space,
2. relationship between the operators and observables,
3. the unitary character of the time evolution of the quantum states, and,
4. the collapse of a state during the process of measurement.

We mentioned that the quantum description of a state has two distinct regime, one of unitary evolution when the system is not under observation and a second a non-unitary process when the system interacts with a measuring apparatus. We pointed out that when we measure some physical observable in a pure system, the result that we obtain is probabilistic which projects out one of the possible eigenstates of the operator that is being measured. Measurement, thus, is an essential postulate of quantum mechanics. It must, however, be pointed out that quantum mechanics postulates that possible results arise with predetermined probabilities but does not provide a mechanism how a particular result actually arises. In this lecture and the next we will discuss the measurement postulate and its various ramification.

2 Measurement Postulates

We have seen that when we make a measurement of a state, the state to which it will collapse would depend on the basis in which we measure the system. For instance, if we have a state $\alpha \mid 0\rangle + \beta \mid 1\rangle$ which is prepared in the computational basis, it would collapse to a state $\mid 0\rangle$ with a probability $\mid \alpha \mid^2$ and to a state $\mid 1\rangle$ with a probability $\mid \beta \mid^2$, *provide, the measurement also is made in the computational basis.* If the same state is measured in, for instance, the diagonal basis, the state would collapse to $\mid +\rangle$ with a probability $\mid \alpha+\beta \mid^2 /2$ and to a state $\mid -\rangle$ with a probability $\mid \alpha-\beta \mid^2 /2$. The process of measurement described in this manner is confusing and it is more elegant to describe it in terms of a mathematical formalism of which the projection operators happen to be special cases. The mathematical formulation of the process of measurement is stated in the form of a separate postulate of quantum mechanics. The essential features of measurement postulates are as follows:

1. A measurement in quantum mechanics is described by a set of operators $\{M_m\}$, where m is an index representing the possible results of the measurement process and if there are n number of possible results, then m could take values from 1 to n.

2. The probability of observing a particular outcome m is given by $\langle\psi \mid M_m^\dagger M_m \mid \psi\rangle$.

3. Since there are only n possibilities, the set of operators must satisfy completeness relation $\sum_{m=1}^{n} M_m^\dagger M_m = I$. Here I is the identity operator.

4. According to the quantum postulates, after one has made a measurement and got a specific result m, the system would have collapsed to a state which is an eigenstate of the operator corresponding to eigenvalue m and the (normalized) post-measurement of the system would be
$$\frac{M_m \mid \psi\rangle}{\sqrt{\langle\psi \mid M_m^\dagger M_m \mid \psi\rangle}}$$

We will now discuss two specific cases of measurements.

3 Projective or Von-Neumann Measurement

A concept which is central to the principle of measurement and wave function collapse is that the process of measurement is repeatable, i.e., if one repeats the measurement on a system after a measurement has been performed on the system (but before the system has interacted with other devices), the result of the measurement will be the same. This implies that the process of measurement has not disturbed the pre-existing state of the system and what one measures in a process of measurement is the property associated with the state of the system which exists before the measurement was performed. The wave function collapse does not worry about such details but treats the process of measurement as a black box which measures the pre-existing state of the system. (This presumption is not tenable; for instance, when one measures the momentum of a particle through a collision experiment, the momentum of the target particle after measurement is not the

same as it was before such a measurement.)
To determine the state $\mid \psi\rangle$, we proceed as follows. We measure a property A and obtain a value. If the corresponding operator has a non-degenerate spectrum. we simply take a large number of identically prepared state of the system and measure the same property and obtain a distribution of the occurrence of N mutually exclusive values from which we reconstruct $\mid \psi\rangle$ using the fact that the state of the system before a measurement was a linear combination of the eigenstates of the operator corresponding to the physical observable that is measured.
If the eigenvalue is degenerate, one must find a 'compatible operator' $\hat{B}$ (which commutes with $\hat{A}$) and determine an eigenstate state $\mid n_A, n_B\rangle$ and obtain the distribution of the corresponding eigenstates. The simultaneous measurement of a compatible set of physical properties determine a *maximal set* for a state vector with N mutually exclusive outcome. Mathematically, a maximal set corresponds to defining N one dimensional projectors P_n (which, in this case is the measurement operator M_n) having the following property

$$P_m^\dagger = P_m \qquad P_m P_n = P_m \delta_{m,n}; \quad \sum_m P_m = I \tag{1}$$

Since $P_m^2 = P_m$, the probability in this case is given by $p(m) = \langle\psi \mid P_m \mid \psi\rangle$. Since its eigenvalues are 0 and 1, P is a positive operator. If we measure a physical variable M with non-degenerate eigenvalues, λ_α is measured. This is equivalent to the spectral resolution of the corresponding operator $\hat{M}$

$$\hat{M} = \sum_{m=1}^{n} m P_m$$

The above set of projectors constitute a **von Neumann or projective measurement**. It may be observed these operators constitute a special set of the general measurement operator defined above with the additional constraints which follow from the representation of P_m in terms of the orthogonal basis set (corresponding to the eigenstate of the operator for the observable that is being measured) which it projects, i.e. $P_m = \mid \psi_m\rangle\langle\psi_m \mid$.

- They are hermitian, which follows from the representation in terms of the states which they project, since $P_m^\dagger = (\mid \psi_m\rangle\langle\psi_m \mid)^\dagger = \mid \psi_m\rangle\langle\psi_m \mid$.
- They are idempotent as $P_m P_m = (\mid \psi_m\rangle\langle\psi_m \mid)(\mid \psi_m.\rangle\langle\psi_m \mid) = (\mid \psi_m\rangle\langle\psi_m \mid) = P_m$.
- They are orthogonal since $P_m P_n = (\mid \psi_m\rangle\langle\psi_m \mid)(\mid \psi_n\rangle\langle\psi_n \mid) = (\mid \psi_m\rangle\delta_{m,n}\langle\psi_n \mid) = P_m\delta_{m,n}$.

To illustrate, consider the following example of the state $\alpha \mid 0\rangle + \beta \mid 1\rangle$ prepared in the computational basis having a matrix representation $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. We define the measurement operators to be $M_0 =\mid 0\rangle\langle 0\mid$ and $M_1 =\mid 1\rangle\langle 1\mid$. The probability of getting the state $\mid 0\rangle$ is given by

$$\begin{aligned}
p(0) &= \langle\psi\,|M_0|\,\psi\rangle \\
&= \langle\psi\,|0\rangle\langle 0\,|\psi\rangle \\
&= (\alpha^*\langle 0\mid +\beta^*\langle 1\mid) \mid 0\rangle\langle 0\mid (\alpha\mid 0\rangle + \beta\mid 1\rangle) \\
&=\mid \alpha\mid^2
\end{aligned}$$

Likewise, the probability of getting the state $\mid 1\rangle$ is $p(1) =\mid \beta\mid^2$.
Suppose the result of the measurement happens to be $\mid 0\rangle$, the post measurement state is given by

$$\frac{M_0\mid\psi\rangle}{\sqrt{\langle\psi\,|M_0|\,\psi\rangle}} = \frac{\mid 0\rangle\langle 0\mid(\alpha\mid 0\rangle + \beta\mid 1\rangle}{\mid\alpha\mid} = \frac{\alpha\mid 0\rangle}{\mid\alpha\mid} = e^{i\varphi}\mid 0\rangle$$

Thus the post-measurement state is the state $\mid 0\rangle$ itself but for an immaterial global phase factor.
A special case of the projective measurement is when the measurement is done in a basis, such as the example of the computational basis shown above. Projective measurements may not be restricted to basis but may be performed in any orthogonal states in the Hilbert space. If we measured the state given above in the diagonal basis, instead of the computational basis, we must first express the given state in terms of the new basis $\{\mid +\rangle, \mid -\rangle\}$, where

$$\mid +\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}; \quad \mid -\rangle = \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}$$

in terms of which

$$\mid 0\rangle = \frac{\mid +\rangle + \mid -\rangle}{\sqrt{2}}; \quad \mid 1\rangle = \frac{\mid +\rangle - \mid -\rangle}{\sqrt{2}}$$

and the state $\mid\psi\rangle$ can be written as

$$\mid\psi\rangle = \frac{\alpha+\beta}{\sqrt{2}}\mid +\rangle + \frac{\alpha-\beta}{\sqrt{2}}\mid -\rangle$$

The probabilities in this case are given by $p(+) = \dfrac{\mid\alpha+\beta\mid^2}{2}$ and $p(-) = \dfrac{\mid\alpha-\beta\mid^2}{2}$.
Suppose we get the state $\mid +\rangle$ on measurement in this basis. We can find the post measurement state in an identical manner

$$\sqrt{2}\frac{M_+\mid\psi\rangle}{\mid\alpha+\beta\mid} = e^{i\varphi}\mid +\rangle$$

At this state, if we made repeated measurement on this post-measurement state, we would continue to get the same state over and over again as the state having already collapsed to the basis will continue to be there.

We now extend our ideas to multi-qubit case. We consider the case of two qubits here. Generalization to more than two qubits is straightforward. Let

$$\mid \psi\rangle = \alpha \mid 00\rangle + \beta \mid 01\rangle + \gamma \mid 10\rangle + \delta \mid 11\rangle$$

In this case one can think of different measurements. Suppose we are interested in measuring only the first qubit, irrespective of what the state of the second qubit might be. in that case, the measurement operator is given by (here the first element of the two qubit refers to the first qubit and the second element to the second qubit)

$$\begin{aligned} M_0 &= M_{00} + M_{01} \\ &= \mid 00\rangle\langle 00 \mid + \mid 01\rangle\langle 01 \mid \\ &= \mid 0\rangle\langle 0 \mid (\mid 0\rangle\langle 0 \mid + \mid 1\rangle\langle 1 \mid) \\ &= \mid 0\rangle\langle 0 \mid \otimes I_2 \end{aligned}$$

where I_2 is the identity operator which acts on the second qubit. The probability of getting the first qubit to be $\mid 0\rangle$ is given by

$$\begin{aligned} \langle\psi \mid M_0 \mid \psi\rangle &= \langle\psi \mid (\mid 0\rangle\langle 0 \mid \otimes I_2) \mid \psi\rangle \\ &= \mid \alpha \mid^2 + \mid \beta \mid^2 \end{aligned}$$

In the same way if we calculated the probability for the first qubit to be $\mid 1\rangle$ we would get it to be $\mid \gamma + \delta \mid^2$.

In the preceding, we talked about the general measurement theory in quantum mechanics. we discussed the special case of projective or Von Neumann measurement for pure states. In this lecture we would first generalize the above to the case of mixed states where the description is in terms of the density matrix rather than the state vector. Later in this lecture we will discuss a non-projective measurement known as the POVM or the Positive Operator Valued Measure.

Before proceeding, it is convenient to state the quantum postulates as modified for the case of mixed systems. They are

1. The ensemble is described by a density matrix ρ in the Hilbert space.
2. The time evolution of the density matrix is given by the Liouville equation.
3. If a measurement of the system is made by a set of measurement operators $\{M_m\}$, the probability of a particular outcome m is given by $\mathrm{Tr}(M_m^\dagger M_m \rho)$. The post-measurement state of the system is given by the new density matrix $\dfrac{M_m^\dagger \rho M_m}{\mathrm{Tr}(M_m^\dagger M_m \rho)}$.

4 Measurement in a mixed state

Recall our discussion on the density operator where the operator ρ was defined by a weighted sum of the pure states in an ensemble

$$\rho = \sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid$$

where p_i is the classical probability of occurrence of the state $\mid \psi_i\rangle$ in the ensemble. If we consider a basis in this Hilbert space $\{\mid e_i\rangle\}$, then we may write $\mid \psi_i\rangle = \sum_j C_{ij} \mid e_j\rangle$, where C_{ij} is the Born probability, given by $\langle e_j \mid\psi_i\rangle$. The probability of finding the state $\mid \psi_i\rangle$ in the basis state $\mid e_j\rangle$ is simply given by $\mid C_{ij} \mid^2$. Thus the probability of observing the system in the basis state basis state $\mid e_j\rangle$ is given by multiplying the probability of picking up the state $\mid \psi_i\rangle$ from the ensemble (which is p_i) with the probability of finding the the state $\mid \psi_i\rangle$ in the basis state $\mid e_j\rangle$ and then summing over all states $\mid \psi\rangle$ of the system. This is given by

$$\begin{aligned}
p(\mid e_j\rangle) &= \sum_i p_i \mid \langle e_j \mid\psi_i\rangle \mid^2 \\
&= \sum_i p_i\langle e_j \mid\psi_i\rangle\langle\psi_i \mid e_j\rangle \\
&= \langle e_j \mid(\sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid)\mid e_j\rangle \\
&= \langle e_j \mid\rho\mid e_j\rangle \qquad (2)
\end{aligned}$$

where $\rho = \sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid$.

Let us recall our formulation of quantum mechanics using density operators. Let the density matrix of the system be ρ. The probability of finding the system in a state $\mid n\rangle$ is given by

$$p_n = \mathrm{Tr}(\rho \mid n\rangle\langle n \mid) = \mathrm{Tr}(\rho P_n) \qquad (3)$$

It is easy to see that the probability found in (2) is the same as that of (3). To see this, examine the following, considering the fact that trace can be calculated in any representation. Considering a basis set $\{\mid e_i\rangle\}$

$$\begin{aligned}
p_n &= \mathrm{Tr}(\rho P_n) \\
&= \sum_i \langle e_i \mid\rho\mid n\rangle\langle n \mid e_i\rangle \\
&= \sum_i \langle e_i \mid\rho\mid n\rangle\delta_{i,n} \\
&= \langle n \mid\rho\mid n\rangle
\end{aligned}$$

Defining the post-measurement density matrix by $\rho' = \mid \psi_{post-measure}\rangle\langle\psi_{post-measure} \mid$

$$\rho' = \frac{P_n \mid \psi\rangle\langle\psi \mid P_n^\dagger}{\langle\psi \mid P_n \mid \psi\rangle} = \frac{P_n \rho P_n}{\mathrm{Tr}(\rho P_n)} \tag{4}$$

Given a complete set of projectors, the probability of a particular result is given by taking the trace of the product of the pre-measurement density matrix with the projection operator for the corresponding event.

$$Pr(m) = \mathrm{Tr}\,\rho P_m = \langle\psi \mid P_m \mid \psi\rangle$$

the last relation in the above equation is valid for pure states only. The completeness of the set of the projection operators ensures probability conservation,

$$\sum_m Pr(m) = \sum_m \mathrm{Tr}\,\rho P_m = \mathrm{Tr}\,\rho \sum_m P_m = \mathrm{Tr}\,\rho = I$$

Example : Let $\mid \psi\rangle = \alpha \mid 0\rangle + \beta \mid 1\rangle$. Suppose we we measure the z component of the spin, using the spectral decomposition of the corresponding operator $\sigma_z = \mid 0\rangle\langle 0 \mid - \mid 1\rangle\langle 1 \mid$, we have using $M_0 = \mid 0\rangle\langle 0 \mid, \rho = \mid \psi\rangle\langle\psi \mid = (\alpha \mid 0\rangle + \beta \mid 1\rangle)(\alpha\langle 0 \mid + \beta\langle 1 \mid)$, we get

$$\begin{aligned}\mathrm{Tr}(M_0^\dagger M_0 \rho) &= \mathrm{Tr}\,[\mid 0\rangle\langle 0 \mid 0\rangle\langle 0 \mid (\alpha \mid 0\rangle + \beta \mid 1\rangle)(\alpha^*\langle 0 \mid + \beta^*\langle 1 \mid)] \\ &= \mathrm{Tr}\,[\mid 0\rangle\alpha(\alpha^*\langle 0 \mid + \beta^* \mid 1\rangle)] = \mid \alpha \mid^2\end{aligned}$$

where we have used the orthonormality of the states and used the fact that $\mathrm{Tr}(\mid a\rangle\langle b \mid) = \langle b \mid a\rangle$. The post measurement state of the system is easily obtained by observing that

$$M_0 \rho M_0^\dagger = \mid \alpha \mid^2 \mid 0\rangle\langle 0 \mid$$

so that

$$\rho_{post-measurement} = \frac{M_0 \rho M_0^\dagger}{\mathrm{Tr}(M_0^\dagger M_0 \rho)} = \mid 0\rangle\langle 0 \mid$$

Another interesting that emerges is what happens to the system if we make yet another measurement on the system after one measurement has been made. Obviously, once the system has collapsed to a particular state when a measurement has been made in one basis, it would continue to reman in the same state on repeating the measurement in the same basis. However, this does not remain true if one makes a change of basis. To illustrate, let us consider the same example as before and assume that a measurement in the computational basis has been made and we found the state collapsing to $\mid 0\rangle$ with a probability $\mid \alpha \mid^2$. If we now make another measurement on the resulting state in the diagonal basis, as the state $\mid 0\rangle$ can be expressed as $\dfrac{\mid +\rangle + \mid -\rangle}{\sqrt{2}}$, the state would collapse to $\mid +\rangle$ and $\mid -\rangle$ with equal probability. Thus the result would be to get the state $\mid +\rangle$

with probability $\frac{|\alpha|^2}{2}$ and a state $|-\rangle$ with the same probability. If however, we had measured the original state first in the diagonal basis, we would have go the state $|+\rangle$ with a probability $\frac{|\alpha+\beta|^2}{2}$ and the state $|-\rangle$ with the probability $\frac{|\alpha-\beta|^2}{2}$. Suppose we got the state to be $|+\rangle$ and we now measure the state in the computational basis. We would get the states $|0\rangle$ and $|1\rangle$ with equal probability. Thus the probability of getting the result $+,0$ is $\frac{|\alpha+\beta|^2}{4}$ which is different from the probability $|\alpha|^2$ of getting $0,+$.
The process of measurement is a *postulate* in quantum mechanics because it is a result of interaction between the system and the measuring apparatus, which cannot be treated quantum mechanically. The measurement process described above is a direct measurement. It may be noted that as a consequence of measurement, all information about the pre-measurement state of the system is obliterated. In a more general measurement, the system is allowed to interact with an external system (the ancilla) and this joint system develops unitarily.

5 POVM

Another special class of measurement that we consider is the POVM measurement, which is an acronym for Positive Operator Valued Measure, a name whose origin is not clear and need not concern us. This is a non-projective and non-orthogonal measurement. (Projective operators, are special cases of this class). Further, for a d- dimensional Hilbert space, there are exactly d orthogonal projectors while in POVM they may be more than the dimensionality.
Consider a composite system in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Assuming that the joint system is prepared in a sate $\rho_A \otimes \rho_B$, we perform a joint measurement of the physical property M_{AB} in the Hilbert space of $\mathcal{H}_A \otimes \mathcal{H}_B$. Denoting the complete set of projectors of the combined space by P_m, the probability of outcome m is

$$p(m) = \mathrm{Tr}_A\mathrm{Tr}_B\left[P_m(\rho_A \otimes \rho_B)\right] = \mathrm{Tr}_A[E_m\rho_A] \tag{5}$$

where

$$E_m = \mathrm{Tr}_B[P_m\rho_B] \tag{6}$$

Comparing the expression (5) with (2), we find that E_m defined by (6) is the projection operator corresponding to the measurement of our system of interest A from the composite system (AB). The set of operators $\{E_m\}$ is known as POVM. These operators act on $\mathcal{H}_A$. It can be seen that they are indeed a set of projectors. In terms of matrix elements, we write

$$(E_m)_{i,j} = \sum_{\mu\nu}(P_m)_{i\mu,j\nu}(\rho^B)_{\mu,\nu}$$

1. Hermiticity :

$$(E_m)^*_{i,j} = \sum_{\mu\nu}(P_m)^*_{i\mu,j\nu}(\rho^B)^*_{\mu,\nu}$$
$$= \sum_{\mu,\nu}(P_m)_{j\nu,i\mu}(\rho^B)_{\nu,\mu}$$
$$= (E_m)_{ji} \tag{7}$$

   In deriving above, we have used hermiticity of P_m and of ρ^B.

2. E_m is a positive operator. Consider a basis in which ρ^B is diagonal : $\rho^B = \sum_\mu p_\mu \mid \mu\rangle\langle\mu\mid$. We have

$$\langle\psi_A\,|E_m|\,\psi_a\rangle = \sum_\mu p_\mu\langle\psi\otimes\mu_B\,|P_m|\,\psi_A\otimes\mu_B\rangle \geq 0$$

   since P_M is a positive operator.

3. Completeness :

$$\sum_m E_m = \mathrm{Tr}_B[\sum_m(P_m)\rho^B] = \sum_m P_m \times \mathrm{Tr}_B\rho^B = I_A$$

   since $\sum_m P_m = I_{AB}$ and $\mathrm{Tr}_B\rho^B = 1$.

However, the projectors E_m are not orthogonal, i.e. $E_mE_n \neq \delta_{m,n}$. We define a set of POVM as any set of operators for which $E_m = E_m^\dagger, E_m \geq 0$ and $\sum_m E_m = I$. Since the measurement operators $\{M_m\}$ act on a quantum system in the state $\mid\psi\rangle$ so that the post measurement state is $\dfrac{M_m\mid\psi\rangle}{\sqrt{\langle\psi\,|M_m^\dagger M_m|\,\psi\rangle}}$, the above properties will be satisfied if we define $E_m = M_m^\dagger M_m$. The probability of measurement is $p(m) = \mathrm{Tr}(E_m\rho)$. If a measurement is made but not read, the state goes to $\rho \to \rho' = \sum_m M_m\rho M_m$. If it is read, it becomes $\rho' = \dfrac{M_m\rho M_m^\dagger}{\mathrm{Tr}[M_m\rho M_m^\dagger]}$. Reading is an important part of any measurement because until a measurement has been made but not read, the state, in principle, remains in a linear combination of possible results but once the reading is done, it collapses into the state that is read. (Reading is essentially a process of interaction with equipment such as a meter, an oscilloscope etc.)

Example

Suppose Alice gives Bob a quantum state prepared in either $\mid\psi_1\rangle = \mid 0\rangle$ and $\mid\psi_2\rangle$. Bob decides to use a POVM basis to find out what the state was. (This is an example from

Nielsen and Chuang's book). Define,

$$E_1 = \frac{\sqrt{2}}{\sqrt{2}+1} \mid 1\rangle\langle 1 \mid$$
$$E_2 = \frac{1}{2}\frac{\sqrt{2}}{\sqrt{2}+1} [(\mid 0\rangle - \mid 1\rangle)(\langle 0 \mid - \langle 1 \mid)]$$
$$E_3 = I - E_1 - E_2$$

Clearly, these are not orthogonal as we are in two dimensional space in which we have defined three operators. The interesting point about this case is it allows us to identify non-orthogonal states correctly with some probability without ever wrongly identifying a state.

Suppose Alice sent Bob a state randomly, which could be either $\mid 0\rangle$ or the state $\dfrac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}$. Bob cannot distinguish these states by making a measurement in the computational basis as the second state has a projection along both $\mid 0\rangle$ and $\mid 1\rangle$, which means that only when he measured the state to be $\mid 1\rangle$, he may be sure that it was the second state but would never be able to identify the first state correctly. Thus in a random situation, he would only be correct 25% time.

The situation improves if he uses POVM instead. If Bob had received $\mid 0\rangle$, he will observe E_2 with zero probability since $\langle 0 \,| E_1 |\, 0\rangle = 0$. Hence if his measurement results in E_1, he will be able to identify the state as $\mid \psi_2\rangle$. Similarly, if his measurement gives E_2, he will be able to identify it as $\mid \psi_1\rangle$. In case his measurement results in E_3, no definite conclusion can be made. Thus if he identifies a state he is never wrong but may not always be able to make a definite statement on some occasions. .