A MATHEMATICAL MODEL FOR MEASUREMENTS IN QUANTUM MECHANICS

TUYEN TRUNG TRUONG

Abstract. Let $V = \mathbb{C}^N$, and $H$ (an observable) a Hermitian linear operator on $V$. Let $v_1, \ldots, v_n$ be an orthonormal basis for $V$. Let $\mathcal{M}$ be a measurement apparatus prepared to measure a state of an observed system and collapses the state to one of the $v_j$'s. The states of $\mathcal{M}$ are represented by $W = \mathbb{C}^m$. Let the time interval to perform a measurement to be $[0, 1]$.

Let $\mathcal{S}$, with state $\xi$, be an observed system to be measured by $\mathcal{M}$. Then, as well-known, the combined system $\mathcal{S} + \mathcal{M}$ can be regarded to have a Schrodinger's unitary evolution $\hat{U}(t) = e^{-iHt/\hbar} : V \otimes W \to V \otimes W$, $t \in [0, 1]$.

By the set up of $\mathcal{M}$, there are $N$ "gates" $W_j := \hat{U}(1)^{-1}(v_j \otimes W)$ from which $\xi \otimes W$ must choose one to go through. To explain the randomness of the results, we define a notion of capacity for the gates, and a complete ordering on capacities of different gates. For a linear subspace $Z \subset V \otimes W$, we let $P_Z$ be the projection onto $Z$. The evolutions $\hat{U}_j(t) := \hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}\hat{U}(1)$ are characteristic for $\mathcal{M}$, satisfy the same Schrodinger's equation as $\hat{U}(t)$, and $\sum_{j=1}^N \hat{U}_j(t) = \hat{U}(t)$.

To obtain the picture of what happens from the view of $V$ alone, we take the traces of $\hat{U}_j(t)$, that is

$$U_j(t) = \text{Tr}_W(\hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}\hat{U}(1)) = \sum_{i=1}^m \langle \hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}(\hat{U}(1))\xi \otimes w_i, w_i \rangle.$$

Here, $w_1, \ldots, w_m$ are an orthonormal basis for $W$. If $\text{Tr}_W(\hat{H}P_{v_j \otimes W}\hat{U}(1)) = \text{Tr}_W(P_{v_j \otimes W}\hat{U}(1))$ for all $j = 1, \ldots, N$, then $\mathcal{S}$ and $\mathcal{M}$ may be regarded as independent after the measurement.

Instead of its compatibility with experiments and many applications, the probability nature of Quantum Mechanics is still much mysterious. In Quantum Mechanics (see, for example [2 1]), an observable is a linear operator $H : V \to V$, where $V$ is a complex Hilbert space. The operator $H$ is assumed to be Hermitian, which means that $H = H^\dagger$ where $H^\dagger$ is the conjugate transpose of a complex matrix $H$. Moreover, it is assumed that the eigenvectors $\{v_i\}_{i \in I}$ of $H$ form a basis for $V$. An eigenstate of $H$ is an eigenvector of $H$. A state (or a ket) is a vector $\xi \in V$. We view vectors $v$ and $cv$ as representing the same state, for any $c \neq 0$ a complex number.

For simplicity, in the rest of this paper we will discuss only the case where $V$ is of finite dimensional. Therefore, we fix $V = \mathbb{C}^N$ with the usual inner product $\langle \cdot, \cdot \rangle$ and $H : \mathbb{C}^N \to \mathbb{C}^N$ a Hermitian linear operator. We denote by $\lambda_1, \ldots, \lambda_m$ the distinct eigenvalues of $H$, and by $V_j$ ($j = 1, \ldots, m$) the corresponding eigenspaces. We denote by $\mathcal{U}(n)$ the group of unitary linear operators, i.e. linear operators $U$ on $\mathbb{C}^N$ such that $U^\dagger U = UU^\dagger = \text{Id}$. Here $U^\dagger$ is the conjugate transpose of a complex matrix $U$.

The most intriguing feature of Quantum Mechanics is the contrast between the smooth evolution of Schrödinger’s evolution on one side, and the abruptly behavior under measurements.
A state $v$ may involve in time, so we denote by $v(t)$ its dependence on time. Let $H(t)$ be the Hamilton of the system at time $t$, then Schrödinger’s equation asserts that the state $v(t)$ will vary according to the rule

$$i\hbar \frac{d}{dt}v = H(t)v,$$

here $\hbar$ is the Planck constant. In particular, if $H(t) = H$ is constant then the evolution of $v(t)$ is unitary and smooth

$$v(t) = e^{-\frac{i}{\hbar}Ht}\varphi,$$

where $\varphi = v(0)$.

If $\xi \in V$ is an eigenstate of eigenvalue $\lambda$ of $H$, then the result of a measurement of $\xi$ will certainly give the value $\lambda$. If $\xi$ is not an eigenstate, then by the collapse of wave functions, a measurement will behave very wild:

1) The result we obtain will be one of the eigenvalues $\lambda_1, \ldots, \lambda_m$. Moreover, we will obtain the value $\lambda_j$ with probability proportional to $|<\xi, P_j\xi>|^2$, where $P_j : \mathbb{C}^N \rightarrow V_j$ is the projection onto the eigenspace $\lambda_j$.

2) After the measurement, $\xi$ will become one vector in $V_j$.

**Remark 0.1.** The above rules are those of the Copenhagen’s interpretation. The experiments that satisfy these rules are called experiments of the first kind. There are however, many more experiments that do not satisfy these rules.

Now the contrast between the Schrödinger’s evolution and the evolution under a measurement comes from the superposition principle in Quantum Mechanics. This principle says that if a system can be in states $v_1, \ldots, v_k$, then it can also be in the state $c_1v_1 + \ldots + c_kv_k$ for any complex numbers $c_1, \ldots, c_k$. Therefore if we measure a state $\xi$ which is not an eigenstate, then before the measurement $\xi$ varies in a deterministic manner, but after the measurement then $\xi$ varies randomly.

The contrast above, known as the measurement problem, has been extensively discussed since the beginning of Quantum Mechanics, with many interpretations (see for example [1, 2, 3, 4]). Here we will follow the well-known approach that the combined "observed system" + "measurement apparatus" should satisfy the Schrödinger’s evolution. One contemporary representative of this interpretation is the quantum decoherence theory. This approach regards the combined "observed system" + "measurement apparatus" as entangled in and even after the measurement process, and thus why after the measurement the wave function collapses has not been explained.

The purpose of the current paper is to give an interpretation of how after the measurement we may regard the observed system and the measurement apparatus to be independent. We arrive equations of what happens in the measurement process, as viewed from the combined "observed system" + "measurement apparatus", or from the observed system alone. The measurement process is modelled as having $N$ gates, and the randomness appear since a choice of which gate to enter needs to be made. The process of how a choice is made is deterministic with respect to
the measurement apparatus, but looks random to a human observer because he may not know many of the measurements that \( \mathcal{M} \) made.

Here we give the mathematical details. Recall that \( V = \mathbb{C}^N \), and \( H \) (an observer) is a Hermitian linear operator on \( V \). We let \( v_1, \ldots, v_n \) be an orthonormal basis for \( V \). A measurement apparatus \( \mathcal{M} \) is designed to measure states of observed systems and reduce the state to one of the \( v_j \)'s. A system \( \mathcal{S} \), with state \( \xi \), will be measured by \( \mathcal{M} \). The time interval of each measurement is taken to be \([0, 1]\).

**Notations.** For a linear subspace \( Z \) of a vector space, we let \( P_Z \) be the projection onto \( Z \). We will use \( \langle, \rangle \) for the usual inner product on either \( V \), \( W \) or \( V \otimes W \).

0.1. **Schroedinger’s evolution of \( \mathcal{S} + \mathcal{M} \).** The result in this subsection is well-known. The combined system \( \mathcal{S} + \mathcal{M} \), may be regarded as closed. As so, they have a definite Hermitian Hamiltonian \( \hat{H} \), and satisfies the Schrodinger’s equation. Therefore, a vector \( z \in V \otimes W \) will evoke as \( \hat{U}(t)z \) where

\[
i \hbar \hat{U}'(t) = \hat{H} \hat{U}(t), \quad \hat{U}(0) = Id.
\]

That is, \( \hat{U}(t) = e^{-i\hat{H}t/\hbar} \).

0.2. **Gates.** We can imagine how the measurement process evolves as follows. First, the measurement apparatus \( \mathcal{M} \) registers the state \( \xi \). This \( \xi \), a state in \( V \), corresponds to the vector space \( \xi \otimes W \subset V \otimes W \). If \( \xi \otimes W \) is to end up at one of the \( v_j \)'s at the end of the measurement process, then it must be brought to one of the vector spaces which are brought to \( v_j \otimes W \) under the unitary operator \( U(1) \). That is, \( \xi \times W \) must go through one of the gates

\[
W_j := \hat{U}(1)^{-1}(v_j \times W),
\]

and then proceeds under the unitary operator \( \hat{U}(t) \).

The projections \( P_{W_j} : V \otimes W \to W_j \) may be regarded as the preparation before the measurement takes place.

**Remark.** The procedure to choose a gate to go into will be discussed in the last subsection.

0.3. **Probabilities.** The probability of \( \xi \otimes W \) entering the gate \( W_j \) is given by how close it is to the gate. Since \( \hat{U}(t) \) is unitary, it is the same as how close \( \hat{U}(1)\xi \otimes W \) is to \( v_j \otimes W \). Therefore, it is proportional to the length of the trace of the projection from \( \hat{U}(1)\xi \otimes W \) to \( v_j \otimes W \). The latter vector is

\[
Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi = \sum_{i=1}^m \langle P_{v_j \otimes W} (\hat{U}(1)\xi \otimes w_i), w_i \rangle.
\]

Here \( w_1, \ldots, w_m \) are an orthonormal basis for \( W \).

Therefore, the probability to choose the gate \( W_j \) is proportional to

\[
|Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2 = |\sum_{i=1}^m \langle P_{v_j \otimes W} (\hat{U}(1)\xi \otimes w_i), w_i \rangle|^2.
\]

Note that this is the same as the probability given by the density matrix in the approach of Quantum Decoherence, see [4].
Remark. In the above, only the probabilities of obtaining the results are given. We will discuss in the last subsection how randomness of the results happens in the measurements.

0.4. Equations of the measurement process. From the point of view of the combined system $S + M$, when the result of a measurement is $v_j$, then the value $v_j \otimes W$ is definite. Hence to characterize the measurement process, we need to go backward from $v_j \otimes W$

$$\hat{U}_j(t) := \hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}\hat{U}(1).$$

Here, for a linear subspace $Z \subset V \otimes W$, $P_Z$ denotes the projection onto $Z$. We note that $\hat{U}_j(t)$ also satisfies the Schrödinger’s equation

$$ih\hat{U}_j'(t) = \hat{H}\hat{U}_j(t).$$

Moreover, since $v_1, \ldots, v_N$ are an orthonormal basis for $V$, it follows that

$$\sum_{j=1}^{N} \hat{U}_j(t) = \hat{U}(t).$$

From the point of view of the observed system $S$ alone, the evolution is the trace of the $\hat{U}_j(t)$’s, hence is characterized by

$$U_j(t)\xi = Tr_W(\hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}\hat{U}(1))\xi = \sum_{i=1}^{m} <\hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}(\hat{U}(1)\xi \otimes w_i), w_i> .$$

Here, $w_1, \ldots, w_m$ are an orthonormal basis for $W$.

0.5. Independence of $S$ and $M$ after the measurement. It is reasonable to regard that if at the end of a measurement, the Hamiltonian of the system $S$ returns to $H$, then $S$ is independent from $M$. From the definitions of $U_j(t)$ we find that

$$ihU_j'(t)\xi = Tr_W(\hat{H}\hat{U}(t)\hat{U}(1)^{-1}P_{v_j \otimes W}\hat{U}(1))\xi .$$

Hence, at $t = 1$

$$ihU_j'(t)\xi = Tr_W(\hat{H}P_{v_j \otimes W}\hat{U}(1))\xi .$$

Hence, if

$$Tr_W(\hat{H}P_{v_j \otimes W}\hat{U}(1)) = HT_{tr_W}(P_{v_j \otimes W}\hat{U}(1)) ,$$

for all $j = 1, \ldots, N$, then we may regard $S$ and $M$ as independent after the measurement.

0.6. Source of randomness. In this subsection, we propose a simple model to explain the randomness of the measurements in Quantum Mechanics.

To each gate $W_j = \hat{U}(1)^{-1}(v_j \otimes W)$, there associates a real number $\rho_j$ called its capacity. We give the following complete ordering on the pairs $(W_j, \rho_j)$ ($j = 1, \ldots, N$). We define $(W_j, \rho_j) > (W_k, \rho_k)$ if: Either $\rho_j > \rho_k$, or $\rho_j = \rho_k$ and $j > k$.

When $M$ measures an $S$ with state $\xi$, then the following happens:

1) The gates for which $Tr_W(P_{v_j \otimes W}\hat{U}(1))\xi = 0$ are disregarded.

2) For every $j = 1, \ldots, N$, $\rho_j$ is changed to $\rho_j + |Tr_W(P_{v_j \otimes W}\hat{U}(1))\xi|^2$. 
3) $S$ will go into the gate $W_{j_0}$, where $(W_{j_0}, \rho_{j_0})$ is the maximum element, under the ordering defined in the previous paragraph, among those with $|Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2 > 0$. Then $\rho_{j_0}$ is changed to $\rho_{j_0} - 1$.

Here we normalize $\xi$ so that

$$\sum_{j=1}^{N} |Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2 = 1.$$ 

The probability of the results of this process can be shown to be the same as the probabilities we computed before.

**Proof.** By the normalized condition

$$\sum_{j=1}^{N} |Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2 = 1,$$

it follows that

$$\sum_{j=1}^{N} \rho_j = C$$

is unchanged, no matter how many measurements have been made. Now assume that a measurement is made in the time interval $[0, 1]$. We let $\rho_j$ be the value of the capacity of the gate $W_j$ at time 0, and let $\rho'_j$ be the value of the capacity of the gate $W_j$ at time 1. Let $\rho''_j = \rho_j + |Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2$ be the values of the capacities after Step 2 of the measurement process. Then

$$\sum_{j=1}^{N} \rho''_j = \sum_{j=1}^{N} \rho_j = C + 1.$$ 

Hence if $\rho_j < -|C| - 2$, there will be one $k$ with $\rho''_k > 0 > \rho''_j$. It follows that the gate $W_j$ will not be chosen, and hence after the measurement $\rho'_j = \rho''_j \geq \rho_j$.

It follows from this that there is a constant $B$ such that we always have

$$\rho_j > B$$

for all $j$. In fact, before the measurement, assume that for all $j = 1, \ldots, N$ we have $\rho_j > B$, for some $B < -|C| - 3$. Then after the measurement, all $\rho_j$ with $\rho_j < -|C| - 2$ will have $\rho'_j \geq \rho_j \geq B$. There is only one $j_0$ with $\rho'_{j_0} \leq \rho_{j_0}$, but for that one then $\rho'_{j_0} \geq \rho_{j_0} - 1 \geq -|C| - 2 - 1 > B$.

We also have an upper bound for $\rho_j$. In fact, for any $j = 1, \ldots, N$ we always have

$$\rho_j = \sum_{k=1}^{N} \rho_k - \sum_{k \neq j} \rho_k = C - \sum_{k \neq j} \rho_k < C - (N-1)B.$$ 

Now let $n$ be a large integer, and let $n_j$ be the number of times that in $n$ measurements, the gate $W_j$ has been chosen $n_j$ times. Then after $n$ measurements we have

$$B < \rho_j = n|Tr_W(P_{v_j \otimes W} \hat{U}(1))\xi|^2 - n_j < C - (N-1)B.$$
Therefore
\[
\lim_{n \to \infty} \frac{n_j}{n} = |\text{Tr}_W (P_{v_j \otimes W} \hat{U}(1))\xi|^2
\]
for all \( j = 1, \ldots, N \).

Now, the result of each measurement is \textbf{deterministic} with respect to the view of the measurement apparatus. However, to the view of a human observer, the results look \textbf{random} because many measurements that \( \mathcal{M} \) made may not be known to him.

We may take the view that these capacities are associated either to measurement apparatus or to observed systems. However, it seems that these capacities are more natural associated to the measurement apparatus than to observed systems. The reason is that before any measurement is made, the observed system has no preference on what states it will collapse to. Therefore, if it was to keep information on a certain orthonormal basis \( v_1, \ldots, v_N \), it must do so for all such orthonormal bases, and this will be somehow burdensome. Therefore, we conclude that the information on the capacities are kept within the measurement apparatus.

\textbf{Remark.} The above model may be interpreted as follows. The capacities are the energies of the gates \( W_j \). When an observed system is measured, then it provides a total 1 unit of energy to the gates corresponding with the probabilities. In this sense, we may say that the observed system \( S \) both goes through all the gates and chooses one specific gate to go. When \( S \) goes through the gate \( W_{j_0} \), then \( W_{j_0} \) must send out the result, thus needs to spend energy. This is why \( W_{j_0} \) should have the maximum energy, and why after the measurement then the energy of \( W_{j_0} \) is decreased by 1 unit. We see that energy is preserved in this measurement process.

\textbf{Generalizations.} The results above extend without difficulties to the case where \( \mathcal{M} \) is designed to collapse the states to one of pairwise orthogonal linear subspaces \( V_1, \ldots, V_l \) of \( V \).

\textbf{Remark.} This model does not explain the entanglement in quantum mechanics.

\textbf{References}

[1] R. I. G. Hughes, \textit{The structure and interpretation of Quantum Mechanics}, Harvard University Press, 7th edition, 2003.
[2] P. A. M. Dirac, \textit{Quantum Mechanics}, Oxford University Press, 4th edition, 1958.
[3] R. Penrose, \textit{The road to reality}, Vintage Books, 2004.
[4] The Wikipedia’s pages on "Wave function collapse" and "Quantum decoherence".

\textit{E-mail address: truongt@indiana.edu}