Quantum Foundations 1 – The role of observables in Quantum Mechanics & the Measurement problem

Angelo Bassi

The axioms of Quantum Mechanics

States. A Hilbert space $\mathcal{H}$ is associated to a given physical system, depending on its degrees of freedom. The state of the system is represented by a normalized vector in $\mathcal{H}$, called state vector. Given two systems with associated Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, to the composite system the tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is associated .

Dynamics. The state vector evolves according to the Schrödinger equation.

Observables. Observable quantities are represented by self-adjoint operators:

$$A \rightarrow \hat{A}$$

The possible outcomes on an experiment designed to measure the observable are the eigenvalues $|a_n\rangle$ of the associated operator.

Born rule. In measurements, out come are random and are distributed with the Born rule:

$$\mathbb{P}[a_n] = |\langle a_n|\psi\rangle|^2$$

where $|\psi\rangle$ represents the state of the system at the time of measurement, and $|a_n\rangle$ the eigenstate of the observable $\hat{A}$ associated to the eigenvalue a_n.

State vector collapse. After a measurement, the state vector $|\psi\rangle$ changes to $|a_n\rangle$. This is also called the von Neumann wave function collapse postulate.

See book of Sakurai [to be completed].

Notes. In presenting the postulates, we assumed that the operator $\hat{A}$ has only discrete eigenvalues, which are not degenerate. From the fundamental point of view not much changes if the spectrum of the operators is richer.

We also described what is called an ideal measurement, where the outcome is perfectly and uniquely identified.

Observables as Operators

The fact that observable quantities, such as the position or the energy of a particle, should be represented by such abstract objects as self-adjoint operators, and that the possible values the observable can take are the eigenvalues of the associated operator, calls for an explanation.

In fact, this axiom can be derived by reconsidering the Born rule and the von Neumann collapse, together with a proper understanding of what happens in a (ideal) quantum measurement.

Consider the phenomenology on a measurement M_A of a certain variable A of a quantum system.

1. The measurement has **outcomes**, which we label as $\mathbf{a_n}$**;** they are assumed to be discrete for the sake of simplicity. they occur randomly.

2. A measurement **changes the initial state of the system;** for example, the statistics for the outcome of the measurement M_A changes if before it another measurement M_B is performed. Moreover, the measurement **selects a specific set of final states** depending on the outcomes; these state are such that, if considered as initial states for another run of the same experiment, the outcome can be predicted with certainty. Let us call $|a_n\rangle$ the state of the system after the measurement, whose outcome was a_n.

Therefore we have the association:

$$M_A \rightarrow \{a_n, |a_n\rangle\}_n$$

3. The Born rule states that if the initial state of the system is $|\psi\rangle$, the probability to obtain outcome a_n is equal to $\mathbb{P}[a_n] = |\langle a_n|\psi\rangle|^2$.

As we see, so far we did not introduce self-adjoint operators.

We can easily characterize the states $|a_n\rangle$, which so far are arbitrary (but normalized).

We invoke again reproducibility: When the same measurement is (immediately) repeated, the outcome of the first measurement is reproduced. Given the Born rule, then the states $|a_n\rangle$, are orthogonal. Moreover, these state must span the entire Hilbert space, otherwise the probability distribution given by the Born rule would not sum to 1. Hence **the states $|a_n\rangle$, form a basis of the Hilbert space** associated to the system being measured.

The **spectral theorem** then allows to associate an operator to the elements $\{a_n, |a_n\rangle\}_n$ according to:

$$\hat{A} = \sum_n a_n |a_n\rangle\langle a_n|$$

and this brings to the association of self-adjoint operators to measurements on quantum systems:

$$M_A \rightarrow \{a_n, |a_n\rangle \text{ basis vector}\}_n \rightarrow \hat{A}$$

Then, self-adjoint operators are a convenient way to collect the information about an experiment on a quantum system. For example, the average value of the observable is

$$\langle A \rangle = \sum_n a_n \mathbb{P}[a_n] = \sum_n a_n |\langle a_n|\psi\rangle|^2 = \langle \psi|\hat{A}|\psi\rangle$$

and so on for all other statistical quantities.

We can then rephrase the axioms of QM as follows.

States. To each physical system a Hilbert space $\mathcal{H}$ is associated (the rest as before).

Dynamics. The state vector evolves according to the Schrödinger equation.

State vector collapse. After a measurement, the state vector $|\psi\rangle$ changes to one among a set of selected states $|a_n\rangle$.

Born rule. Measurement outcomes are random and are distributed with the Born rule:

$$\mathbb{P}[a_n] = |\langle a_n|\psi\rangle|^2$$

where $|\psi\rangle$ represents the state of the system at the time of measurement. This implies that the states $|a_n\rangle$ from a basis of the Hilbert space.

Then we have the following **theorem**: the statistics of the outcomes of the measurement is fully represented by the self-adjoint operator

$$\hat{A} = \sum_n a_n |a_n\rangle\langle a_n|$$

Self-adjoint operators do not play any fundamental role in QM. But they have a very important practical role.

Actually, canonical quantization is more than than simply assuming that outcomes of measurements are represented by self-adjoint operators. The basic properties of these operators are defined by the **correspondence rule** between classical Poisson brackets and quantum commutators of canonical variables:

$$\{A, B\} \quad \rightarrow \quad \frac{[\hat{A}, \hat{B}]}{i\hbar}$$

Also the correspondence rule can be derived as a theorem, again by a proper analysis of what a measurement is.

Le us first remind what conjugate variables are. Let us consider a one-dimensional particle on unit mass, with Lagrangian

$$L = L(x, \dot{x}) = \frac{1}{2}\dot{x}^2 + V(x)$$

where x is its position. The conjugate momentum is:

$$p = \frac{\partial L}{\partial \dot{x}} = \dot{x}$$

x and p are canonical variables and satisfy the Poisson bracket

$$\{x,p\} = 1$$

The key point for our discussion is that momentum is the time derivative of the position, and therefore it is measured ad distance covered over time.

Let us now move to the **quantum case**. In the following, we set $\hbar$=1. We assume the particle to be free

The position of the particle can take any value on the real line, therefore the associated operator $\hat{x}$ has eigenvalues x ∈ $\mathbb{R}$ and eigenvectors |x>, which from a basis of the Hilbert space.

On the wave functions ψ(x) = <x|ψ>, the operator acts as a multiplication operator:

$$\hat{x}\psi(x) = x\psi(x)$$

since $\langle x|\hat{x}|\psi\rangle = x\langle x|\psi\rangle = x\psi(x)$

Suppose that, as dictated by the Schrödinger equation, the dynamics is linear and admits a wave solution:

$$\psi(x,t) = \int dp\, c(p) e^{i(px-\omega_p t)}$$

If ω is independent of p, the wave packet does not move in space over time, therefore the model is not physical. If ω is linear in p: ω = vp, then any wave packet moves with velocity v, and as such the velocity is not a dynamical variable (like the speed of light c). Therefore the interesting case is when ω is at least quadratic on p:

$$\omega_p = \frac{1}{2}p^2$$

This is what is predicted by the Schrödinger equation, for a particle with the Lagrangian L written before.

Therefore:

$$\psi(x,t) = \int dp\, c(p) e^{i(px - p^2 t/2)}$$

The group velocity is v = p: a wave packet moves with a group velocity equal to the velocity of a classical particle (of unit mass). But being the dispersion relation quadratic in p, the dynamics is dispersive. Assuming for simplicity that the initial wave packet is a Gaussian centered in the origin (see book of Duerr for the general case), we have:

$$\psi(x,0) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2} + ip_0 x}$$

Its Fourier transform is:

$$\psi(k,0) = \sqrt{\frac{\sigma}{\sqrt{\pi}}} e^{-\frac{\sigma^2}{2}(p-p_0)}$$

The wave function at time t is relatively easy to compute and the solution can be found on textbooks. Here we are interested in its square modulus, which reads

$$|\psi(x,t)|^2 = \frac{1}{\sqrt{\pi}\sigma_t} e^{-\frac{(x-vt)^2}{\sigma_t^2}}$$

With v = p_0 and

$$\sigma_t^2 = \sigma^2 \left(1 + \frac{t^2}{\sigma^4}\right) \quad \xrightarrow[t \gg \sigma^2]{} \quad \frac{t^2}{\sigma^2}$$

This means that for $t >> \sigma^2$:

$$|\psi(x,t)|^2 \sim \frac{\sigma}{\sqrt{\pi}t} e^{-\frac{\sigma^2}{t^2}(x-vt)^2}$$

Now, since the wave packet is initially centered in the origin, to say that a particle has momentum p means that it has moved to x = pt. In other words, we first measure the position of the particle at time t=0 (and we find it in the origin), then after some time t we measure the position again and we find it at x = pt, and we conclude that it had a momentum p. We can compute such a probability density:

$$\begin{aligned} \mathbb{P}_p^{\psi_t}([p, p+dp]) &= \mathbb{P}_x^{\psi_t}([pt, pt+dx]) = |\psi(pt,t)|^2 dx = \\ &= \frac{\sigma}{\sqrt{\pi}t} e^{-\sigma^2(p-p_0)^2} dx = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2(p-p_0)^2} dp = \\ &= |\psi(p,0)|^2 dp \end{aligned}$$

Therefore the distribution in momentum is equal to the square modulus of the Fourier transform of the initial wave function at time 0. This result is not related to having used a Gaussian wave packet and can be proven in a general case. But with this example we can easily appreciate the different steps.

COMMENTO sul significato del limite $t >> \sigma^2$

We further have:

$$\begin{aligned}\langle p\rangle &= \int dp\, p|\psi(p,0)|^2 = \int dp\, \psi^*(p,0)p\psi(p,0)\\ &= \int dx\, \psi^*(x,0)\left[-i\frac{d}{dx}\right]\psi(x,0)\\ &= \langle\psi_0|\hat{p}|\psi_0\rangle\end{aligned}$$

with (remember we are in the position representation):

$$\hat{p} = -i\frac{d}{dx}$$

From which it also follows that

$$[\hat{x},\hat{p}] = i$$

The argument can be extended to more complex situations. The argument has been presented by position and momentum, but works for any two **canonical variables A and B** such that B = Å, which is true if the Lagrangian is quadratic in A (in all physical situations, let aside the electromagnetic case, which can be treated separately) and if the Schrödinger equation is a second order differential equation with respect to A.

The measurement problem

The measurement problem is the acknowledgement that the axioms of QM are inconsistent because they assume two completely different types of dynamics:

- The Schrödinger dynamics, which is linear and deterministic
- The collapse of the wave function, which is nonlinear and stochastic

without telling exactly when one applies in place of the other.

In a way, there is no surprise in the fact that these two dynamics are present: the Schrödinger dynamics applies to well isolated system (like two quantum particles) while the collapse of the wave function refers to measurements, when a quantum system interacts with a large macroscopic device. Therefore **one might think that the collapse is a simplified way of treating the system-device interaction, which is ultimately described by the Schrödinger equation**, or one of its generalizations like the Dirac equation.

We show that **this is not the case**. Let us consider, in rather general terms, a measurement process.

Microscopic system. For the sake of simplicity, we assume that it is described by a finite dimensional Hilbert space $\mathcal{H}_S$, of dimension N. Let as call $|a_n\rangle$ a basis of the space.

Macroscopic apparatus. It has a controllable degree of freedom, like the position of a pointer, or of a needle, or some outcome on a screen, together with many (mainly microscopic) uncontrollable/uncontrolled degrees of freedom.

We label its (normalized) state vector as $|A, \alpha\rangle$, where A denoted the controllable degree of freedom and α of the other degrees of freedom (potentially also those of the surrounding environment, if needed).

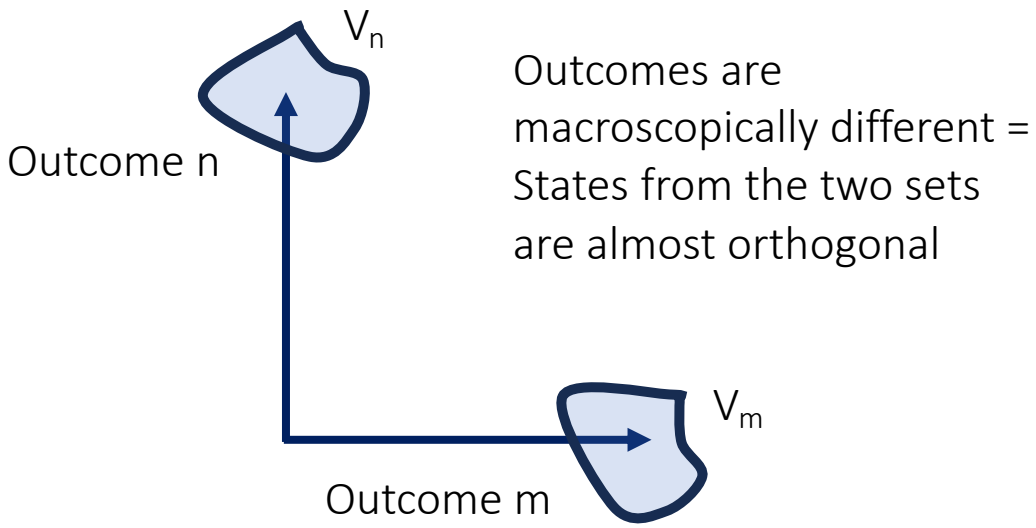
We ask the device to have a ready state (pointer or needle set to 0, or empty screen) which we denote as $|A_0, \alpha\rangle$, and at least N **macroscopically different states** $|A_n, \alpha\rangle$ (n = 1, …, N) corresponding to the n possible outcomes. By macroscopically different states we mean the following. Let us call:

V_n = {α such that $|A_n, \alpha\rangle$ represents outcome n, which is macroscopically different from all other outcomes}

We say that two outcomes A_n and A_m (n ≠ m) are macroscopically different if the two associated sate vectors are almost **orthogonal**

$$\inf_{|A_n,\alpha\rangle \in V_n, |A_m,\beta\rangle \in V_m} \| |A_n, \alpha\rangle - |A_m, \beta\rangle \| \geq \sqrt{2} - \eta$$

with $\eta \ll 1$. (η takes into account the possibility that wave functions have tails).



System-Apparatus interaction. The device is initially prepared in the ready state $|A_0, \alpha\rangle$; we do not have control on the degrees of freedom label by α, therefore we assume that they are distributed with some probability $\mathbb{P}(\alpha)$. In other words, the state of the state of the apparatus is described by a statistical mixture, typically represented by a density matrix. Since the density matrix formalism is ambiguous when discussing foundational issues (see later discussion), we keep working at the wave function level.

Suppose the microscopic system is in the state $|a_n\rangle$. The joint initial state is:

$$|a_n\rangle \otimes |A_0, \alpha\rangle, \quad \alpha \text{ distributed with prob. } \mathbb{P}(\alpha)$$

The measurement interaction is unitary and maps the initial state in some final state

$$|a_n\rangle \otimes |A_0, \alpha\rangle \quad \rightarrow \quad |F, a_n, \alpha\rangle$$

So, far, any physical interactions behave like that. But here we are considering special cases called measurement processes. What we demand is that the interaction is such that the device measures the observable to which, according the previous discussion, the operator having $\{|a_n\rangle , n = 1, .., N\}$ as eigenstates is associated; and it does so with sufficient accuracy. This means that $|F, a_n, \alpha\rangle$ should correspond, most of the times, to outcome n. Formally, if we define

$$J_n = \{\alpha : |F, a_n, \alpha\rangle \in V_n\}$$

then:

$$\mathbb{P}(J_n) \geq 1 - \epsilon$$

with ε a suitably small number.

From this it follows that

$$J_{n,m} \equiv J_n \cap J_m \;\Rightarrow\; \mathbb{P}(J_{n,m}) \geq 1 - 2\epsilon$$

Now we can state the measurement problem. Suppose that the microscopic system is initially prepared in the superposition of e.g. two states $|a_n\rangle$ and $|a_m\rangle$, for some fixed n and m. The by linearity of the interaction:

$$\frac{1}{\sqrt{2}}[|a_n\rangle + |a_m\rangle] \otimes |A_0, \alpha\rangle \;\rightarrow\; |F, a_n + a_m, \alpha\rangle \equiv \frac{1}{\sqrt{2}}[|F, a_n, \alpha\rangle + |F, a_m, \alpha\rangle]$$

The question is: does $|F, a_n + a_m, \alpha\rangle$ correspond to some measurement outcome? The answer is that, in most cases, it does not. Thake $\alpha \in J_{n,m}$, which is highly probable. Then:

$$\begin{aligned} \| |F, a_n + a_m, \alpha\rangle - |F, a_n, \alpha\rangle \| &= \| (1/\sqrt{2} - 1)|F, a_n, \alpha\rangle + 1/\sqrt{2}|F, a_m, \alpha\rangle \| \\ &\leq 1 - 1/\sqrt{2} + 1/\sqrt{2} = 1 \end{aligned}$$

This means that $|F, a_n + a_m, \alpha\rangle$ does not belong to V_n because it violates the orthogonality condition stated before, and as such it does not correspond to outcome n; by a similar argument one proves it it does not correspond to outcome m, and more generally to any macroscopic definite state.

This is essentially the Schrödinger's cat paradox, rephrased by taking into account the complexity of the macroscopic device (or of the cat), which does not help in changing the conclusion: **the Schrödinger's dynamics is not capable of taking into account the collapse of the wave function; if this is removed from the theory, then microscopic superpositions easily turn into macroscopic superpositions, which are not part of our experience**.

It is our conviction that a reinterpretation of the wave function (many worlds, consistent histories, model interpretations, relational approaches) does not suffice to solve the measurement problem. We believe that this problem indicates that the theory needs to be modified/replaced. In the following, we present two attempts in this direction.

Decoherence does not solve the measurement problem

The previous argument has shown that the quantum-to-classical transition, which is ensured by the collapse of the wave function, cannot be explained within the standard quantum formalism.

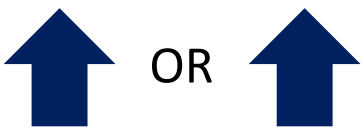
Yet, very often it is claimed that decoherence, i.e. the interaction of a system with its environment, provides such an explanation: microscopic systems can be isolated from the surrounding environment and show a quantum behavior; macroscopic object cannot and therefore behave classically.

There is truth in this; yet decoherence alone in not capable to solve the measurement problem. The reason is the following.

Superposition state

$$|\psi\rangle = \frac{1}{\sqrt{2}}[|L\rangle + |R\rangle]$$

Localized states



$$50\%\ |L\rangle, \quad 50\%\ |R\rangle$$

Density matrix (in the |L>, |R> basis)

$$\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In the language of the density matrix, to pass from a delocalized state to a localized state, one has to find a mechanics that changes a non diagonal density matrix (in the position representation) into a diagonal one.

Decoherence does the job! Therefore we have:

1. Take a delocalized state (Schrödinger's cat)	$\|\psi\rangle = \frac{1}{\sqrt{2}}[\|L\rangle + \|R\rangle]$
2. Write the state as a density matrxi (in position)	$\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
3. Turn on decoherence and compute the reduced density matrix for the system	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
4. Interpret the result as representing a localized state	$50\%\ \|L\rangle, \quad 50\%\ \|R\rangle$

The problem is that le last step is wrong. The reason is that the association between statistical mixtures and statistical operators is not 1-to-1 but many-to-1:

$50\%\ |L\rangle, \quad 50\%\ |R\rangle$

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$50\%\ |L\rangle, \quad 50\%\ |R\rangle$

Not only it is not legitime to associate a specific statistical mixture to a given density matrix, without further information. In this case, the specific association is wrong.

Consider two particles in a singlet state; suppose one of the two is lost. Then we have to trace over the degrees of freedom of the lost particle in order to have the reduced density matrix of the remaining one. The reduced density matrix (in the spin basis) is (1/2)I, as the one encountered before. Can we conclude that the particle has a definite spin?

The answer is no: if we were capable of recovering the other particle, we can perform a Bell test to check that the two share an entangled state. And in an entangled state none of the particles has a definite spin, according to the standard doctrine.

Similarly, in the case of a macroscopic object interacting with the environment, in principle we could do a Bell test involving the system and the particles of the environment: the conclusion – in the Schrödinger equation is correct – is that they share an entangled state. The system is not localized somewhere. Of course it is practically impossible to perform such a test, but here we are not worried about that. The important thing it that the test is not forbidden by the rules of QM.

The conclusion is very simple: due to the interaction with the environment, the system's wave function becomes rapidly entangled with it. The entanglement is such that, if a measurement (a quantum measurement with all the interpretative problems it carry along) is performed, one does not see interference in position. But the wave function does not corresponds to that of a localized object.

As we anticipated, decoherence is an important physical phenomenon. But, alone, it does not explain the emergence of a classical world from a quantum one.