3. Axioms of Quantum Mechanics

3.1 Introduction
3.2 The axioms of quantum mechanics
3.2.1 Observables and State Space
3.2.2 Quantum measurement
3.2.3 Law of motion
3.3 Strong measurements
3.3.1 Expectation values
3.3.2 Uncertainty relationships
3.3.3 Repeated measurements and Quantum Zeno Effect

3.1 Introduction

Every physical theory is formulated in terms of mathematical objects. It is thus necessary to establish a set of rules to map physical concepts and objects into mathematical objects that we use to represent them[5]. Sometimes this mapping is evident, as in classical mechanics, while for quantum mechanics the mathematical objects are not intuitive. In the same way as classical mechanics is founded on Newton's law or electrodynamics on the Maxwell-Boltzmann equations, quantum mechanics is also based on some fundamental laws, which are called the postulates or axioms of quantum mechanics. The axioms we are going to see apply to the dynamics of closed quantum systems. We want to develop a mathematical model for the dynamics of closed systems: therefore we are interested in defining states, observables, measurements and evolution. Some subtleties will arise since we are trying to define measurement in a closed system, when the measuring person is instead outside the system itself. We give below (and explain in the next few sections) one formulation of the QM axioms. Different presentations (for example starting from density operators instead of state vectors) are possible.

1. The properties of a quantum system are completely defined by specification of its state vector $|\psi\rangle$. The state vector is an element of a complex Hilbert space $\mathcal{H}$ called the space of states.
2. With every physical property $\mathcal{A}$ (energy, position, momentum, angular momentum, ...) there exists an associated linear, Hermitian operator A (usually called observable), which acts in the space of states $\mathcal{H}$. The eigenvalues of the operator are the possible values of the physical properties.

3.a If $|\psi\rangle$ is the vector representing the state of a system and if $|\varphi\rangle$ represents another physical state, there exists a probability $p(|\psi\rangle, |\varphi\rangle)$ of finding $|\psi\rangle$ in state $|\varphi\rangle$, which is given by the squared modulus of the scalar product on $\mathcal{H}$: $p(|\psi\rangle, |\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$ (Born Rule).

3.b If A is an observable with eigenvalues a_k and eigenvectors $|k\rangle$ ($A|k\rangle = a_k|k\rangle$), given a system in the state $|\psi\rangle$, the probability of obtaining a_k as the outcome of the measurement of A is $p(a_k) = |\langle k|\psi\rangle|^2$. After the measurement the system is left in the state projected on the subspace of the eigenvalue a_k (Wave function collapse).

4. The evolution of a closed system is unitary. The state vector $|\psi(t)\rangle$ at time t is derived from the state vector $|\psi(t_0)\rangle$ at time t_0 by applying a unitary operator $U(t, t_0)$, called the evolution operator: $|\psi(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle$.

[5] See: Leslie E. Ballentine, "Quantum Mechanics A Modern Development", World Scientific Publishing (1998). We follow his presentation in this section.

3.2 The axioms of quantum mechanics

3.2.1 Observables and State Space

A physical experiment can be divided into two steps: preparation and measurement. The first step determines the possible outcomes of the experiment, while the measurement retrieves the value of the outcome. In QM the situation is slightly different: the first step determines the *probabilities* of the various possible outcomes, while the measurement retrieve the value of a particular outcome, in a statistic manner. This separation of the experiment is reflected into the two types of mathematical objects we find in QM. The first step corresponds to the concept of a *state* of the system, while the second to observables.
The state gives a complete description of the set of probabilities for all observables, while these last ones are all dynamical variables that in principle can be measured. All the information is contained in the state, irrespectively on how I got the state, of its previous history. For the moment we will identify the state with the vectors of an Hilbert space $|\psi\rangle$. We will see later on that a more general definition exists in terms of state operators ρ.
All physical observables (defined by the prescription of experiment or measurement) are represented by a linear operator that operates in linear inner product space (an Hilbert space in case of finite dimensional spaces). States of the system are represented by the direction/ray (not a vector) in the linear inner product space (again Hilbert space in the finite dimensional case).

3.2.2 Quantum measurement

The value of the measurement of an observable is one of the observable eigenvalues. The probability of obtaining one particular eigenvalue is given by the modulus square of the inner product of the state vector of the system with the corresponding eigenvector. The state of the system immediately after the measurement is the normalized projection of the state prior to the measurement onto the eigenvector subspace.
Let A be the observable with eigenvalues a_k and eigenvectors $|k\rangle$: $A|k\rangle = a_k|k\rangle$. Given a system in the state $|\psi\rangle$, the probability of obtaining a_k as the outcome of the measurement of A in this system is

$$p(a_k) = |\langle k|\psi\rangle|^2.$$

We can also write this in terms of the k^{th} eigenvector projector $P_k = |k\rangle\langle k|$: $p(a_k) = \langle\psi|P_k|\psi\rangle$. Since here we are considering strong, projective measurement, also called Von Neumann measurements, immediately after a measurement that gave us the result a_k, the state of the system is in the $|k\rangle$ eigenstate. More precisely, the normalized output state after the measurement is

$$|\psi'\rangle = \frac{P_k|\psi\rangle}{\sqrt{|\langle\psi|P_k|\psi\rangle|}}.$$

If we repeat the experiment after the first measurement, we will obtain again the same result (with probability 1). If $|\psi\rangle$ is an eigenstate of A, $A|\psi\rangle = a_\psi|\psi\rangle$, then we will measure a_ψ with probability unity. This is the well-known collapse of the wavefunction.
The collapse of the wavefunction is of course a source of confusion and contradictions: as stated above it appears as an almost instantaneous evolution of the system from a given state to another one, an evolution which is not unitary (as evolution should be per axiom # 4). The source of contradiction stems from the fact that in this simple description of the measurement, the observer (or the measurement apparatus) are *external* to the system (thus the assumption of closed system is not respected) and might not even be quantum-mechanical. A more advanced theory of measurement attempts to solve these issues[6].
On the other side, we note that operationally the wavefunction collapse is required to define a well-formulated theory. The collapse allows the experimenter to check the result of the measurement by repeating it (on the system just observed) thus giving confidence on the measurement apparatus and procedure. If this were not the case, no measurement could be ever believed to be the correct one, so no confirmation of the theory could be done.
Reference
M. Brune, E. Hagley, J. Dreyer, X. Matre, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche *Observing the Progressive Decoherence of the Meter in a Quantum Measurement*, Phys. Rev. Lett.**77**, 4887 - 4890 (1996)

[6] In addition to the "strong" or projective measurement presented here, generalized models for measurement exist, see for example POVM in Prof. Preskill's online notes

3.2.3 Law of motion

We can define the time evolution operator U, such that

$$|\psi'\rangle = U|\psi\rangle, \text{ with } U^\dagger U = 1.$$

Since the state has all the information about the system at time t, the state of the system at the time $t+dt$ depends only on the state at time t and on the evolution operator $U(t, t+dt)$ (that thus depends only on the times t and $t+dt$, not on any previous times, otherwise it would bring extra information to the system).
The unitarity of the evolution is equivalent to the following statement regarding the evolution of the state vector. The dynamics of the system are generated by the system Hamiltonian $\mathcal{H}$ (the observable corresponding to the total energy of the system), as described by Schrödinger equation:

$$i\hbar\frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

where $\hbar$ is the reduced Planck's constant[7] (1.0545×10^-34 Js).
We would like to link this second statement (Schrödinger equation) to the previous statement regarding the unitarity of the evolution. To do so we first look at the evolution for an infinitesimal time dt.
For an infinitesimal evolution we have then: $|\psi(t+dt)\rangle = |\psi\rangle - idt\mathcal{H}|\psi\rangle$. It follows that $U(t, dt) = \mathbb{1} - i\mathcal{H}dt$. Since the Hamiltonian is a self-adjoint operator, to the same order of approximation we retrieve the fact that U is unitary: $UU^\dagger = (\mathbb{1} - i\mathcal{H}dt)(\mathbb{1} + i\mathcal{H}dt) \approx \mathbb{1} + o(dt^2)$.
We can build the dynamics for any time duration in terms of infinitesimal evolutions, $U(t,t') = U(t', t'-dt)\dots U(t+2dt, t+dt)U(t+dt, t)$ since the propagator U depends only on the time t.
If the Hamiltonian is time independent (and setting $t' = 0$), we obtain: $|\psi(t)\rangle = U(0,t)|\psi(0)\rangle$, where the evolution operator U is given by $U = e^{-i\mathcal{H}t}$, i.e. U is an exponential operator.

? Question: Show from the infinitesimal time product and the Taylor expansion for the exponential that this is indeed the case. □

Equivalently, we can find a differential equation for the dynamics of the propagator: from $U(t+dt, t_0) - U(t, t_0) = -\frac{i}{\hbar}\mathcal{H}U(t, t_0)$ we have the Schrödinger equation for the time evolution operator (propagator):

$$i\hbar\frac{\partial U}{\partial t} = \mathcal{H}U$$

This equation is valid also when the Hamiltonian is time-dependent (and we will see later on a formal solution to this equation).
As the Hamiltonian represents the energy of the system, its *spectral representation* is defined in terms of the energy eigenvalues ϵ_k, with corresponding eigenvectors $|k\rangle$: $\mathcal{H} = \sum_k \epsilon_k |k\rangle\langle k|$. The evolution operator is then: $U = \sum_k e^{-i\epsilon_k t}|k\rangle\langle k|$. The eigenvalues of U are therefore simply $e^{-i\epsilon_k t}$, and it is common to talk in terms of eigen-phases $\epsilon_k t$. If the Hamiltonian is time-independent we have also $U^\dagger = U(-t)$, it is possible to obtain an effective inversion of the time arrow.

? Question: What is the evolution of an energy eigenvector $|k\rangle$?
First consider the infinitesimal evolution: $|k(t+dt)\rangle = U(t+dt, t)\,|k(t)\rangle = (\mathbb{1} - i\mathcal{H}dt)\,|k(t)\rangle = (1 - i\epsilon_k dt)\,|k(t)\rangle$. Thus we have the differential equation for the energy eigenket: $\frac{d|k\rangle}{dt} = -i\epsilon_k\,|k\rangle$, so that $|k(t)\rangle = e^{-i\epsilon_k t}\,|k(0)\rangle$. We can also use the spectral decomposition of U: $|k(t)\rangle = U(t,0)\,|k(0)\rangle = (\sum_h e^{-i\epsilon_h t}\,|h\rangle\,\langle h|)\,|k(0)\rangle = e^{-i\epsilon_k t}\,|k(0)\rangle$. □

In conclusion, our picture of QM is a mathematical framework in which the system is completely described by its state, which undergoes a **deterministic** evolution (and invertible evolution). The measurement process, which connects the mathematical theory to the observed experiments, is probabilistic.

[7] We will quite often set $\hbar = 1$, that is, we will measure the energies in frequency units

3.3 Strong measurements

3.3.1 Expectation values

Although the result of a single measurement is probabilistic, we are usually interested in the average outcome, which gives us more information about the system and observable. The average or *expectation value* of an observable for a system in state $|\psi\rangle$ is given by

$$\langle A\rangle = \langle\psi|A|\psi\rangle$$

? Question: Prove that this can be simply derived from the usual definition of average
$\langle A\rangle = \sum_n p(a_k)a_k = \sum_n |\langle\psi|a_k\rangle|^2 a_k = \sum_n \langle\psi|n\rangle\langle n|\psi\rangle a_k = \langle\psi|(\sum_n a_k|k\rangle\langle k|)|\psi\rangle$, and we get the desired result from $A = \sum_n a_k|k\rangle\langle k|$. □

3.3.2 Uncertainty relationships

$\mathcal{D}$: **Compatible Observables** Two observables A, B are said to be compatible if their corresponding operators commute $[A, B] = 0$ and incompatible otherwise.

$\mathcal{D}$: **Degeneracy** If there exist two (or more) eigenstates of an operator A with the same eigenvalues, they are called degenerate.

We have already seen how commuting operators have common eigenvectors and how a compatible observables can be used to distinguish between degenerate eigenvectors. We now look from a more physical point of view at the meaning of commuting (or compatible) observables. Suppose we first measure A, given a state $|\psi\rangle$. We retrieve e.g. the eigenvalue a and the state is now projected into the eigenstate $|a\rangle$. Allowing for the presence of degenerate eigenstates, we actually have a superposition state $|\psi\rangle_{\text{Post-Meas}} = \sum_{i=1}^{d} c_i\,|a, b_i\rangle$, where d is the degree of degeneracy of the eigenvalue a. We then measure B obtaining one of the b_i, $\tilde{b}$. The state is thus projected into $\left|a, \tilde{b}\right\rangle$. If we now measure again A we will retrieve the same result as before: the two measurements of commuting observables A and B do not interfere.
Consider now non-commuting observables. As $AB|\psi\rangle \neq BA|\psi\rangle$ we cannot define a state $\left|a, \tilde{b}\right\rangle$ which is described by the (separate) eigenvectors of the two observables. Also, if we repeat the same 3 successive measurements as above, we obtain a different result. In particular, the second measurement of A does not in general retrieve the same eigenvalue as the first one.

? Question: Show why measurement of non-commuting observables are not compatible.
Given a state $|\psi\rangle$ we measure A, with result a. The state is now projected into the eigenstate $|a\rangle$ as before (we neglect here degeneracy). Now we rewrite this state in the basis of the operator B (which is not the same as the basis for A, so $|a\rangle \notin \{|b_i\rangle\}$): $|a\rangle = \sum_i c_i(a)\,|b_i\rangle$. When we measure B we will therefore obtain an eigenvalue b_i with probability $|c_i(a)|^2$, and the state is projected into: $\frac{P_i|a\rangle}{\sqrt{|\langle a|P_i|a\rangle|}} = \frac{|b_i\rangle\langle b_i|a\rangle}{|\,|a\rangle(|b_i\rangle\langle b_i|)\langle a|\,|^{1/2}} = |b_i\rangle$.
Again, this can be written as a non-trivial superposition of eigenstates of A: $|b_i\rangle = \sum_j c_j(b_i)\,|a_j\rangle$ so that it is now possible to obtain a measurement $a_j \neq a$ when we again measure A. □

$\mathcal{D}$: **Variance of an operator.** We define an operator $\Delta A = A - \langle A\rangle$ for any observable A. The expectation value of its square is the variance of A: $\langle\Delta A^2\rangle = \langle A^2\rangle - \langle A\rangle^2$.

■ Theorem: (**Uncertainty relation**). For any two observables, we have

$$\langle\Delta A^2\rangle\langle\Delta B^2\rangle \geq \frac{1}{4}|\,\langle[A, B]\rangle\,|^2$$

From Schwartz inequality ($|\langle\psi|\varphi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\varphi|\varphi\rangle$) we have $\left\langle \Delta A^2\right\rangle \left\langle \Delta B^2\right\rangle \geq |\left\langle \Delta A \Delta B\right\rangle|^2$. Now $\Delta A \Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}$ (where we defined the anticommutator $\{A, B\} = AB + BA$). Taking the expectation value (noting that $\langle[\Delta A, \Delta B]\rangle = \langle[A, B]\rangle$) we have

$$\langle \Delta A \Delta B\rangle = \frac{1}{2}\langle[A, B]\rangle + \frac{1}{2}\langle\{\Delta A, \Delta B\}\rangle .$$

Now we can show that $[A, B] = iC$ and $\{A, B\} = D$ where C and D are hermitian operators. Then the first term in the RHS is purely imaginary and the second purely real. Thus we have:

$$\left\langle \Delta A^2\right\rangle \left\langle \Delta B^2\right\rangle \geq |\left\langle \Delta A \Delta B\right\rangle|^2 = \frac{1}{4}|\left\langle[A, B]\right\rangle|^2 + \frac{1}{4}|\left\langle\{\Delta A, \Delta B\}\right\rangle|^2 \geq \frac{1}{4}|\left\langle[A, B]\right\rangle|^2.$$

□

3.3.3 Repeated measurements and Quantum Zeno Effect

A. Photon Polarization

In the same way an electromagnetic wave can be polarized, also individual photons possess a polarization. In particular they can be in a state of linear or circular polarization (the most general case, is called elliptical polarization). We consider a photon polarizer. This can be thought as a filter that ensures photons coming out of it are only of the right polarization.

— In-class demonstration with polarizer filters —

The photon polarizer (a polarization filter) is very similar to a measurement process and indeed it can be described by a projector. Let's assume that light can be described as either being in the horizontal $|h\rangle$ or vertical $|v\rangle$ polarization. Then , for an horizontal polarizer, for example, we have $P_H = |h\rangle\langle h|$. If we send a photon in the state $|\psi\rangle$ through this linear (horizontal) polarizer, its state after the polarizer will be $|h\rangle$. However the photon will emerge only with a probability $|\langle h|\psi\rangle|^2$. If we then send the photon to an orthogonal (vertical) polarizer $P_V = |v\rangle\langle v|$, the probability of a photon coming out is just zero. This situation is very similar to the case of repeated measurement. Thus the polarizer is a measurement process.
Now let's send an horizontally polarized photon ($|h\rangle$) through a 45 degrees polarizer. This polarizer can be described by the projector operator $P_{45} = |h + v\rangle\langle h + v|/2$. The state after the polarizer is then $(|h + v\rangle/\sqrt{2}$, and the probability of coming out is $\frac{1}{2}$. If now we send this photon through a $|v\rangle\langle v|$ polarizer, we obtain as a final state $|v\rangle$, and the total probability is $1/4$ (compare to zero before).
We can extend this even further. Assume we have a large number of polarizers each ensuring a polarization at a growing angle, each in a small step ϑ with the horizontal (that is, the first polarizer's angle is ϑ, the second 2ϑ etc.). The relevant projector is then

$$P_n(\vartheta) = (\cos(n\vartheta)|h\rangle + \sin(n\vartheta)|v\rangle)(\cos(n\vartheta)\langle h| + \sin(n\vartheta)\langle v|).$$

We start with a photon horizontally polarized $|\psi\rangle_0 = |h\rangle$. After the first polarizer, the photon emerges through in the state $|\psi\rangle_1 = \cos\vartheta|h\rangle + \sin\vartheta|v\rangle$ with probability $p_1(\vartheta) = |(\cos\vartheta\langle h| + \sin\vartheta\langle v|)|h\rangle|^2 = \cos^2\vartheta$. Now passing through the second polarizer the photon will emerge again with probability $\cos^2\vartheta$ and in the state $|\psi\rangle_1 = \cos(2\vartheta)|h\rangle + \sin(2\vartheta)|v\rangle$. After n polarizers, the state of the emerging photon is

$$|\psi\rangle_n = \cos(n\vartheta)|h\rangle + \sin(n\vartheta)|v\rangle .$$

Of course, we could get no photon at all, however the combined probability of getting a photon is $\cos(\vartheta)^{2n} \approx 1$ if the angle ϑ is small and the number of polarizer n is large. Thus we obtain an evolution of the system by using a measurement process.

B. Quantum Zeno effect

We consider a photon polarization rotator, whose action is to rotate the polarization about the propagation axis. By denoting $\{h, v\}$ the horizontal and vertical polarization, respectively, the polarization rotator achieves the following transformation:

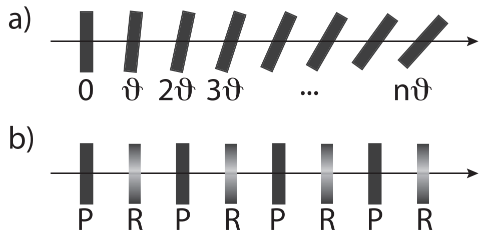

Fig. 1: a) Rotation by measurement. b) Quantum Zeno effect

$$|h\rangle \to \cos(\vartheta)|h\rangle + \sin(\vartheta)|v\rangle$$
$$|v\rangle \to \cos(\vartheta)|v\rangle - \sin(\vartheta)|h\rangle$$

This corresponds to the evolution matrix U:

$$U = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

? Question: What are the eigenstates of U?
By diagonalizing the matrix, we find the eigenvectors corresponding to right and left polarization:

$$R = (|h\rangle + i\,|v\rangle)/\sqrt{2}$$
$$L = (-i\,|h\rangle + |v\rangle)/\sqrt{2}$$

With eigenvalues $e^{i\vartheta}$ and $e^{-i\vartheta}$ respectively. The evolution given by the polarization rotator is unitary. □

Now assume another experiment in which we alternate a polarizer rotator and an horizontal polarizer. First consider just a set of polarizer rotators, each described by the formula above:

$$U(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

After n of these rotators, the photon is rotated to $U(\vartheta)^n\,|h\rangle = U(n\vartheta)\,|h\rangle = \cos(n\vartheta)\,|h\rangle + \sin(n\vartheta)\,|v\rangle$. Now if we alternate with the horizontal polarizer, every time the photon is transmitted with probability $\cos^2\vartheta$ and rotate back to $|h\rangle$. Again for ϑ small, the probability of a photon emerging tends to 1, and the final state of the photon is $|h\rangle$. This is a phenomenon called quantum Zeno effect[8] or we can call it a "watched milk never boils" phenomenon. The repeated measurements inhibit a (slow) evolution.

References

- B. Misra and E. C. G. Sudarshan, *The Zeno's paradox in quantum theory* J. Math. Phys. **18**, 756 (1977).
- W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, *Quantum Zeno effect*, Phys. Rev. A **41**, 2295 - 2300 (1990).
- Saverio Pascazio, Mikio Namiki, Gerald Badurek and Helmut Rauch *Quantum Zeno effect with neutron spin*, Physics Letters A, **179**, 155 - 160, (1993).

[8] Zeno's paradoxes are a set of problems (8 of which surviving) generally thought to have been devised by Zeno of Elea to support Parmenides's doctrine that "all is one" and that in particular, contrary to the evidence of our senses, motion is nothing but an illusion. The arrow paradox as related by Aristotle, (Physics VI:9, 239b5) states that "The third is ... that the flying arrow is at rest, which result follows from the assumption that time is composed of moments ... . he says that if everything when it occupies an equal space is at rest, and if that which is in locomotion is always in a now, the flying arrow is therefore motionless." To make the argument more similar to the QM version, we can rephrase it as: If you look at an arrow in flight, at an instant in time, it appears the same as a motionless arrow. Then how do we see motion?

22.51 Quantum Theory of Radiation Interactions
Fall 2012