On the Postulates of Quantum Mechanics 
and their Interpretation.

Marcel Nooijen
University of Waterloo

One arrives at very implausible theoretical conceptions, if one attempts to maintain the thesis that the statistical quantum theory is in principle capable of producing a complete description of an individual physical system. On the other hand, those difficulties of theoretical interpretation disappear, if one views the quantum mechanical description as the description of ensembles of systems. Albert Einstein (1949).

Quantum mechanics has given rise to a lot of discussion since its conception, as some things are so weird. I will try to clarify some of the issues here. One thing is very important from the outset. Quantum mechanics is a statistical theory. It tells us the various possible outcomes of experiments and the corresponding probabilities if we would do a large number of identical experiments on individual quantum systems. Identical experiments are necessarily idealizations, but this is not much of a restriction in practice, as many variables (e.g. what's is going on in Sidney or on the next bench in the lab) are irrelevant. In this context taking a spectrum of a sample in the gas phase appears to be a single experiment but it really amounts to doing measurements on many individual quantum systems. The systems are not all identical but this is the same type of fluctuation that occurs in classical statistical descriptions. At first sight the situation may not appear very different therefore from the description provided by classical statistical mechanics. In that case however, we have an underlying description (classical mechanics) that provides a complete (i.e. non-statistical) description of the world, which in general is far too complex, however, to be of use. Throwing a dice is a good example. Here we find the probabilities to obtain either of the outcomes 1, 2, 3, 4, 5, 6 equally.
likely (1/6). However, in principle it is possible to throw the dice always in precisely the same way such that it would yield a 6 say every time. It is unlikely that you would master this skill yourself but it appears well possible to build a device that would accomplish this. Classical mechanics says that in principle it can be done. The dice in the quantum world is different. It is *intrinsically impossible to predict the outcome of a single event* (unless the probability is unity). In fact this type of 'single event' experiment is not even the subject of quantum theory. We can only predict statistics and this requires that one does many identical experiments. Let me emphasize this aspect as it is easily to get confused on this issue. For example in every day language we might hear that there is a 70% chance that it will rain tomorrow. What does this mean? How can we verify that this 70% is a valid prediction? These kind of predictions only have a meaning when evaluated over many tomorrows.

Every description of an experiment on a microscopic system (even single molecule spectroscopy) is essentially statistical. Typically one performs an experiment on a sample consisting of similar microscopic systems. In an idealized theoretical description we view such an experiment as equivalent to performing a sequence of measurements on each (now supposedly identical) microscopic system in isolation. This generates a definite result for each individual experiment, and *the statistics of the distribution of results is described by quantum mechanics*. This theoretically predicted distribution of results may not quite agree with the experimental result for a variety of reasons. The actual experimental sample will contain a distribution of different microsystems; it will involve some interaction between different microsystems or between microsystems and the environment, and so forth. Some of these effects can be taken into account by using statistical mechanics. This is beyond where I want to go, however. Let us simply assume that the quantum description would agree very well with the experimental result, and analyze what it says and what it means.

The above abstraction of an experiment shows that quantum mechanics describes the statistical outcome of an experiment performed on an *ensemble* of identical
microsystems. *The wavefunction in quantum mechanics describes the properties of the ensemble.* This is very different from classical theory, where the laws of physics describe individual systems, and we can imagine how they relate to ‘reality’. We cannot claim that this is true in quantum mechanics. The wavefunction truely relates to an ensemble of identical sytems. In most textbooks on quantum mechanics the wavefunction is said to describe the 'system'. This is slightly inaccurate and it is analogous to the fact that the weather forecast does not describe tomorrow but an ‘average of possible tomorrows’ given the current situation in the atmosphere. The reason that I emphasize this seemingly minor point is that it helps to explain many of the ‘strange’ results of quantum mechanics. As long as we strictly adhere to the so-called ensemble interpretation of quantum mechanics we seldom run into difficulties. It is unsatisfactory perhaps that this is all we (can?) know about the microscopic world, but this is the (current) state of affairs.

Below I will discuss the postulates of quantum mechanics. We will phrase these postulates using just the $x$-coordinate of a single particle. The generalization to many-particles in 3d space is straightforward, but leads to less transparent equations unless some new notation would be introduced. Also we will discuss everything in the so-called coordinate representation. A more general formulation is not needed for our purposes.

1. **Measurements in Quantum Mechanics.**

In McQuarrie & Simon math chapter 2 you find a discussion on the statistical analysis of a repeated measurement of the quantity $A$ on a given ensemble. Let us assume that if we measure the quantity $A$, we have possible outcomes $a_1, a_2, \ldots, a_i$. The probability $P(a_i) = P_i$ to find the value $a_i$ for our given ensemble is $N_i / N_{tot}$, where $N_i$ is the number of times we obtain $a_i$ out of a total number of measurements $N_{tot}$. In the limit of a very large set of measurements these probabilities will converge to definite numbers $0 \leq P(a_i) \leq 1$. This leads to the purely classical definitions of the average value $\bar{A}$ or $\langle A \rangle$

$$\bar{A} = P_1 a_1 + P_2 a_2 + \ldots + \sum_k P_k a_k$$

(4.1)

and the standard deviation $\sigma(A)$, or similarly its square, the so-called variance:
The variance can be written alternatively as
\[
\sum_k P_k (a_k - \bar{A})^2 = \sum_k P_k (a_k^2 - 2a_k \bar{A} + (\bar{A})^2)
\]
\[
= \sum_k P_k a_k^2 - 2\langle A \rangle \sum_k P_k a_k + \langle A \rangle^2 \sum_k P_k
\]
\[
= \langle A^2 \rangle - \langle A \rangle^2 \geq 0
\]
and this is the form that turns out to be very convenient in QM.

In Quantum Mechanics there are very strict rules on what is a measurable quantity or a so-called ‘observable’. An observable classical quantity \( A \) corresponds to a linear \textit{Hermitean operator} \( \hat{A} \) in quantum mechanics (Operators carry hats to distinguish them from numbers. They act on functions.) Classical quantities are functions of the basic variables, the positions and momenta of the particles (energy, angular momentum are some examples). This allows us to make the corresponding operators in QM: Replace \( x, y, z \) by the multiplication operators \( \hat{X}, \hat{Y}, \hat{Z} \), and replace the momenta by differential operators: \( p_x \rightarrow \hat{P}_x = -i\hbar \frac{\partial}{\partial x} \) and so forth. This allows you to construct most operators that occur in QM. In addition there are three operators with no classical analog. They are the spin operators \( \hat{S}_x, \hat{S}_y \) and \( \hat{S}_z \). These will be discussed later on.

An operator \( \hat{A} \) is linear if for any linear combination of functions (in its domain)
\[
\hat{A}(c_1 f(x) + c_2 g(x)) = c_1 \hat{A}(f(x)) + c_2 \hat{A}(g(x))
\]
while it is called Hermitean if for any two functions (in its domain)
\[
\int_{-\infty}^{\infty} g^*(x) \hat{A}(f(x)) dx = \int_{-\infty}^{\infty} f(x)(\hat{A}g(x))^* dx = \int_{-\infty}^{\infty} (\hat{A}g(x))^* f(x) dx
\]
problem (This is again tied to the integration. Think of the particle in the box, or a particle on the ring). Such functions that satisfy the boundary conditions are in the so-called domain of the operator. Hermitean operators have some very important properties that make them suitable for the representation of observables which play a crucial role in QM. In particular they have real eigenvalues \( a_k \), and their eigenfunctions can always be chosen to be orthonormal:

Hermitean operator \( \hat{A} \): orthonormal eigenfunctions \( \phi_k(x) \), real eigenvalues \( a_k \):

\[
\hat{A} \phi_k(x) = a_k \phi_k(x)
\]  \hspace{1cm} (4.6)

\[
\int_{-\infty}^{\infty} \phi_k^*(x) \phi_l(x) dx = \delta_{kl}
\]  \hspace{1cm} (4.7)

where \( \delta_{kl} = 1, k = l \); \( \delta_{kl} = 0, k \neq l \); is the the Kronecker delta symbol.

Moreover, any function \( \Psi(x) \) (that satisfies the boundary conditions) can be expressed as a linear combination of the eigenfunctions of a Hermitean operator \( \hat{A} \) (this is the the so-called completeness property)

\[
\Psi(x) = c_1 \phi_1(x) + c_2 \phi_2(x) + \ldots
\]  \hspace{1cm} (4.8)

where the expansion coefficients are given by

\[
c_k = \int_{-\infty}^{\infty} \phi_k^*(x) \Psi(x) dx
\]  \hspace{1cm} (4.9)

**Proof that eigenvalues of Hermitian operators are real:**

\[
\int_{-\infty}^{\infty} \phi_k^*(x) (\hat{A} \phi_k(x)) dx = a_k \int_{-\infty}^{\infty} \phi_k^*(x) \phi_k(x) dx
\]

\[
= \int_{-\infty}^{\infty} (\hat{A} \phi_k(x))^* \phi_k(x) dx = \int_{-\infty}^{\infty} (a_k \phi_k(x))^* \phi_k(x) dx = a_k^* \int_{-\infty}^{\infty} \phi_k^*(x) \phi_k(x) dx
\]

or

\[
a_k = a_k^*, \text{ i.e. } a_k \text{ is real.}
\]
Proof that eigenfunctions corresponding to different eigenvalues are orthogonal:

\[
\int_{-\infty}^{\infty} \phi_k^*(x)(\hat{A}\phi_l(x))dx = a_l \int_{-\infty}^{\infty} \phi_k^*(x)\phi_l(x)dx
\]

\[
= \int_{-\infty}^{\infty} (\hat{A}\phi_k(x))^*\phi_l(x)dx = \int_{-\infty}^{\infty} (a_k \phi_k(x))^*\phi_l(x)dx = a_k \int_{-\infty}^{\infty} \phi_k^*(x)\phi_l(x)dx
\]

using that the eigenvalues are real. Therefore

\[
(a_i - a_k) \int_{-\infty}^{\infty} \phi_i^*(x)\phi_l(x)dx = 0 \rightarrow \int_{-\infty}^{\infty} \phi_i^*(x)\phi_l(x)dx = 0 \text{ if } a_i \neq a_k
\]

If two or more eigenfunctions have the same eigenvalue, then linear combinations are also eigenfunctions, and we can choose appropriate linear combinations such that these functions are orthonormal. Following this construction the complete set of eigenfunctions can be chosen to form an orthonormal set.

These mathematical definitions and properties lead to a very precise description of measurements on an ensemble of microscopic systems. The statements below are a concise form of the postulates of QM related to measurements. The ensemble in QM is described by a normalized wave function \( \Psi(x) : \int_{-\infty}^{\infty} \Psi(x)^*\Psi(x)dx = 1 \), and the results of measurement of a quantity \( A \) for the entire ensemble can be described as follows:

- Only eigenvalues \( a_k \) can be obtained, one of them for ‘each measurement on an individual system in the ensemble’. Collecting the data on the individual measurements yields the statistical information.

- Each possible value \( a_k \) shows up with probability \( P_k = |c_k|^2 = c_k^*c_k \), where the (in general complex) coefficient \( c_k \) is defined in Eqn. (4.9).

- After measuring \( A \) the initial ensemble is split up into subensembles, one corresponding to each of the possible eigenvalues \( a_k \). The wave function that describes the subensemble corresponding to \( a_k \) is the corresponding (normalized) eigenstate \( \phi_k(x) \).

In the above description we have assumed that each of the eigenvalues of the operator \( \hat{A} \) is non-degenerate. If there is more than one independent eigenfunction, say \( \phi_{k_1}(x), \phi_{k_2}(x) \)
corresponding to the same eigenvalue \(a_k\) the rules are slightly more complicated. It adds little to our present discussion however, and I will not pursue this side track.

=============== Average values, expectation values and variances in QM.===============

The above is a fairly precise description of the information content of QM. In particular we can obtain information on the averages and variances discussed above:

\[
\overline{A} = \langle A \rangle = \langle \hat{A} \rangle \equiv \int_{-\infty}^{\infty} \Psi^*(x) \hat{A} \Psi(x) dx \tag{4.10}
\]

This average quantity is also called the expectation value. For the variance we have

\[
\sigma^2(A) = \langle A^2 \rangle - \langle A \rangle^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle = \int_{-\infty}^{\infty} \Psi^*(x) (\hat{A} - \langle \hat{A} \rangle)^2 \Psi(x) dx \tag{4.11}
\]

These relations are not extra postulates. They follow from the previous statements and their classical definitions as can be verified by expanding \(\Psi(x)\) in terms of eigenfunctions of \(\hat{A}\). e.g.

\[
\langle \hat{A} \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \hat{A} \Psi(x) dx = \int_{-\infty}^{\infty} \sum_k c_k^* \phi_k^*(x) \hat{A} \sum_l c_l \phi_l(x) dx = \sum_k c_k^* \sum_l c_l \int_{-\infty}^{\infty} \phi_k^*(x) \hat{A} \phi_l(x) dx = \sum_k c_k^* \sum_l c_l P_{lk} = \sum_k c_k^* c_k a_k = \langle a_k \rangle \tag{4.12}
\]

There are important general aspects of this proof that you are expected to know. In the first step we substitute the expansion for \(\Psi(x)\). Please note that we use a different name for the summation indices to represent \(\Psi(x)\) and its complex conjugate. This is completely general technique: never use the same summation label twice. In the second step we use linearity, both of the operator \(\hat{A}\) and of performing integration. In going from the third to the fourth line we use that the expansion functions are eigenfunctions of \(\hat{A}\), and finally we use that the eigenfunctions are orthonormal. The integral in the fourth line is non-zero (unity in fact) only if \(l = k\), which gets rid of the sum over \(l\).
Similarly the QM variance can be shown to reduce to \( \sigma^2(A) = \sum_k P_k(a_k - \bar{A})^2 \), which is the classical expression. The variance vanishes if and only if only one term in the sum contributes, in which case the average equals this particular eigenvalue \( a_i \) say. It follows therefore that the variance vanishes only if the wavefunction \( \Psi(x) \) that describes the ensemble is an eigenfunction of \( \hat{A} \). One of the postulates was that the subensemble obtained by measuring an eigenvalue \( a_i \) was described by the eigenfunction \( \phi_i(x) \). Here it is seen that if we measure \( A \) again on this particular subensemble we find the eigenvalue \( a_i \) with absolute certainty. This is what we would expect classically. Measuring a quantity is therefore a good way to prepare an ensemble of identical microsystems(!), as the subensemble corresponding to a particular eigenvalue is collectively described by the same wavefunction after the measurement.

For those of you who are familiar with linear algebra it may be illustrative to point out the analogy of the orthonormal eigenfunctions of an operator \( \hat{A} \) with an orthonormal basis for a vector space. The functions \( \phi_1(x), \phi_2(x), \ldots \) can be thought to form orthonormal basis vectors (coordinate system, or axes). \( \Psi(x) \) represents an arbitrary vector in the space, with coordinates \( (c_1, c_2, \ldots) \) with respect to this basis. The inner product between two vectors (needed to define the vector space) is defined as:

\[
\Psi_a(x) = a_1\phi_1(x) + a_2\phi_2(x) + \ldots \\
\Psi_b(x) = b_1\phi_1(x) + b_2\phi_2(x) + \ldots
\]

\[
\Psi_a \cdot \Psi_b = \int_{-\infty}^{\infty} \Psi_a^*(x) \Psi_b(x) dx = a_1^*b_1 + a_2^*b_2 + \ldots
\]

The act of measurement is then described as decomposing the full ensemble corresponding to eigenfunction \( \Psi(x) = \sum_k c_k \phi_k(x) \) into subensembles corresponding to eigenvalues \( a_i \) that are each described by a particular basis vector \( \phi_i(x) \). The probability to find eigenvalue \( a_i \) equals \( |c_i|^2 \). It is not correct to think of the ensemble to consist of subensembles in the first place (this many particles in state \( \phi_1(x) \), that many in state \( \phi_2(x) \), etc.). This would give the wrong result whenever so-called interference effects play a role (see below). Using this vector analogy different measurements \( A, B \)
correspond to different sets of coordinate axes (different eigenfunctions in general), and a
different decomposition of the wavefunction $\Psi(x)$ that describes the ensemble. The
vector analog is an intuitive (and rigorously valid) picture that may help you to
understand this somewhat abstract chapter.

----- Commutators and the uncertainty principle.

Above we discussed the variance $\sigma_A^2$ of measuring the quantity $A$ corresponding to an
operator $\hat{A}$. The variance depends on the ensemble (of course), hence on the
wavefunction $\Psi(x)$, and we found that $\sigma_A = 0$ if and only if the wave function $\Psi(x)$ that
describes the ensemble is actually an eigenfunction of $\hat{A}$. The easiest way to prepare an
ensemble in such a state is by performing the measurement (as discussed). Now we can
measure a quantity $B$ after measuring $A$ and ask the question: "Can we prepare an
ensemble such that both $\sigma_A$ and $\sigma_B$ are zero?" This would mean that both the quantity $A$
and $B$ are uniquely specified for each element of the ensemble. Following the postulates
this would be the case if the wavefunction $\Psi(x)$ is an eigenfunction of both the operators
$\hat{A}$ and $\hat{B}$. A more elaborate question would be "Can we create an ensemble that has sharp
values for $A$ and $B$ for all possible outcomes of measuring $A$ then $B$?". This would
imply that there is a function for each possible pair of eigenvalues $(a_i, b_j)$ that is an
eigenfunction of both $\hat{A}$ and $\hat{B}$. The operators $\hat{A}$ and $\hat{B}$ in such a case should have a
complete set of common eigenfunctions

$$\phi_{a_i,b_i}(x), \phi_{a_j,b_j}(x), \ldots$$

where

$$\hat{A}\phi_{a_i,b_j}(x) = a_i \phi_{a_i,b_j}(x)$$
$$\hat{B}\phi_{a_i,b_j}(x) = b_j \phi_{a_i,b_j}(x)$$

(4.14)

From this it follows that

$$\hat{A}\hat{B}\phi_{a_i,b_j}(x) = \hat{B}\hat{A}\phi_{a_i,b_j}(x) = a_i b_j \phi_{a_i,b_j}(x)$$

and since any $\Psi(x)$ can be expanded in this set of eigenfunctions (the completeness
property) it follows that
\[ \hat{A}\hat{B}\Psi(x) = \hat{B}\hat{A}\Psi(x) \text{ for any } \Psi(x) \]  
(4.15)

or

\[ [\hat{A}, \hat{B}] = 0 \]  
(4.16)

Here we have shown that if it is in principle possible to create an ensemble that has sharp values for \( A \) and \( B \) for any possible pair of values of these observables only if \( \hat{A} \) and \( \hat{B} \) commute. Moreover it is clear that the eigenvalues come in specific pairs, as they have to correspond to specific common eigenfunctions. How can we prepare such an ensemble?

Simple. Measure a quantity \( A \), take the subensemble that corresponds to eigenvalue \( a_i \) and go on to measure the quantity \( B \). If we now pick the subensemble corresponding to \( b_j \) it would be described by a common eigenfunction of \( \hat{A} \) and \( \hat{B} \), \( \phi_{a_i,b_j}(x) \). We can reverse the order of the measurements to create the specific ensemble, and we can do it for any pair of commensurate eigenvalues \( a_i \) and \( b_j \). This also works if \( \hat{A} \) and \( \hat{B} \) have degenerate eigenvalues, although I will not venture a derivation of this result here.

Finally the converse is also true: if \([\hat{A}, \hat{B}] = 0 \) then \( A \) and \( B \) can be measured to indefinite precision. This can be more accurately phrased as: It is possible to prepare specific ensembles that will yield definite results for the (sequential) measurement of \( A \) and \( B \).

From the postulates it also transpires what happens if we subsequently measure observables \( A \) and \( B \) whose corresponding operators \( \hat{A} \) and \( \hat{B} \) do not have common eigenfunctions. If we measure \( \hat{A} \) and continue with the subensemble corresponding to eigenvalue \( a_i \), which is described by the eigenfunction \( \phi_i(x) \) we can continue to measure \( \hat{B} \). However by assumption \( \phi_i(x) \) is not an eigenfunction of \( \hat{B} \) and we will therefore not get a sharp value for \( B \). There is necessarily a spread. After measuring \( \hat{B} \) and continuing with the subensemble corresponding to \( b_j \) described by eigenfunction \( \psi_j(x) \) (eigenfunction of \( \hat{B} \) but not \( \hat{A} \)) we could decide to measure \( \hat{A} \) again. However, the new wavefunction \( \psi_j(x) \) is not an eigenfunction of \( \hat{A} \) and we would get a distribution of values for the observable \( A \). This means that measuring \( B \) has destroyed the sharp value for \( A \) we created by the very act of measuring \( A \). If the operators \( \hat{A} \) and \( \hat{B} \) do not
commute in general the measurement of $\hat{B}$ destroys the information we obtained by measuring $\hat{A}$. This is the content of the uncertainty principle. I might add that it is possible that $\hat{A}$ and $\hat{B}$ do not commute, but still have *some* eigenfunctions in common. In such a case the corresponding eigenvalues can be specified sharply and are not disturbed by subsequent measurements of $\hat{A}$ or $\hat{B}$. Let me also note that the notion of *simultaneous* measurement does not enter the discussion. Using our ensemble interpretation the quest is for an ensemble that would yield a sharp value for the observables upon measurement. Be it in sequence or simultaneous or in combination. None of this matters if $\hat{A}$ and $\hat{B}$ commute. (This is strictly true only if $\hat{A}$ and $\hat{B}$ also commute with the Hamiltonian, see below).

For your information I will quote the general result about the product of standard deviations for operators that do not commute. The minimum spread depends on the initial wavefunction that describes the ensemble and the general result reads

$$\Delta A \Delta B \geq \frac{1}{2} \left| \left[ \hat{A}, \hat{B} \right] \right|$$

(for given state $\Psi(x)$) \hspace{1cm} (4.17)

For the quantum afficionados in the class the proof follows by defining

$$\phi_{\lambda}(x) = \{ (\hat{A} - \langle \hat{A} \rangle) + \lambda (\hat{B} - \langle \hat{B} \rangle) \} \Psi(x);$$

\hspace{1cm} (4.18)

Then using that $\int_{-\infty}^{\infty} \phi_{\lambda}^*(x) \phi_{\lambda}(x) dx \geq 0 \ \forall \ \lambda$ (this depends on the Hermiticity of $\hat{A}$ and $\hat{B}$) you can derive the general result (4.17). Give it a try!

Finally we have the special case: $[\hat{A}, \hat{B}] = c$ (a constant), which implies $\Delta A \Delta B \geq \frac{1}{2} |c| :$

This leads to a lower bound on the precision of measuring $A$ and $B$ that is *independent of the initial ensemble (wave function)*. Famous example: $\Delta x \Delta p_x \geq \frac{1}{2} |\hbar| = \frac{\hbar}{2}$, Heisenbergs uncertainty principle.
Some tricks when evaluating Commutators:

\[(\hat{A}\hat{B} - \hat{B}\hat{A})\Psi(x,\ldots) = i\hbar \hat{C}\Psi(x,\ldots), \text{ for any } \Psi(x) \quad \rightarrow \quad [\hat{A}, \hat{B}] = \hat{C}\]

examples \([\hat{p}_x, \hat{p}_y] = 0; \quad [\hat{x}, \hat{y}] = 0; \quad [\hat{x}, \hat{p}_y] = 0; \quad [\hat{x}, \hat{p}_z] = i\hbar\]

\[
\begin{align*}
[\hat{A}, \hat{B} + \hat{C}] &= [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]; \\
[\hat{A}, \hat{B}\hat{C}] &= [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]
\end{align*}
\]

e.g. \([\hat{x}, \hat{\rho}_z] = [\hat{x}, \hat{p}_z] + \hat{p}_z[\hat{x}, \hat{\rho}_z] = 2i\hbar\hat{\rho}_z\]
2. The Time-Dependent Schrödinger equation.

Another important postulate in quantum mechanics concerns the time-dependence of the wave function. This is governed by the time-dependent Schrödinger equation

\[ i\hbar \frac{\partial \Psi(x,t)}{\partial t} = \hat{H}\Psi(x,t) \]  

(4.19)

where \( \hat{H} \) is the Hamiltonian operator of the system (the operator corresponding to the classical expression for the energy). This is a first order differential equation in \( t \), which means that if we specify the wave function at an initial time \( t_0 \), the wave function is determined at all later times. Let me emphasize that this means the wavefunction has to be specified for \( all \) \( x \) at initial time \( t_0 \). These initial conditions are familiar from wave equations as discussed in MS Chapter 2. In classical physics we often deal with second order differential equations and in addition the time derivative \( \frac{\partial \Psi(x,t)}{\partial t} / \partial t \) would then need to be specified for all \( x \). Let me emphasize here that although the experimental results that can be predicted from QM are statistical in nature, the Schrödinger equation that determines the wave function as a function of time is completely deterministic.

=*=*=*=*=* Special solutions: Stationary States (only if \( \hat{H} \) is time-independent).

If we assume that the wave function can be written as a product: \( \Psi(x,t) = \phi(x)\gamma(t) \) we can separate the time dependence from the spatial dependence of the wave function in the usual way. The separation constant is called \( E \) and will turn out to be the energy of the system for such solutions

\[ \rightarrow \ldots i\hbar \frac{d\gamma(t)}{dt} = E\gamma(t) \rightarrow \gamma(t) = e^{-iE(t-t_0)/\hbar} \]  

(4.20)

\[ \hat{H}\phi(x) = E\phi(x) \rightarrow \hat{H}\phi_n(x) = E_n\phi_n(x) \]  

(4.21)

equation (4.21) is called the time-independent Schrödinger equation and plays a central role in all of chemistry. Since the operator \( \hat{H} \) is Hermitean, the eigenfunctions form a complete and (can be chosen to be an ) orthonormal set of functions. Using these
eigenfunctions of $\hat{H}$ special solutions to the time-dependent Schrödinger equation can be expressed as

$$\Psi(x,t) = \phi_n(x)e^{-iE_n(t-t_0)/\hbar}; \quad \Psi(x,t_0) = \phi_n(x)$$

For these special solutions of the Schrödinger equation, all measurable properties are independent of time. For this reason they are called stationary states. For example the probability distribution

$$|\Psi(x,t)|^2 = |\Psi(x,t_0)|^2 = |\phi_n(x)|^2,$$

but also

$$\langle \hat{A} \rangle_t = \langle \hat{A} \rangle_{t_0} \quad \forall \hat{A},$$

as is easily verified, by substituting the product form of the wave function. Also the probabilities to measure an eigenvalue $a_k$ are independent of time, as seen below

$$\hat{A}\phi_k(x) = a_k\phi_k(x) \rightarrow P_k(t) = \left| \int_{-\infty}^{\infty} \phi_k^*(x)\Psi(x,t) \right|^2 = P_k(t_0)$$

The common element in each of these proofs is that the time-dependent phase factor cancels because we have both $\Psi(x,t)$ and $\Psi^*(x,t)$ in each expression. Let me note that the stationary solutions are determined by the initial condition. If you start off with a stationary state at $t_0$, the wave function is a stationary state for all time.

The general solution of the time-dependent Schrödinger equation (TDSE) can be written as a time-dependent linear combination of stationary states. If we assume that the initial state is given by

$$\Psi(x,t_0) = c_1\phi_1(x) + c_2\phi_2(x) + \ldots.$$  

(it can always be written in this fashion as the eigenfunctions of $\hat{H}$ form a complete set), then it is easily verified that

$$\Psi(x,t) = c_1e^{-iE_1(t-t_0)/\hbar}\phi_1(x) + c_2e^{-iE_2(t-t_0)/\hbar}\phi_2(x) + \ldots$$

satisfies the TDSE and the initial condition. For this general linear combination of eigenstates of $\hat{H}$ (the general case) properties do depend on time. This is true for expectation values and probabilities, and is due to the fact that different ‘components’ in the wave function oscillate with different time factors. In calculating expectation values
we get cross terms and the time-dependent phase factors do not cancel out. I will discuss a specific example of this in class, and it is part of a homework problem. The material will be presented in the solution manual.

Independent of the initial wave function, energy is always conserved (as would be expected from classical physics), and also probabilities to measure a particular energy $E_k$:

$$\left| \int_{-\infty}^{\infty} \phi^*_k(x) \Psi(x, t) dx \right|^2 = |c_k e^{-iE_k (t - t_0)}|^2 = |c_k|^2$$

Further Remarks:

- All depends on the initial wave function, which is arbitrary in principle. The most common way is to specify it by means of a measurement!
- Stationary states: the initial state is defined to be an eigenfunction of $\hat{H}$. In this case nothing moves except the phase of the wave function. In our current version of QM we would find infinite lifetimes of excited states! (This is because the e.m. field is missing from our treatment and we have assumed a time-independent $\hat{H}$)
- In general properties oscillate in time (e.g. electron density) $\rightarrow$ radiation!? Again there is a need to include e.m. field. In the real world, systems do not satisfy our time-dependent Schrödinger equation indefinitely. The system interacts with the electromagnetic field, and in this way makes a transition to the ground state (eventually). This occurs even if no radiation field is present (spontaneous emission). These are the reasons that stationary states and in particular the ground state is so important.