Formalism of quantum mechanics

After some exposure to basic wave mechanics in 1D let us recast and systematize the knowledge we have accumulated in a more general form.

Like any other theory, the quantum theory essentially deals with two kinds of things: objects and actions on those objects. In the case of quantum mechanics the objects are wave functions (or state vectors) and the actions are defined by operators. Since the the fundamental equations that lie at the heart of quantum mechanics (e.g. the Schrödinger equation) are linear, the natural language of it is linear algebra.

Any fundamental theory is based on a set of postulates. A postulate is a basic principle or property that is not derived from something.

The foundation of quantum mechanis consists of several postulates. They can be cast in somewhat different forms, yet the essence remains the same. How do we know if those postulates are correct and consistent? The proof comes from the experiment and so far quantum mechanics has been found to be capable describe the physical phenomena very well, at least in the domain where it is supposed to be applicable.

Now let us list those postulates of quantum mechanics:

Postulate 1 The dynamical state of a quantum-mechanical system is completely described by its wave function.

The wave function can be complex and it contains all that can possibly be known about the system. In order to be physically admissible the wave function must satisfy several conditions:

a) It must be normalizable
b) It must be single valued
c) It must be continuous.

These are follow naturally from the physical interpretation of the wave function. Condition a) is essentially equivalent to stating that the probability of finding a quantum particle anywhere must be 1. Condition b) ensures that the probabilitie and expectation values can be determined unambiguosly. Finally, Condition c) follows from the Schrödinger equation, which we know is a second order differential equation. If the wave function were not continuous then the expectation values that involve the second derivative (e.g. energy = $\langle H \rangle$) could not be made well defined (finite).

Postulate 2 For every measurable property of the system there exists a corresponding operator

The operators corresponding to most common observables are listed below:

observable	operator	symbol
Position	x	$\hat{x}$
Momentum	$-i\hbar\frac{\partial}{\partial x}$	$\hat{p}_x$
Kinetic energy	$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$, $\frac{p_x^2}{2m}$	$\hat{T}$, $\frac{\hat{p}_x^2}{2m}$
Potential energy	$V(x)$	$\hat{V}(x)$
Total energy	$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$	$\hat{H}$
Angular momentum	$-i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$	$\hat{L}_x$
	$-i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$	$\hat{L}_y$
	$-i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$	$\hat{L}_z$

Physically observed quantities are real (i.e. not complex). This puts a restriction on the operators that are associated with the observable quantities. The operators must be Hermitian (we will give a definition of a Hermitian operator a bit later)

The action of operators is generally order-dependent:

$$\hat{A}\hat{B}f(x) \neq \hat{B}\hat{A}f(x)$$

When written in a sequence, operators are assumed to operate from left to right, i.e. in $\hat{A}\hat{B}f(x)$, $\hat{B}$ is applied first, then $\hat{A}$ is applied.

Postulate 3 For any measurement involving an observable that corresponds to an operator, the only values that

will be measured are the eigenvalues of the operator.

The eigenvalues of an operator are defined as:

$$\hat{A}\,\varphi_i(x) = a_i\,\varphi_i(x) \qquad a_i \text{ is an eigenvalue}$$

The number of eigenvalues and eigenfunctions/eigenvectors depends on the nature of the operator.

The eigenvalues of Hermitian operators are known to be always real (there is no exception to that).

Postulate 4 If a system is in a state described by a wave function and the value of an observable corresponding to operator $\hat{A}$ is measured just once on many identically prepared systems, the average value of all measurements will be

$$\langle a \rangle = \frac{\int \psi^*(x)\,\hat{A}\,\psi(x)\,dx}{\int \psi^*(x)\,\psi(x)\,dx} = \int \psi^* A \psi\,dx \quad \text{(assuming } \psi \text{ is normalized)}$$

If the wave function happens to be an eigenfunction of $\hat{A}$ then

$$\langle a \rangle = \int \psi^* A \psi\,dx = \int \varphi_j^* A\,\varphi_j\,dx = a_j \int \varphi_j^* \varphi_j\,dx = a_j$$

The fourth postulate tells us what will happen when a large number of identical systems are interrogated one time.

What happens in an individual measurement? We will get a value equal to one of the eigenvalues. Each eigenvalue may occur with a different frequency (probability) and this frequency depends on the wave

function. However, there is no way to know in advance which outcome (eigenvalue) will be produced in a particular measurement, unless the wave function is equal to some eigenfunction (with no admixture of other eigenstates). After a large number of measurements the result will converge to $\langle a \rangle$.

In the Copenhagen interpretation of quantum mechanics (which is the most often used one) the act of the measurement is believed to cause the system to "collapse" into a single eigenstate of the operator corresponding to the measured quantity. In the absence of an external perturbation it will remain in that eigenstate.

Postulate 5 The time evolution of the wave function is described by the Schrödinger equation:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$$

For the particular case of a time-independent potential ($V = V(x)$) the above equation allows a separation of variables, which results in the time-independent (stationary) Schrödinger equation:

$$H\psi_n(x) = E_n \psi_n(x)$$

The n-th particular solution is then

$$\Psi_n(x,t) = \psi_n(x)\, e^{-\frac{iE_n t}{\hbar}}$$

The sufficient condition for this expansion to be valid for a general wave function $\Psi(x,t)$ is that set of

functions $\psi_n(x)$ form a complete set of basis functions.

Hilbert space It is a generalization of the notion of Euclidean space to the case of any (finite or infinite) number of dimensions. In Hilbert space we define

a) vectors (points in space)

b) the addition operators on vectors

c) the inner product (scalar product)

Let us consider illustrative examples:

$$|\alpha\rangle \longrightarrow \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \qquad |\beta\rangle \longrightarrow \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

N can be infinite!

Inner product of vectors $|\alpha\rangle$ and $|\beta\rangle$:

$$\langle\alpha|\beta\rangle = a_1^* b_1 + a_2^* b_2 + \ldots + a_N^* b_N$$

Now compare it to the scalar product of two functions f and g:

$$\langle f|g\rangle = \int f^*(x)\, g(x)\, dx$$

If in some basis $\{\phi_i\}$ we can represent f and g as

$$f(x) = \sum_i a_i \phi_i \qquad g(x) = \sum_j b_j \phi_j$$

then assuming that the basis is orthonormal ($\int \phi_i^* \phi_j dx = \delta_{ij}$) we get:

$$\langle f|g\rangle = \int \left(\sum_i a_i^* \phi_i^*\right)\left(\sum_j b_j \phi_j\right) dx = \sum_{ij} a_i^* b_j \underbrace{\int \phi_i^* \phi_j \, dx}_{\delta_{ij}} =$$

$$= \sum_{ij} a_i^* b_j \delta_{ij} = \sum_i{}' a_i^* b_i = a_1^* b_1 + a_2^* b_2 + \ldots a_N^* b_N$$

Normalizable (square integrable) wave functions (state vectors) are possible only when $\sum_i a_i^* b_i$ is finite. When N is infinite the series $\sum_{i=1}^{\infty}{}' a_i^* b_i$ must be converging.

Linear operators can always be represented by matrices (which may be infinitely large):

$$|\beta\rangle = M|\alpha\rangle = \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1N} \\ m_{21} & m_{22} & \ldots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & \ldots & m_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

Note that for $N = \infty$ the elements of matrix M do not have to converge to some limit when indicies i and j in m_{ij} increase, not to mention the convergence to zero, while the components a_j must converge to 0 when $j \to \infty$.

The elements of matrices representing operators may be complex. However, since the eigenvalues of the matrices that correspond to observables must be real it puts a constraint on those matrices. They must be Hermitian.

A matrix is Hermitian (or as they often say self-adjoint) if $M_{ij} = M_{ji}^*$

For a Hermitian matrix

$$\langle\alpha|M|\alpha\rangle = \vec{a}^{H} M \vec{a} \quad \text{is real for any } |\alpha\rangle$$

Here the subscript H stands for hermite conjugation of a vector:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \qquad \vec{a}^{H} = (a_1^*, a_2^*, \ldots, a_N^*)$$

The usual vector-matrix-vector multiplication rule reads

$$\langle\alpha|M|\alpha\rangle = \sum_{ij}{}' a_i^* M_{ij} a_j$$

Compare it to the scalar product of ψ and $\hat{M}\psi$:

$$\int \psi^*(x)\, \hat{M}\psi(x)\, dx = \int \left(\sum_i a_i^* \phi_i^*\right) \hat{M} \left(\sum_j a_j \phi_j\right) dx = \sum_{ij}{}' a_i^* a_j \int \phi_i^* M \phi_j dx$$

$$= \sum_{ij}{}' a_i^* a_j M_{ij} \qquad \text{where} \quad M_{ij} \equiv \int \phi_i^* \hat{M} \phi_j \, dx$$

The eigenvalues of Hermitian matrices are known to be real, while their eigenvectors corresponding to distinct eigenvalues are mutually orthogonal, i.e.

if $M|\mu_i\rangle = \lambda_i |\mu_i\rangle$ then $\langle\mu_i|\mu_j\rangle = \delta_{ij}$ (if $\lambda_i \neq \lambda_j$)

The eigenvectors corresponding to the same (degenerate) eigenvalue can always be rearranged into linear combinations (which will also be eigenvectors) and made orthogonal

Now an operator $\hat{Q}$ is called Hermitian if

$$\langle f|\hat{Q}g\rangle = \langle \hat{Q}f|g\rangle$$

A related concept is the adjoint operator (or Hermitian conjugate operator):

$$\langle f|\hat{A}g\rangle = \langle \hat{A}^{+}f|g\rangle \quad \Rightarrow \quad \hat{A}^{+} \text{ is adjoint of } \hat{A}$$

If $\hat{A}^{+} = \hat{A}$ then $\hat{A}$ is Hermitian (or self-adjoint)

Example: the momentum operator $\hat{p} = -i\hbar\frac{d}{dx}$

$$\langle f|\hat{p}g\rangle = \int_{-\infty}^{+\infty} f^{*}\left(-i\hbar\frac{dg}{dx}\right)dx \xrightarrow{\text{integrate by parts}} \underbrace{-i\hbar f^{*}g\Big|_{-\infty}^{+\infty}}_{0 \text{ at the end points}} -$$

$$-\int_{-\infty}^{+\infty}(-i\hbar)\frac{df^{*}}{dx}g\,dx = \int_{-\infty}^{+\infty}\left(-i\hbar\frac{d}{dx}f\right)^{*}g\,dx = \langle \hat{p}f|g\rangle$$

We just verified that $\hat{p}$ is a Hermitian operator. Its eigenvalues are real.