An introduction to quantum mechanics: arrival to Schrodinger equation

Shyam Biswas
School of Physics, University of Hyderabad,
C.R. Rao Road, Gachibowli, Hyderabad-500046, India
Date: July 6, 2014

Abstract
Introduction to Schrodinger equation is rarely found in standard text books on quantum mechanics which is very often introduced from Planck’s theory on black body radiation problem. Planck’s theory although is good enough to explain the black body radiation problem, which actually is a many body problem of photons in thermal equilibrium, can actually be explained from Planck-Einstein formula for energy of a photon and Bose statistics for energy distribution of photons. Essentially a quantum field theory is rooted in to the black body radiation problem. So, indirectly, quantum mechanics is very often introduced not for a purely mechanical system rather for a thermodynamic system of large number of photons. In this note (teaching material), we tried to introduce quantum mechanics, for problems of a purely mechanical system, in particular, from Hamilton’s optical- mechanical analogy and the facts of wave-particle duality.

Adequacy and inadequacy of classical mechanics
Let us start from Newton’s equation of motion:
\[
\frac{dp}{dt} = \vec{f}(r) \equiv -\vec{\nabla}V(r).
\]  

1For any correspondence, please email to <sbsp [at] uohyd.ac.in>.
2Planck-Einstein formula for energy of a photon was given in the early 20th century. That time photon was known as light’s energy quanta.
3Sir Isaac Newton gave the equation in around 1687.
It is the equation of motion for a point particle whose position, with respect to a locally inertial frame (OXYZ), is given by \( r = x\hat{i} + y\hat{j} + z\hat{k} = |r|\hat{r} = r\hat{r} \), and momentum with respect to the same inertial frame is given by \( p = p\hat{x} + p\hat{y} + p\hat{z} = |p|\hat{p} = p\hat{p} \). Here \( \vec{f}(r) \) is the force acting on the particle, and for conservative case, it is equal to \(-\nabla V(r)\) where \( V(r) \) is called as potential energy. Let us clarify something. OXYZ is an orthogonal (Cartesian) coordinate frame. Any coordinate frame by default is locally inertial.

Let us ask a few question whose answers are yet to be known for clarity. What is the meaning of locally inertial frame? For a certain observation in a laboratory with respect to a laboratory fixed frame, if the law of force \( \vec{f}(r) \) acting on a particle is unaltered for all possible motion in the laboratory, then the lab fixed frame of reference can be called as locally inertial frame. What is meant by motion? Motion of a particle at time \( t \) is defined as its mechanical state denoted by \( \{ r(t), p(t) \} \) or simply by \( \{ r, p \} \).

What is equation of motion? Implicit relation between \( r(t) \) and \( p(t) \) in terms of force is called equation of motion. Why is the equation of motion necessary at all? It is necessary to know the time evolution of the mechanical state \( \{ r(t), p(t) \} \) if the initial mechanical state \( \{ r(0), p(0) \} \) is known. Newton gave an equation of motion as in Eqn. (1).

\[
m \frac{d^2 r}{dt^2} = \vec{f}(r). \tag{2}
\]

As an example, let us consider the case of a 1-D harmonic oscillator of angular frequency \( \omega \). For this case, Newton’s equation of motion takes the form

\[
\frac{d^2 x}{dt^2} = -\omega^2 x. \tag{3}
\]

We can easily solve this (second order) differential equation, and write the general so-
\[ x(t) = x(0) \sin(\omega t) + \frac{p_x(0)}{m\omega} \cos(\omega t). \]  

(4)

Therefrom we can easily get the time evolution of momentum using \( p_x(t) = m \frac{dx}{dt} \) as

\[ p_x(t) = -p_x(0) \sin(\omega t) + m x(0) \omega \cos(\omega t). \]  

(5)

Although position and momentum in Newtonian formalism are independent, they are implicitly related by the equation of motion. Solving the equation of motion (second order differential equation) we can in principle know the time evolution of position \( r(t) \) and therefrom that of momentum \( p(t) \). But, to know the same we need to know the initial state \( \{r(0), p(0)\} \). This ends our little journey to the Newtonian formalism in classical mechanics. To come to the inadequacy of classical mechanics, let us make a little journey to the other formalisms of classical mechanics.

**Little journey to Lagrangian and Hamiltonian formalism**

In Lagrangian formalism\(^5\), coordinates and their corresponding velocities are considered to be independent, and the equations of motion are obtained from a general principle called least action principle\(^6\). According to this principle, action \( S = \int_{t_1}^{t_2} L(q_1, q_2, q_3, \ldots q_s, \dot{q}_1, \dot{q}_2, \dot{q}_3, \ldots, \dot{q}_s, t) dt \) of a mechanical system (e.g. a point particle, a system of point particles, etc) is to be extremized. Here \( q_1, q_2, q_3, \ldots, q_s \) are the minimum number \((s \leq 3N)\) of coordinates necessary to specify uniquely the position of the system, \( \dot{q}_j = \frac{dq_j}{dt} \) is the velocity of the system corresponding to the coordinate \( q_j \), and \( L(q, \dot{q}, t) \) is a suitable function (called Lagrangian) of the generalized coordinates \( \{q_j\} \), generalized

---

\(^5\)It was originally introduced without the principle of least action (but within calculus of variations) by Leonhard Euler in around 1744 and by Joseph-Louis Lagrange in around 1788.

\(^6\)The known form of least action principle was introduced by William Rowan Hamilton in around 1835. Before this a different form of least action (time) principle was introduced for optics by Pierre de Fermat in around 1662. Fermat’s principle was later modernized to an action principle by Gottfried Leibniz, Johann Bernoulli and Pierre Louis Maupertuis defining the action as an integral of its inverse speed along its path length in around 1707-1746. Leonhard Euler (and, possibly, Leibniz) defined action for a material particle as the integral of the particle’s speed along its path through space in around 1744.
velocities $\{\dot{q}_j\}$ and time $t$, such that correct equation of motion can be reproduced extremizing the action. For fixed values of $q(t_1)$ and $q(t_2)$, $q$ may have infinite possibilities for each intermediate values of $t$. But for only one solution $q(t)$, the action is minimized. This solution $q(t)$ is the correct path. Variations of paths around the correct path leads to the Lagrange’s equation of motion\(^7\)

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j}$$

In this formalism, generalized momentum is defined as $p_j = \frac{\partial L}{\partial \dot{q}_j}$. Now, all together, we actually have a second order differential equations. The Lagrange’s equation of motion, to a certain extent, is a generalization of Newton’s equation of motion. Newton’s equation is there inside the correct form of the Lagrangian and subsequent Lagrange’s equation of motion. But, the least action principle is actually a generalized principle which accommodates Newtonian formalism at least for the conservative cases. Lagrangian formalism is not only restricted for the description of motion of a mechanical system but also for fields, etc. It principle does not work for dissipative cases. But, Newtonian formalism works for this case too. Thus, for a mechanical system, once Lagrange’s equation(s) of motion is/are given, there is principle no difference between the two formalisms. There of course may be some advantage and disadvantage of the two formalisms in evaluating time evolution of the mechanical state for complex type of motion of the system.

In Lagrangian formalism, if the Lagrangian has no explicit time dependence, $E = \sum_{j=1}^n \dot{q}_j p_j - L(q, \dot{q}, t)$ can be interpreted as an integral of motion (constant), and can be defined as the energy of the system. In Hamiltonian formalism, coordinates and their corresponding momenta are considered to be independent, and an Hamiltonian is defined as $H(q, p, t) = \sum_{j=1}^n \dot{q}_j p_j - L(q, \dot{q}, t)$. This definition and Lagrange’s equation

\(^7\)Lagrange originally obtained the equation of motion directly from the Newtonian formulation of mechanics introducing Lagrange multiplier in around 1788. Like Lagrange’s equation of motion, d’Alembert’s principle $\sum_i (f_i - m_i \ddot{r}_i) \cdot \delta r_i = 0$ introduced in around 1744 is considered to be another alternative of Newton’s equation of motion particularly for constrained motion. D’Alembert’s principle, however, was possibly written previously in the variational form by Lagrange. D’Alembert’s contribution was to demonstrate that in the totality of a dynamic system the forces of constraint vanish. That is to say that the generalized forces $\{Q_j = \sum_{i=1}^n f_i \cdot \frac{\partial}{\partial q_j}\}$ do not need constraint forces to be included, and consequently, Lagrange’s equation of motion are unchanged under constrained motion.
of motion lead to 2s + 1 first order differential equation as follows.
\[ \dot{q}_i = \frac{\partial H}{\partial p_i}; \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}; \quad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t}. \] (7)

These are called Hamilton’s canonical equations. The first two equations interestingly have a symmetry. And the third equation in many occasions particularly for conserved energy cases are not important. The first equation essentially defines velocity, and the second equation can be recast as a second order differential equation for a generalized coordinate of a mechanical system. For the correct choice of Hamiltonian of the mechanical system, it is essentially Newton’s equation of motion. Hence, once Hamilton’s canonical equations are given, there is no principle difference between the Hamiltonian and Newtonian formalisms. It is useful (like Lagrangian formalism) even for the description of motion of fields. Hamiltonian formalism unlike Newtonian formalism opens a phase-space \((p - q)\) structure for the description of motion. This formalism has an advantage in describing motion under conserved energy. For the 1-D harmonic oscillator, Hamiltonian is given by \(H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2\). Consequently, Hamilton’s canonical equation can be written as \(\ddot{x} = \frac{p}{m}\) and \(\dot{p}_x = -m\omega^2 x\). Constant energy \(E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2\) additionally gives an elliptic trajectory in phase-space. Thus, restrictions on motion can have an illustration in phase-space trajectory.

Hamiltonian formalism provides an advantage to know about time evolution of an arbitrary variable \(f(q, p, t)\) as
\[ \frac{df}{dt} = -(H, f) + \frac{\partial f}{\partial t}. \] (8)
where \(\{H, f\} = \sum_{i=1}^s \left\{ \frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} \right\}\) is a Poisson bracket. This equation essentially is a generalization of Hamilton’s canonical equation. Hamiltonian formalism also provides relationship of momentum and energy with action. An infinitesimal change in action \((S = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt)\) can always be written as
\[ \delta S = \frac{\partial L}{\partial \dot{q}} \delta q \bigg|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} \right] \delta q dt, \] (9)
where variation, for \(\delta q(t_1) = \delta q(t_2) = 0\), leads to the correct part \((q(t))\) which follows Lagrange’s equation of motion. Let us now adiabatically change the boundary conditions such that ends of the integration interval are displaced along the correct/actual
path. This implies that \( \delta q(t_1) \neq 0; \delta q(t_2) \neq 0 \) and \( \frac{d}{dq} - \frac{\partial}{\partial q} = 0 \). For this case, let us assign \( t_1 = t_0 \) and \( t_2 = t \). This implies  

\[
\frac{dS}{dt} \quad  \frac{\partial S}{\partial q} dq(t_0) - \frac{\partial S}{\partial q_0} dq_0.
\]

On the other hand, action calculated along the correct path should be uniquely determined by its initial and final paths, so that we must have \( S = S(q, q_0) \). This implies  

\[
\frac{dS}{dt} \quad \frac{\partial S}{\partial q} dq + \frac{\partial S}{\partial q_0} dq_0.
\]

From the above two equations, we can write  

\[
p = \frac{\partial S}{\partial q} \quad \text{and} \quad p_0 = -\frac{\partial S}{\partial q_0}.
\]

A generalization of the above equation can be represented as  

\[
p = \nabla S(r, t)
\]

where \( S(r, t) \) can be interpreted as a surface of constant action (\( S(r, t) = \text{const.} \)), but, the action is fixed on each and every point the surface and not fixed while it is in motion. For motion of such a surface, we can write  

\[
\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial r} \cdot \frac{dr}{dt} + \frac{\partial S}{\partial r} \cdot \frac{dr}{dt} = \frac{\partial S}{\partial t} + \nabla S(r, t) \cdot \frac{dr}{dt} + p \cdot v
\]

On the other hand, by definition \( \frac{dS}{dt} = L \) and \( p \cdot v = L = H \). Thus, combining above two

---

8Here from one can write the abbreviated action (i.e. the explicitly time independent part of \( S \)) for a correct (or stationary) path as \( S_0 = \sum_{i=1}^{n} p_i dq_i \). The same, for a closed path, can be written as \( S_0 = \oint p_i dq_i \) where \( p_i \) is the angular momentum and \( \theta \) is the angle. For closed circular path of radius \( r \) with constant angular frequency \( \omega \), we can write the angular momentum as \( p_\theta = mr^2 \omega = L \), and if the winding number for the circularly closed path be \( 1 \), then we can easily write \( S_0 = \oint p_i dq_i = 2\pi L \).

9It implies correct path is unique.
equations, we can have Hamilton-Jacobi equation \(^{10}\)

\[ H = \frac{\partial S}{\partial t} \quad \text{or} \quad \frac{\partial S}{\partial t} = -\frac{|\nabla S|^2}{2m} - V(r). \]  

Now action can alternatively be defined as solution of Hamilton-Jacobi equation. Hamilton-Jacobi equation is of course compatible with Lagrange’s equations equation of motion as well as with Newton’s equation of motion.

**Correspondence between geometrical optics and classical mechanics**

Since geometrical optics is a limiting case of wave (physical) optics, let us start from wave optics writing down the form of electric part of the electromagnetic field as

\[ E = E_0(r,t) \sin \left( \frac{\chi(r,t)}{\lambda} \right), \]  

where \( \chi(r,t) = \phi(r,t) \) is the phase and \( \lambda \) is the wavelength which is small comparable to the linear dimensions of the region occupied by the field (i.e. \( E_0(r,t) \equiv \text{constant} \))\(^{11}\). Above form of the electric field must satisfy Maxwell’s equation\(^{12}\):

\[ \nabla^2 E = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}, \]

where \( c \) is the speed of light or phase velocity of the wave. Now, for smaller values of \( \lambda \), we can write

\[ \frac{\partial E}{\partial t} \approx E_0 \frac{\partial \chi}{\partial t} \cos \left( \frac{\chi}{\lambda} \right). \]

\(^{10}\)We can write the second form of the Hamilton-Jacobi equation if \( L = \frac{1}{2} m \dot{r}^2 - V(r) \) or simply \( H = \frac{p^2}{2m} + V(r) \). The second form with conserved/constant energy (\( H = E \)) takes is known as time independent Hamilton-Jacobi equation \( \frac{p^2}{2m} + V(r) = E \) which can be shown to be compatible with eikonal equation obtained for geometrical optics (\( \lambda \to 0 \)). While Hamilton obtained this equation in around 1834, Carl Gustav Jacobi obtained the same in around 1837.

\(^{12}\)James Clerk Maxwell gave the equation in around 1864 stressing light as the electromagnetic wave. Long before this, in around 1704, Newton tried to establish light to be made up of extremely subtle corpuscles (particles, now called photons) contradicting wave nature of light as thought by Christaan Huygens in around 1690.
and
\[ \frac{\partial^2 E}{\partial t^2} \cong E_0 \frac{\partial^2 \chi}{\partial \lambda t^2} \cos \left( \frac{\chi}{\lambda} \right) - E_0 \frac{1}{\lambda^2} \left( \frac{\partial \chi}{\partial t} \right)^2 \sin \left( \frac{\chi}{\lambda} \right). \tag{19} \]

For the limiting case of geometric optics (i.e. for \( \lambda \to 0 \)), we can write
\[ \frac{\partial^2 E}{\partial t^2} \cong -E_0 \frac{1}{\lambda^2} \left( \frac{\partial \chi}{\partial t} \right)^2 \sin \left( \frac{\chi}{\lambda} \right). \tag{20} \]

Similarly, for \( \lambda \to 0 \), we have
\[ \nabla^2 E \cong -E_0 \frac{1}{\lambda^2} \left( \nabla \chi \right)^2 \sin \left( \frac{\chi}{\lambda} \right). \tag{21} \]

Now, putting above two expressions, in to the wave equation, we get a first-order differential equation for the phase \( \phi = \frac{\chi}{\lambda} \), in the geometric optics limit, as
\[ (\nabla \phi)^2 = \frac{1}{c^2} \left( \frac{\partial \phi}{\partial t} \right)^2. \tag{22} \]

This is known as an special form of eikonal equation, in partic ular, for free space \(^{13}\).

On the other hand, in the limiting case of a plane wave \([E = E_0 \sin(k \cdot r - \omega t)]\), the phase is given by
\[ \phi = k \cdot r - \omega t, \tag{23} \]

which implies wave vector or wave number to take the form
\[ k = \frac{\partial \phi}{\partial r} = \nabla \phi, \tag{24} \]

angular frequency to take the form
\[ \omega = -\frac{\partial \phi}{\partial t}. \tag{25} \]

\(^{13}\)For a dispersive medium of locally uniform refractive index \( \mu(r) \), \( c \) is to be replaced by \( c/\mu(r) \) which is the local speed of light at position \( r \).
and the wave equation to lead

\[ k^2 = \frac{\omega^2}{c^2}. \]  

According to the equation for phase, above equation for plane wave is also satisfied by the quantities \( \nabla \phi \) and \( \frac{\partial \phi}{\partial t} \) in the expression of almost plane wave (16)\(^{14}\). It follows that we can consider Eqn.(24) and Eqn.(25) as definitions of wave vector and frequency of an almost plane wave. The wave vector is directed along the normal to a surface of constant phase \( \phi = \text{const} \). The propagation of an almost plane wave may be represented by the displacement in space of a family of surfaces of constant phase so that \( \phi \) is not only constant on each and every point of the surface but may also be the same when it is in motion, as because, for motion of each of such surfaces, \( d\phi = \frac{\partial \phi}{\partial t} \cdot dr + \frac{\partial^2 \phi}{\partial t^2} \cdot dt = 0 \), leads to the phase velocity as \( v_p = \frac{dr}{dt} = \frac{\partial \phi}{\partial r} = \frac{\omega}{k} = c \).

Now comparing Eqns.(13) and (15) obtained for motion of a particle with Eqns.(24) and (25) obtained for geometric optics (considering \( \lambda \to 0 \)), we can conclude that surface of constant action of a particle is propagated similar to a surface of constant phase: the momentum of a particle is similar to the wave vector, the particle energy is similar to the angular frequency.

For a packet (or collection) of (traveling) plane (or almost plane) waves having an angular frequency dispersion (\( \omega = \omega(|k|) \)) of wave vectors (\(|k|\)), group velocity

\[ v_g = \frac{\partial \omega}{\partial k} \]  

in general is different from phase velocity (\( c \)) represents propagation of envelope (or over all shape) of the packet of waves\(^{15} \). A similar to the above equation (obtained for geometrical optics) is also obtained as one of the Hamilton equations

\[ \mathbf{v} = \mathbf{r} = \frac{\partial H}{\partial \mathbf{p}} \]  

\(^{14}\)For almost plane wave \( \mathbf{\phi} = \mathbf{k} \cdot \mathbf{r} + \epsilon(\mathbf{k} \cdot \mathbf{r})^2 - \omega t - \delta(\omega t)^2 + O((\mathbf{k} \cdot \mathbf{r})^3,(\omega t)^3), \) where \( \epsilon, \delta \ll 1 \).

\(^{15}\)Amplitude of the packet may not in general be constant, it gets a periodicity according to Fourier's theorem \( \text{sum of periodic functions is another periodic function} \). Amplitude profile represents the envelope of the packet of waves.
from which we can conclude, that, the group velocity of a wave packet or simply the velocity of a wave packet can be similar to the velocity of a particle. Thus, analogy between mechanics and geometrical optics established a correspondence between a particle and a wave packet.

This optical-mechanical analogy was established by Hamilton in 1835\textsuperscript{16}. However, its physical significance was not well understood up to 1925.

**Difference between wave optics and classical mechanics**

Classical mechanics is analogous to geometrical optics and by no means to wave (physical) optics. The difference between mechanics and wave optics can be illustrated by examples of diffraction (and subsequently of interference) phenomena, as because, for $\lambda \rightarrow 0$ is the limit for geometrical optics, and diffraction effect is nil in this limit.

Single slit diffraction, double slit interference, xray diffraction, etc. of monochromatic wave need aperture(s) of width ($a$) comparable to wavelength ($\lambda$) of light so that $\frac{\lambda}{a}$ significantly differs from 0 i.e. from the geometrical optics limit \textsuperscript{17}. If we observe at least one of these phenomena for a single particle (or for identical particles), then we can claim that Hamilton-Jacobi equation, which is compatible to the geometrical optics limit, is not enough to describe the diffraction of particle. In the limiting case ($\lambda/a \rightarrow 0$), Hamilton-Jacobi equation (or Lagrange’s/Newton’s equation of motion) of course correctly describes other mechanical properties for motion of the particle. So, for the description of diffraction and other mechanical properties for motion of a particle, we need a generalized equation such that for $\lambda/a \rightarrow 0$ we get back Hamilton-Jacobi equation.

Before trying to get the generalized formula, let us know that a particle may have diffraction as well as wave property.

\textsuperscript{16}It can be shown, that, this analogy is relativistically invariant.

\textsuperscript{17}Remember, for single slit diffraction, fringe pattern gets intensity-minima at diffraction-angles $\theta = \sin^{-1}\left(\frac{n\lambda}{a}\right)$ where $n = \pm 1, \pm 2, \pm 3, \ldots$, $\lambda$ is the wavelength of the incident monochromatic light originating from a coherent source, and $a$ is the slit-width. Also remember, for double slit diffraction/interference, particularly for $a \ll d$, fringe pattern gets intensity-maxima at diffraction-angles $\theta = \sin^{-1}\left(\frac{m\lambda}{d}\right)$ where $m = 0, \pm 1, \pm 2, \pm 3, \ldots$, $\lambda$ is the wavelength of the incident monochromatic light originating from a coherent source, $a$ is the slit-width, and $d$ is distance between two slits.
Wave-particle duality

Wave-particle duality means existence of either known particle’s (eg. electron’s, neutron’s, etc’s) wave nature or known wave’s (light’s) particle nature. Since the beginning of 1700s there had been a controversy whether light was a macroscopic wave or a collection of microscopic particles. The controversy was mainly between Huygens (who was in favour of wave nature) and Newton (who was in favour of particle nature). A century later, Young’s double slit experiment performed in around 1800 confirmed light to be a wave. Later on, Maxwell’s equations, obtained in around 1864, stressed that light is an electromagnetic wave. These equations for electromagnetic wave were confirmed experimentally by Heinrich Rudolf Hertz in around 1888, and the wave theory became widely accepted.

In the beginning of 20th century, in connection with the explanation of photo electric effect\textsuperscript{18}, in 1905 and more precisely in 1917 Einstein proposed wave-particle duality for light (electromagnetic wave) in a way that light of wave vector \( \mathbf{k} \) and angular frequency \( \omega \) is composed of a collection of identical particles (called photons) each having energy \( \epsilon = \hbar \omega \) and momentum \( \mathbf{p} = \frac{\hbar \omega}{c} \hat{\mathbf{k}} = \hbar \mathbf{k} \).\textsuperscript{19, 20} Einstein’s energy equation as well as prediction of wave-particle duality for light (proposed in connection with the explanation of photo-electric effect) was experimentally verified in a complementary way by Robert Millikan\textsuperscript{22} in 1916 and by Arthur Compton\textsuperscript{23} in 1923. This way Einstein made a revolution in the theory of optics and resolved more than 200 years’ debate on the concept of light placing Newton’s corpuscular (particle) theory in a novel form in the background of Huygens’ wave(let) or more precisely of Maxwell’s electromagnetic wave theory.

\textsuperscript{18}Photoelectric effect was discovered by Heinrich Rudolf Hertz in 1887 and by Philipp Lenard in around 1901. Electron was discovered, before Lenard’s observation, in 1897 by Joseph John Thomson.

\textsuperscript{19}The expression of energy and momentum apart from the constant (\( \hbar \)) is compatible with Hamilton’s optical-mechanical analogy.

\textsuperscript{20}Einstein’s theory on photo-electric effect, of course, was compatible with Max Planck’s quantum hypothesis (energy of a harmonic oscillator can have certain discrete values: \( \epsilon_n = n \hbar \omega \) where \( \epsilon = \hbar \omega \) is energy element) proposed in connection with the explanation of intensity distribution of black body radiation in 1900-01.

\textsuperscript{21}That, the name photon came later in 1926 and that, rest mass of photon be zero, was understood later in 1930 after Paul Dirac’s work.

\textsuperscript{22}Millikan verified the energy equation for photo-electric effect.

\textsuperscript{23}Compton verified the momentum equation for photo-electric effect. The momentum equation, which was implicitly apparent in Einstein’s 1917 year’s work on quantum theory of radiation, was explicitly written by Compton.
wave theory. Wave particle duality for known particle eg. for an electron of a hydrogen atom was implicitly mentioned by Niels Bohr in 1913. For circular motion of the electron he stoke with Planck’s ad hok formula of quantized energy of a simple harmonic oscillator even for a rotating object, and considered average kinetic energy of a hydrogen atom as half of the total energy of a simple harmonic oscillator even for a rotating object, and considered average kinetic energy of a hydrogen atom as \( \frac{1}{2} \hbar \omega \). This way Bohr obtained discrete energy levels for and electron at the stable/stationary orbits of hydrogen atom as 
\[
\epsilon_n = \frac{m_e e^4}{2(4\pi \varepsilon_0)^2 \hbar^2 n^2},
\]
and explained two decades’ old Balmer’s empirical formula for discrete spectral lines of hydrogen atom. Considering Bohr’s result for angular momentum as a postulate for a circularly rotating object with uniform angular velocity, and the classical results 
\[
p = \frac{\hbar}{i} q, \quad E = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2}, \quad S = -Et + S_0 \quad \text{and} \quad S_0 = \oint p\, dq = \oint p_i\, d\theta = 2\pi L,
\]
for the circular motion, W. Wilson and Arnold Sommerfeld, in around 1916, quantized the abbreviated action \( S_0 \) for a closed path of a conditionally periodic motion of a mechanical system of conserved energy as
\[
\oint p\, dq_j = 2\pi\hbar(n_j + \gamma_j),
\]
where \( n_j = 0, 1, 2, 3, ... \) and \( \gamma_j \) although is a number of order unity which depends on the nature of the boundary conditions for the generalized coordinate considered, was ignored initially, and the ignorance is justified for large values of \( n_j \). This result is known as Bohr-Sommerfeld-Wilson quantization or simply Bohr-Sommerfeld quantization. Quantized energy levels can be easily obtained form this semi-classical result. This result founded the so called old-quantum mechanics.

In 1924 Louis-Victor de Broglie hypothesized wave-particle duality for all particles generalizing the same for photon. He postulated that wave number, frequency, phase
velocity, and group velocity associated with a particle of momentum $p$ and energy $\epsilon$ would be:

$$ k = \frac{p}{\hbar}, $$

$$ \omega = \frac{\epsilon}{\hbar}, $$

$$ v_p = \frac{\epsilon}{p}, $$

and

$$ v_g = \frac{\partial \epsilon}{\partial k} $$

respectively. According to optical-mechanical analogy, if $v_p \neq v_g$, then particles velocity would be represented no longer by $v_p$ but by $v_g$ of a wave packet. He applied this hypothesis to the problem of a circularly periodic stationary motion of a particle. If radius of the circular path be $r$, then wavelength associated with the oscillations, as expected for the wave nature or from the wave particle duality, would be $\lambda = \frac{2\pi}{k} = \frac{2\pi r}{n}$ where $n = 1, 2, 3, \ldots$ is an integer which counts periodicity of the phase of the associated wave around the circle. Now, from $k = p/\hbar$, one can easily get $pr = L = nh$ which is Bohr’s quantized formula for angular momentum.

So, a wave or simply a wave function is associated around the stationary/classical path of a particle. If it is a fact, then, what is the wave equation for the motion of a particle?

Einstein’s theory of relativity (proposed in 1905), he obtained a vanishingly small mass of ‘photon’. To him there was essentially no difference between a ‘photon’ and a usual particle (eg. electron) except one thing that ‘photon’. For photon’s (of non-zero mass, and energy $\hbar \omega$), within relativistic framework, he derived wave vector of a ‘photon’ as $k = \frac{\omega}{c} \hat{e} = \frac{\max}{\sqrt{1-v^2/c^2}} = \frac{\max}{\hbar}$, and for photon’s periodic motion in a force field, he obtained Bohr-Sommerfeld quantization formula from the hypothesis of wave particle duality.

$^{28}$These relations are compatible with optical-mechanical analogy.

$^{29}$Here stationary means, motion is stable only on the circle i.e., the motion is oscillatory around the circle.
Wave mechanics

The wave equation, for both the non-relativistic and the relativistic motion of a point particle, was postulated by Erwin Schrödinger in early 1926 (considering Hamilton-Jacobi equation true for $\hbar \to 0$) as\(^{30}\)

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

where the wave function $\psi(r, t)$ (which usually is a complex variable) was soon after interpreted in 1926 by Max Born as probability amplitude for position of the particle $r$ at time $t$ so that $\psi(r, t)^* \psi(r, t) = |\psi(r, t)|^2$ (like diffraction/interference phenomenon in wave optics) represents probability (density) of the same event. For non-relativistic case, Schrödinger equation can be recast as\(^{31}\)

\[
i \hbar \frac{\partial}{\partial t} \psi(r, t) = \left( \frac{\hat{p}^2}{2m} + V(r) \right) \psi(r, t) = \frac{\hbar^2}{2m} |\nabla \psi(r, t)|^2 + V(r) \psi(r, t) = \left( - \frac{\hbar^2}{2m} \nabla^2 + V(r) \right) \psi(r, t),
\]

where

\[
\hat{p} = -i\hbar \nabla
\]

is the momentum operator in differential form. This equation unlike Bohr-Sommerfeld-Wilson quantization is not semi-classical. Exploring wave nature around classical/stationary path, as considered by de Broglie from Bohr’s semi-classical principle (quantization of angular momentum), is not its primary issue. While Bohr-Sommerfeld-Wilson quantization is applicable only for classically periodic motion of a high energetic particle, Schrödinger equation is applicable for classically all kind of motion with all possible energy.

Soon after the discovery of Schrödinger equation, Schrödinger himself solved the wave equation for free particle, for electron of a hydrogen atom, etc. He not only obtained the same discrete energy levels as obtained by Bohr for hydrogen atom problem

\(^{30}\)Check, that, this equation is compatible with Hamilton-Jacobi equation. Check also that, as $\hbar \to 0$, wavelength, as postulated in wave-particle duality, goes to zero, i.e., geometrical optics limit is recovered.

\(^{31}\)Check, that, this equation reproduces both the Hamilton-Jacobi equation and the optical-mechanical analogy for $\hbar \to 0$. 

14
but also correctly obtained \( l \) degeneracy and \( m \) degeneracy as was postulated by Bohr-Sommerfeld in an \textit{ad hoc} manner. Here-from solving Schrodinger equation, for different mechanical problems, has been a prime interest of physics. This way old quantum mechanics, which is governed by Bohr-Sommerfeld-Wilson quantization, is suppressed by the new quantum mechanics whose one part, called wave mechanics, is governed by the Schrodinger equation. The hypothesis of de Broglie on the wave-particle duality can be considered to be an inter-mediator of old and new quantum mechanics. This hypothesis as well as wave particle duality, in particular the diffraction pattern an electron passing through a crystal, was experimentally verified in 1927 by Clinton Davisson, Lester Germer, and George Paget Thomson.

Matrix mechanics

Another part of the new quantum mechanics is known as matrix mechanics which was formulated by Werner Heisenberg, Max Born, and Pascual Jordan in around 1925 postulating an operator algebra for two canonically conjugate variables eg. commutation of position operator \( \hat{x} \) (Now \( \hat{x} \) is an infinite dimensional square matrix,) and momentum operator \( \hat{p}_x \) (Now \( \hat{p}_x \) is not differential operator but an infinite dimensional square matrix) was postulated as

\[
[\hat{x}, \hat{p}] = i\hbar I.
\]

Noncommutativity of canonically conjugate variables, like the above, physically means they are not theoretically (and of course experimentally) measurable simultaneously in any single state.

Matrix mechanics was the first conceptually autonomous and logically consistent formulation of quantum mechanics. It extended the Bohr model by describing how the quantum jumps occur. It did so by interpreting the physical properties of a mechanical system as matrices that evolve in time. The time evolution for a matrix operator (\( \hat{O}(t) \)) is known as Heisenberg equation of motion

\[
\frac{d\hat{O}}{dt} = -\frac{[\hat{H}, \hat{O}]}{i\hbar} + \frac{\partial \hat{O}}{\partial t},
\]

where \( \hat{H} = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 + V(\hat{x}, \hat{y}, \hat{z}) \) is the Hamiltonian operator, for the system, in matrix form, and \([\hat{H}, \hat{O}] = \hat{H}\hat{O} - \hat{O}\hat{H} \) is the commutator of \( \hat{H} \) and \( \hat{O} \). Heisenberg equation of
motion in matrix mechanics is the counter equation to the Schrödinger equation in wave mechanics.

Later on, in 1927, Heisenberg obtained an uncertainty relation between two canonically conjugate variables, as a consequence of their operator algebra, e.g., for position ($\hat{x}$) and momentum ($\hat{p}_x$) operators, as

\[ \Delta x \Delta p_x \geq \frac{\hbar}{2}. \tag{39} \]

Matrix mechanics was shown to be equivalent to wave mechanics collectively by Schrödinger, Dirac and John von Neumann in around late 1926. In contrast to the wave mechanics, it produces spectra of energy operators by purely algebraic, ladder operator, methods. Based on this methods, Wolfgang Pauli derived the Hydrogen atom spectrum in 1926 before the development of wave mechanics.

Quantum mechanics

Matrix mechanics and wave mechanics together form new quantum mechanics or simply the quantum mechanics. The two different formalism, to a certain extent, was shown to be equivalent by Schrödinger in 1926. He noticed existence of wave mechanics in matrix mechanics, and related each differential operator of wave mechanics to a matrix operator of the matrix mechanics. Dirac in around 1926 pointed out the difference between states and observables of a physical system, a distinction which was present in wave mechanics but not in matrix mechanics. He interpreted eigenvectors of matrices (observables) as states of the system. He further introduced unitary transformation of a complete set of orthonormal states, like canonical transformation in classical mechanics, in Hilbert space using Dirac delta function. This way he showed equivalence of matrix mechanics and wave mechanics, and unified them in a single framework called Dirac formulation of quantum mechanics. Later in 1939, Paul Dirac introduced bra-ket notation in quantum mechanics.

Dirac’s approach required essential use of delta functions, which were suspect from the standpoint of mathematical rigor. In 1927 von Neumann placed quantum mechanics on a rigorous mathematical foundation with separable Hilbert space and rigorously proved (i.e., without the use of delta functions) of the equivalence of matrix and wave mechanics.

\[ \text{A relativistic theory, for quantum mechanics, was also developed by Paul Dirac in around 1928.} \]
mechanics. However, Dirac’s formulation of quantum mechanics is very useful and influential despite its lack of mathematical rigor. It is used extensively by physicists and it inspired some powerful mathematical developments in functional analysis. Eventually, mathematicians developed a suitable framework for placing Dirac’s formal framework on a firm mathematical foundation, which is known as a *rigged* Hilbert space $^{33}$.

To conclude, a rigorous mathematical formulation for quantum mechanics was developed in late 1920s (and even after that) from collective works of David Hilbert, Paul Dirac, John von Neumann, Hermann Weyl, and others keeping both the matrix mechanics and wave mechanics in a single framework. This formulation is distinguished from mathematical formalisms for theories developed prior to the early 1900s by the use of abstract mathematical structures, such as infinite-dimensional Hilbert spaces and differential operators on these spaces. Many of these structures are drawn from functional analysis, a research area within pure mathematics that was influenced in part by the needs of quantum mechanics. In brief, values of physical observables such as energy and momentum were no longer considered as values of functions on phase space, but as eigenvalues of linear operators in Hilbert space. Motion of a system/particle can now be understood as rotation of a state vector in Hilbert space.

Referred Books: [1, 2, 3, 4, 5]

---

Bibliography


