January 28, 2006

Fundamentals of Quantum Mechanics for Laser Remote Sensing

Professor Xinzhao Chu

Reference Books:

1. Dirac, P. A. M., *The principles of Quantum Mechanics*
2. Landau, L. D., and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory)*
3. Claude Cohen-Tannoudji et al., *Quantum Mechanics (Volume One and Two)*
4. Schiff, L. I., *Quantum Mechanics*

Outline

I. The Principle of Uncertainty (Indeterminacy)
II. The Principle of Superposition of States
III. The Equations of Motion – Schrödinger Equation
IV. Postulates of Quantum Mechanics

In the present state of scientific knowledge, quantum mechanics plays a fundamental role in the description and understanding of natural phenomena. In fact, phenomena that occur on a very small (atomic or subatomic) scale cannot be explained outside the framework of quantum physics. For example, the existence and the properties of atoms, the chemical bound and the propagation of an electron in a crystal cannot be understood in terms of classical mechanics. Even when we are concerned only with macroscopic physical objects (i.e., whose dimensions are comparable to those encountered in everyday life), in principle, it is necessary to begin by studying the behavior of their various constituent atoms, ions, electrons, in order to arrive at a complete scientific description. There are many phenomena that reveal, on a macroscopic scale, the quantum behavior of nature. It is in this sense that it can be said that quantum mechanics is the basis of our present understanding of all natural phenomena, including those traditionally treated in chemistry, biology, etc …

From a historical point of view, quantum ideas contributed to a remarkable unification of the concepts of fundamental physics by treating material particles and radiation on the same footing. At the end of the nineteenth century, people distinguished between two entities in physical phenomena: matter and radiation. Completely different laws were used for each one. To predict the motion of material particles, the laws of Newtonian mechanics were utilized. Their success was very impressive. With regard to radiation, the theory of electromagnetism, thanks to the introduction of Maxwell's equations, had produced a unified interpretation of a set of phenomena which had previously been considered as belonging to different domains: electricity, magnetism, and optics. In particular, the electromagnetic theory of radiation had been spectacularly confirmed experimentally by the discovery of Hertzian waves. Finally, interactions between radiation and matter were well explained by the Lorentz force. This set of laws had brought physics to a point that could be considered satisfactory, in view of the experimental data at the time.

However, at the beginning of the twentieth century, physics was to be marked by the profound upheaval that led to the introduction of relativistic mechanics and quantum mechanics. The relativistic revolution and the quantum revolution were, to a large extent, independent, since they challenged classical physics on different points. Classical laws cease to be valid for material bodies traveling at very high speeds, comparable to that of light (relativistic domain). In addition, they are also found to be wanting on an atomic or subatomic scale (quantum domain). However, it is important to note that classical physics, in both cases, can be seen as an approximation of the new theories, an approximation that is valid for most phenomena on an everyday scale. For example, Newtonian mechanics enables us to predict correctly the motion of a solid body, providing it is non-relativistic (speeds much smaller than that of light) and macroscopic (dimension much greater than atomic ones). Nevertheless, from a fundamental point of view, quantum theory remains indispensable. It is the only theory that enables us to understand the very existence of a solid body and the values of the macroscopic parameters (such as density, specific heat, and elasticity, etc.) associated with it.

I. The Principle of Uncertainty (Indeterminacy)

A. Classical Causality and Laplacian Determinacy

In classical theory, there is so-called "Laplacian Determinacy" or "Classical Causality". It is a kind of philosophy in the classical physics. Its content is as below.

Pierre Simon Laplace stated in 1776: "An intelligence that knows all of the relations of the entities of the universe at one instant could state their positions, motions, and general effects at any instant in the past and future."

If translating Laplace's statement into classical mechanics, the determinacy and causality is expressed in the followings:

"A complete description of the state of a classical system is to state all its coordinates and velocities at a given instant. With these initial data, the equations of motion completely determine the behavior of the system at all subsequent instants."

According to this determinacy, two major features of classical mechanics are (1) all coordinates and velocities of a classical system can be determined with infinite accuracy at the same time; (2) a classical system is completely predictable: if you know the state of a system at one instant time, you can predict its states to ultimate accuracy at any future time by using the classical equations of motion.

B. Breakdown of the Determinacy

Classical mechanics and electrodynamics work very well in classical system, mainly macroscopic dimension world. However, when we attempt to apply classical mechanics and electrodynamics to explain atomic phenomena, they break down. One example is to apply classical electrodynamics to a model of an atom in which the electrons move round the nucleus in classical orbits. Such motion is an accelerated motion of charges, so the electrons would have to emit electromagnetic waves continually. By this emission, the electrons would lose their energy, and this would eventually cause them to fall into the nucleus. Thus, according to classical electrodynamics, the atom would be unstable, which does not agree with reality at all.

Another example is the experimentally observed phenomenon, known as "electron diffraction". When a homogeneous beam of electrons passes through a crystal, the emergent beam exhibits a pattern of alternate maxima and minima of intensity, very similar to the diffraction pattern observed in the diffraction of electromagnetic waves. Thus, under certain conditions, the behavior of material particles (in this case, the electrons) displays features belonging to wave processes.

How does the above phenomenon contradict our usual ideas of motion? Let's look at another example – the double-slit interference experiment with electrons.

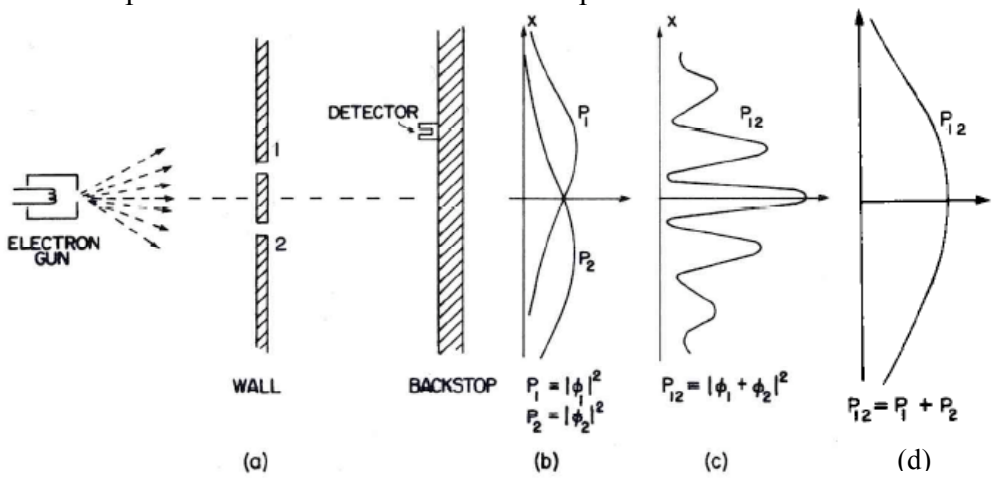

Figure 1. Double-slit interference experiment with electrons.

On observing the passage of a beam of electrons through one of the slits, the other being covered, we obtain some pattern of intensity distribution (P_1) on a continuous screen placed behind the slit. Similarly, by uncovering the second slit and covering the first, we obtain another pattern (P_2). If uncovering both slits, on the basis of ordinary classical ideas, we should expect a pattern which is a simple superposition of the other two: each electron, moving in its path, passes through one of the slits and has no effect on the electrons passing through the other slit, i.e., $P_{12} = P_1 + P_2$, shown as the rightmost plot in Figure 1. However, the reality tells us otherwise. An interference pattern similar to the Young's double slit experiment with light is displayed for the electron interference. This phenomenon can be explained only by the quantum mechanics. We cannot predict the path of an electron any more – the breakdown of the determinacy. Instead, only the probability of the electron occurring at certain position of the screen may be predicted.

C. Light Quanta and the Planck-Einstein Relations

Newton considered light to be a beam of particles, able to bounce back upon reflection from a mirror. During the first half of the nineteenth century, the wavelike nature of light was demonstrated (interference, diffraction, etc.). This enabled optics to be integrated into electromagnetic theory. In this framework, the speed of light, c, is related to electric and magnetic constants and light polarization phenomena can be interpreted as manifestations of the vectorial character of the electric field.

However, the study of blackbody radiation, which electromagnetic theory could not explain, led Planck to suggest the hypothesis of the quantization of energy (1990): for an electromagnetic wave of frequency ν, the only possible energies are integral multiples

of the quantum hν, where h is a new fundamental constant. Generalizing this hypothesis, Einstein proposed a return to the particle theory (1905): light consists of a beam of photons (whose rest mass is zero), each possessing an energy hν. Einstein showed how the introduction of photons made it possible to understand, in a very simple way, certain as yet unexplained characteristics of the photoelectric effect. Twenty years had to elapse before the photon was actually shown to exist, as a distinct entity, by the Compton effect (1924).

These results lead to the following conclusion: the interaction of an electromagnetic wave with matter occurs by means of elementary indivisible processes, in which the radiation appears to be composed of particles, the photons. Particles parameters (the energy E and the momentum p of a photon) and wave parameters (the angular frequency $\omega = 2\pi\nu$ and the wave vector $\vec{k}$, where $\left|\vec{k}\right| = 2\pi/\lambda$, with ν the frequency and λ the wavelength) are linked by the fundamental relations:

$$E = h\nu = \hbar\omega$$

$$\vec{p} = \hbar\vec{k} \qquad \text{(Planck-Einstein relations)}$$

where $\hbar = h/2\pi$ is defined in terms of the Planck constant h= 6.626 × 10^{-34} J s. During each elementary process, energy and total momentum must be conserved.

D. Wave-Particle Duality of Light

What is light – wave or particle? Analyzing Young's well-known double-slit experiment will lead us to the following conclusion: a complete interpretation of the phenomena can be obtained only by conserving both the wave aspect and the particle aspect of light (although they seem a priori irreconcilable). We shall then show how this paradox can be resolved by the introduction of the fundamental quantum concepts.

The experiment is similar to the schematic shown in Figure 1. Only the electron gun was replaced by a monochromatic light source. The light falls on a screen (a photographic film, for example) pierced by two narrow slits 1 and 2. If we block slit 2, we obtain a light intensity distribution $I_1(x)$ on the screen, which is the diffraction pattern of slit 1. Similarly, when the slit 1 is blocked, the diffraction pattern of slit 2 is described by $I_2(x)$. When the two slits 1 and 2 are open at the same time, we observe a system of interference fringes on the screen. In particular, we note that the corresponding intensity I(x) is not the sum of the intensities produced by slit 1 and slit 2 separately: $I(x) \neq I_1(x) + I_2(x)$.

Let us concentrate on the interference phenomenon (ignore the diffraction pattern). From the particle picture, we could attempt to explain it by an interaction between the photons that pass through the slit 1 and those that pass through the slit 2. Such an explanation would lead to the following predictions: if the intensity of the light source is decreased until the photons strike the screen practically one by one, the interaction between the photons must diminish, and eventually, varnish. The interference fringes should therefore disappear.

The wave theory of light provides a completely natural interpretation of the fringes. The light intensity at a point of the screen is proportional to the square of the amplitude of the electric field at this point. If $E_1(x)$ and $E_2(x)$ represent, in complex

notation, the electric fields produced at x by slits 1 and 2 respectively (the slits behave like secondary sources), the total resultant field at this point when slits 1 and 2 are both open is: $E(x) = E_1(x) + E_2(x)$. We then have the light intensity:

$$I(x) \propto |E(x)|^2 = |E_1(x) + E_2(x)|^2 = |E_1(x)|^2 + |E_2(x)|^2 + 2\vec{E}_1 \cdot \vec{E}_2 .$$

Since the intensity $I_1(x)$ and $I_2(x)$ are proportional to $|E_1(x)|^2$ and $|E_2(x)|^2$ respectively, above equation shows that I(x) differs from $I_1(x)+ I_2(x)$ by an interference term $2\vec{E}_1 \cdot \vec{E}_2$ that depends on the phase difference between E_1 and E_2 and whose presence explains the fringes. The wave theory thus predicts that decreasing the intensity of the light source will simply cause the interference fringes to diminish in intensity but not to varnish.

What actually happens when the light source emits photons practically one by one? Neither the predictions of the wave theory nor those of the particle theory are true. In fact:

(1) If we use a photographic film as the screen and increase the exposure time so as to capture a large number of photons, we observe that the fringes have not disappeared. Therefore, the purely corpuscular interpretation, according to which the fringes are due to interaction between different photons, must be rejected.

(2) On the other hand, we can expose the photographic film during a time so short that it can only receive a few photons. We then observe that each photo produces a localized impact on the film and not a very weak interference pattern. Therefore, the purely wave interpretation must also be rejected.

In reality, as more and more photons strike the photographic film, the following phenomenon occurs. Their individual impacts seem to be distributed in a random manner, and only when a great number of them have reached the file does the distribution of the impacts begin to have a continuous aspect. The density (i.e., the probability) of the impacts at each point of the file corresponds to the interference fringe: maximum on a bright fringe and zero on a dark fringe. It can thus be said that the photons, as they arrive, build up the interference pattern.

The result of this experiment leads to a paradox. Within the framework of the particle theory, it can be expressed in the following way. Since photon-photon interactions are excluded, each photon must be considered separately. But then it is not clear why the phenomena should change drastically according to whether only one slit or both slits are open. For a photon passing through one of the slits, why should the fact that the other is open or closed have such a critical importance?

In the above experiment, we did not seek to determine through which slit each photon passed before it reached the screen. In order to obtain this information, we can imagine placing detectors (photomultipliers) behind slits 1 and 2. It will then be observed that if the photons arrive one by one, each one passes through a well-determined slit (a signal is recorded either by the detector placed behind slit 1 or by the one behind slit 2 but not by both at once). But, obviously, the photons detected in this way are absorbed and do not reach the screen. Remove the PMT which blocks slit 1, then the one that remains behind slit 2 tells us that, out of a large number of photons, about half pass through slit 2. We conclude that the others (which can continue as far as the screen) pass through slit 1. But the pattern that they gradually construct on the screen is not an interference pattern any more, since slit 2 is blocked. It is only the diffraction pattern of slit 2.

The above analysis shows that it is impossible to explain all the phenomena observed if only one of the two aspects of light, wave or particle, is considered. We must accept the possibility that these classical wave and particle concepts may not be valid in the new ("microscopic") domain that we are entering. For example, an essential characteristic of this new domain appeared when we placed counters behind Young's slits: *when one performs a measurement on a microscopic system, one disturbs it in a fundamental fashion*. This is a new property since in the macroscopic domain we always have the possibility of conceiving measurement devices whose influence on the system is practically as weak as one might wish. This critical revision of classical physics is imposed by experiment and must be guided by experiment.

Let us consider the paradox stated above concerning the photon that passes through one slit but behaves differently depending on whether the other slit is open or closed. We saw that if we try to detect the photons when they cross the slits, we prevent them from reaching the screen. More generally, a detailed experimental analysis shows that *it is impossible to observe the interference pattern and to know at the same time through which slit each photon has passed*. Thus, in order to resolve the paradox, it is necessary to give up the idea that a photon inevitably passes through a particular slit. This also leads us to question the concept of a particle's trajectory that is one of the fundamental concepts of classical physics. In other words, we may not be able to know all the coordinates and velocities of a particle at the same time.

Furthermore, as the photons arrive one by one, their impacts on the screen gradually build up the interference pattern. This implies that for a particular photon, we are not certain in advance where it will strike the screen, i.e., the position of the photon is not predictable. Since these photons are all emitted under the same conditions, the experiment result also tells us that the initial conditions cannot completely determine the subsequent motion of a particle. Thus, another classical idea – the Laplacian determinacy – has been destroyed. We can only say that when a photon is emitted, the probability of its striking the screen at position x is proportional to the intensity *I(x)* calculated using wave theory, i.e., to $|E(x)|^2$.

Based upon the above discussion, **the concept of wave-particle duality of light is formulated. It is schematically summarized as follows**:

(1) The particle and wave aspects of light are inseparable. *Light behaves simultaneously like a wave and like a flux of particles (photons)*. The wave enables us to calculate the probability of the manifestation of a particle. The particle defines a finite and quantized energy and momentum.

(2) Predictions about the behavior of a photon can only be probabilistic.

(3) The information about a photon at time t is given by the wave E(r,t), which is a solution of Maxwell's equations. This wave characterizes the state of the photons at time t. E(r,t) is interpreted as the probability amplitude of a photon appearing at time t and at the position r. This means that the corresponding probability is proportional to $|E(r,t)|^2$.

E. Material Particles and Matter Waves – the de Broglie Relations

In 1923, in analogies to the wave-particle duality of light, de Broglie put forth the following hypothesis: *material particles, just like photons, can have a wavelike aspect*. Electron diffraction experiments by Davisson and Germer in 1928 (using a cubic nickel

crystal) and independently by Thompson in 1928 (using thin films of celluloid, gold, platinum or aluminium) strikingly confirmed the existence of a wavelike aspect of matter by showing that *interference patterns could be obtained with material particles such as electrons*. The wave associated with material particles is named *de Broglie matter wave*.

The expression of de Broglie matter wave is that all material particles possess a wave-particle duality nature. A material particle with energy E and momentum $\vec{p}$ is associated with a matter wave whose angular frequency $\omega = 2\pi\nu$ and wave vector $\vec{k}$ by the similar relations as for photons:

$$E = h\nu = \hbar\omega$$

$$\vec{p} = \hbar\vec{k}$$

In other words, the corresponding wavelength of the matter wave is given by

$$\lambda = \frac{2\pi}{|\vec{k}|} = \frac{h}{|\vec{p}|} = \frac{h}{p} \qquad \text{(de Broglie relation).}$$

The very small value of the Planck constant h (6.626×10^{-34} Js) explains why the wavelike nature of matter is very difficult to demonstrate on a macroscopic scale. Complement A_I discusses the orders of magnitude of the de Broglie wavelengths associated with various material particles.

F. Heisenberg Uncertainty Relation

Werner Heisenberg wrote in 1924: "There is a fundamental limit on the accuracy to which position and velocity can be determined." In quantum mechanics, a complete description of the state of a physical system like classical system is in principle impossible, because the coordinates and the corresponding velocities cannot exist simultaneously. Thus, the state of a quantum system is described by means of a smaller number of quantities than in classical mechanics, i.e., it is less detailed than a classical description.

A very important consequence follows this is the nature of the predictions made in quantum mechanics. Whereas a classical description is sufficient to predict the future motion of a mechanical system with complete accuracy, the less detailed description given in quantum mechanics evidently cannot be enough to do so. This means that even if an electron is in a state described in the most complete manner possible in quantum mechanics, its behavior at subsequent instants is still in principle uncertain. Therefore, quantum mechanics cannot make completely definite predictions concerning the future behavior of the electron. For a given initial state of the electron, a subsequent measurement can give various results. The problem in quantum mechanics becomes determining the probability of obtaining various results on performing this measurement. It is understood that in some cases the probability of a given result of measurement may be equal to unity, i.e., certainty, so that the result of that measurement is unique. These unique states are eigenstates that will be introduced below.

Quantitative expression of the Heisenberg uncertainty principle will be described after the principle of superposition of states is introduced next.

II. The Principle of Superposition of States

A. Definition of State

According to Dirac in 1930, a state of a system may be defined as an undisturbed motion that is restricted by as many conditions or data as are theoretically possible without mutual interference or contradiction.

In practice, the conditions could be imposed by a suitable preparation of the system, consisting perhaps in passing it through various kinds of sorting apparatus, such as slits and polarizers, the system being left undisturbed after the preparation.

The word "state" here refers to the state of a system at an instant time. The change of the state with time is called a "state of motion".

Let us take any atomic ("microscopic") system, composed of particles or bodies with specified properties (mass, moment of inertia, etc.) interacting according to specified laws of force. There will be various possible motions of the particles or bodies consistent with the laws of force. Each such motion is called a *state* of the system. According to classical ideas, one could specify a state by giving numerical values to all the coordinates and velocities of the various component parts of the system at some instant of time, the whole motion being them completely determined. Now we learned that we cannot observe a small system with that amount of detail which classical theory supposes. The limitation in the power of observation puts a limitation on the number of data that can be assigned to a state. Thus a state of an atomic system must be specified by few or more indefinite data than a complete set of numerical values for all the coordinates and velocities at some instant of time.

B. Superposition of States

The general principle of superposition of quantum mechanics applies to the states of any one physical system. It requires us to assume that between these states there exist peculiar relationships such that whenever the system is defined in one state, we can consider it as being partly in each of two or more other states. The original state must be regarded as the result of a kind of *superposition of the two or more new states*, in a way that cannot be conceived on classical ideas.

Any state may be considered as the result of a superposition of two or more other states, and indeed in an infinite number of ways. Conversely any two or more states may be superposed to give a new state. The procedure of expressing a state as the result of superposition of a number of other states is a mathematical procedure that is always permissible, independent of any reference to physical conditions, like the procedure of resolving a wave into Fourier components. Whether it is useful in any particular case, though, depends on the special physical conditions of the problem under consideration.

The nature of the relationships that the superposition principle requires to exist between the states of any system is of a kind that cannot be explained in terms of familiar physical concepts. One cannot in the classical sense picture a system being partly in each of two states and see the equivalence of this to the system being completely in some other state. There is an entirely new idea involved, to which one must get accustomed and in

terms of which one must proceed to build up an exact mathematical theory, without having any detailed classical picture.

When a state is formed by the superposition of two other states, it will have properties that are in some vague way intermediate between those of the two original states and that approach more or less closely to those of either of them according to the greater or less "weight" attached to this state in the superposition process. The new state is completely defined by the two original states when their relative weights in the superposition process are known, together with a certain phase difference. The exact meaning of weights and phases is provided in the general case by the mathematical theory.

The non-classical nature of the superposition process is brought out clearly if we consider the superposition of two states, A and B, such that there exists an observation which, when made on the system in state A, is certain to lead to one particular result, a, and when made on the system in state B, is certain to lead to some different result, b. What will be the result of the observation when made on the system in the superposed state? The answer is that the result will be sometimes *a* and sometimes *b*, according to a probability law depending on the relative weights of A and B in the superposition process. It will never be different from both a and b. The intermediate character of the state formed by superposition thus expresses itself through the probability of a particular result for an observation being intermediate between the corresponding probabilities for the original states, not through the result itself being intermediate between the corresponding results for the original states.

In this way, we see that such a drastic departure from ordinary ideas as the assumption of superposition relationships between the states is possible only on account of the recognition of the importance of the disturbance accompanying an observation and of the consequent indeterminacy in the result of the observation. When an observation is made on any atomic system that is in a given state, the result will not be determinate in general, i.e., if the experiment is repeated several times under identical conditions, several different results may be obtained. It is a law of nature that if the experiment is repeated a large number of times, each particular result will be obtained in a definite fraction of the total number of times, so that there is a definite probability of its being obtained. This probability is what the quantum theory sets out to calculate. Only in special cases when the probability for some result is unity is the result of the experiment determinate.

It is important to remember that the superposition that occurs in quantum mechanics is of an essentially different nature from any occurring in the classical theory, as is shown by the fact that the quantum superposition principle demands indeterminacy in the results of observations in order to be capable of a sensible physical interpretation. Therefore, we do not encourage the analogies of quantum mechanics to classical mechanics by so-called "wave mechanics", as they are liable to be misleading.

C. Eigenstates and Eigenvalues

In the above example of a state being a superposition of two states A and B, the states A and B are called the *eigenstates* if observations are certain to give the same result a when made on A and are certain to give the same result b when made on B. The results a and be are called the *eigenvalues*.

The eigenstates can be defined as the states of a system that the probability of measurements on any one of these states giving a particular value is unity. The particular value associated with each eigenstate is defined as the eigenvalue.

D. Principle of Spectral Decomposition

Armed with the ideas introduced above, we are now going to discuss another simple optical experiment, whose subject is the polarization of light. This will permit us to introduce the fundamental concepts that concern the measurement of physical quantities.

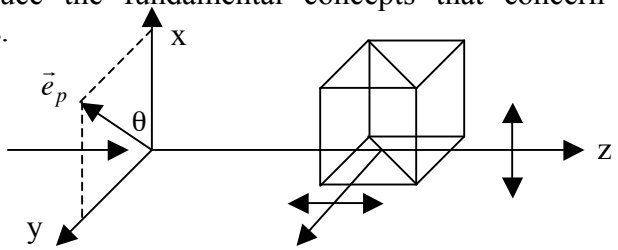

Figure 2. Experiment of light polarization

As shown in Figure 2, the experiment consists of directing a polarized plane monochromatic light through a polarized beam splitter (the cubic crystal). The propagation direction of the light is along z-axis, and the polarization of the light is denoted as a unit vector $\vec{e}_p$. The polarized beam splitter transmits light polarized parallel to x-axis and reflects light polarized parallel to y-axis. The classical description of this experiment (a description which is valid for a sufficiently intense light beam) is the following. The polarized plane wave is characterized by an electric field of the form: $\vec{E}(\vec{r},t) = E_0\vec{e}_p e^{i(kz-\omega t)}$, where E_0 is a constant. The light intensity I is proportional to $|E_0|^2$. After its passage through the polarized beam splitter, the plane waves polarized along x-axis and z-axis are given respectively by $\vec{E}'(\vec{r},t) = E_0'\vec{e}_x e^{i(kz-\omega t)}$ and $\vec{E}''(\vec{r},t) = E_0''\vec{e}_z e^{i(ky-\omega t)}$. The intensities of these two waves are given by $I' = I\cos^2\theta \propto |E_0'|^2$ and $I'' = I\sin^2\theta \propto |E_0''|^2$. Here, $\vec{e}_x$ is the unit vector of the x-axis and θ is the angle between $\vec{e}_x$ and $\vec{e}_p$.

What will happen on the quantum level, i.e., when I is weak enough for the photons to reach the polarized beam splitter one by one? We then place a photon detector on each side of the beam splitter. The detector never registers a "fraction of a photon". Either the photon transmits through the beam splitter and hits the detector placed along z-axis, or it is entirely reflected and strikes the detector placed along y-axis. In addition, we cannot predict with certainty whether a given incident photon will transmit or reflect. We can only know the corresponding probabilities. If we send out a large number N of photons one after the other, the result will correspond to the classical law, in the sense that about $N\cos^2\theta$ photons will be transmitted and $N\sin^2\theta$ photons will be reflected.

We shall retain the following ideas from the above description:

(1) The measurement device can give only certain privileged results, which is called eigen results. In the above experiment, there are only two possible results: the photon transmits through the beam splitter or is reflected by the beam splitter. There is quantization of the result of the measurement, in contrast to the classical case where the transmitted intensity can vary continuously, according to the value of θ, between 0 and 1.

(2) To each of these eigen results corresponds an eigenstate. Here, the two eigenstates are characterized by: $\vec{e}_p = \vec{e}_x$ and $\vec{e}_p = \vec{e}_y$. If $\vec{e}_p = \vec{e}_x$, we know with certainty that the photon will traverse the beam splitter; if $\vec{e}_p = \vec{e}_y$, it will be reflected by the beam splitter. The correspondence between eigen results and eigenstates is therefore the following: before the measurement, if the particle is in one of the eigenstates, the result of the measurement is certain and it can only be the associated eigen result.

(3) When the state before the measurement is arbitrary, only the probabilities of obtaining the different eigen results can be predicted. To find these probabilities, one decomposes the state of the particles into a linear combination of the various eigenstates. Here, for an arbitrary $\vec{e}_p$, we write: $\vec{e}_p = \vec{e}_x \cos\theta + \vec{e}_y \sin\theta$. The probability of obtaining a given eigen result is then proportional to the square of the absolute value of the coefficient of the corresponding eigenstate. The proportionality factor is determined by the condition that the sum of all these probabilities must be equal to 1. We thus deduce that each photon has a probability $\cos^2\theta$ of traversing the beam splitter and a probability of $\sin^2\theta$ of being reflected (we know $\cos^2\theta + \sin^2\theta = 1$). This rule is called in quantum mechanics the *principle of spectral decomposition*. Note that the decomposition to be performed depends on the type of measurement device being considered, since one must use the eigenstates that correspond to it.

(4) After passing through the polarized beam splitter, the light is completely polarized along $\vec{e}_x$. This means that the state of the photons is the eigenstate characterized by $\vec{e}_x$ after they traverse the beam splitter. Therefore, there has been an abrupt change in the state of the particles. Before the measurement, this state was defined by a vector $\vec{E}(\vec{r},t)$ that was collinear with $\vec{e}_p$. After the measurement, the state of the photon is described by a different vector, which is now collinear with $\vec{e}_x$. This expresses the fact that the measurement disturbs the microscopic system (here, the photon) in a fundamental fashion.

If we return to the Young's double-slit interference experiment with light, it is not difficult to reach such a conclusion that the phenomena can only be explained by each single photon interferes with itself. Some time before the discovery of quantum mechanics, people realized that the connection between light waves and photons must be of a statistical character. What they did not clearly realize, however, was that the wave function gives information about the probability of one photon being in a particular place and not the probable number of photons in that place. The importance of the distinction can be made clear in the following way. Suppose we have a beam of light consisting of a large number of photons split up into two components of equal intensity. On the assumption that the intensity of a beam is connected with the probable number of photons in it, we should have half the total number of photons going into each component. If the two components are now made to interfere, we should require a photon in one component to be able to interfere with one in the other. Sometimes these two photons would have to annihilate one another and other times they would have to produce four photons. This would contradict the conservation of energy. The new theory, which connects the wave function with probabilities for one photon, gets over the difficulty by making each photon go partly into each of the two components. Each photon then interferes only with itself. Interference between two different photons never occurs.

E. Wave function and Schrödinger Equation

In accordance with above ideas, we are led to the following formulation:

(1) For the classical concept of a trajectory, we must substitute it with the concept of a time-varying *state*. The quantum state of a particle such as the electron is characterized by a *wave function* $\psi(\vec{r},t)$, which contains all the information it is possible to obtain about the particle.

(2) $\psi(\vec{r},t)$ is interpreted as a *probability amplitude* of the particle's presence. Since the possible positions of the particle form a continuum, the probability $dP(\vec{r},t)$ of the particle being at time t in a volume element $d^3r = dx\ dy\ dz$ situated at the point $\vec{r}$ must be proportional to d^3r and therefore infinitesimal. $|\psi(\vec{r},t)|^2$ is then interpreted as the corresponding *probability density*, with $dP(\vec{r},t) = C\,|\psi(\vec{r},t)|^2\, d^3r$, where C is a normalization constant.

(3) The *principle of spectral decomposition* applies to the measurement of an arbitrary physical quantity:

— The result found must belong to a set of eigen results {a}.

— With each eigenvalue *a* is associated with an eigenstate, i.e., an eigen-function $\psi_a(\vec{r})$. This function is such that if $\psi(\vec{r},t_0) = \psi_a(\vec{r})$ (where t_0 is the time at which the measurement is performed), the measurement will always yield *a*.

— For any $\psi(\vec{r},t)$, the probability P_a of finding the eigenvalue *a* for a measurement at time t_0 is given by decomposing $\psi(\vec{r},t_0)$ in terms of the eigen-function $\psi_a(\vec{r})$:

$$\psi(\vec{r},t_0) = \sum_a c_a \psi_a(\vec{r})$$

Then we have: $P_a = \dfrac{|c_a|^2}{\sum_a |c_a|^2}$. The presence of the denominator insures that the total probability is equal to 1: $\sum_a P_a = 1$.

— If the measurement indeed yields a, the wave function of the particle immediately after the measurement is $\psi'(\vec{r},t_0) = \psi_a(\vec{r})$.

(4) The equation describing the evolution of the function $\psi(\vec{r},t)$ remains to be written. This fundamental equation is called the *Schrödinger equation*. When the particle (of mass m) is subjected to the influence of a potential $V(\vec{r},t)$, the Schrödinger equation takes on the form:

$$i\hbar \frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\Delta\psi(\vec{r},t) + V(\vec{r},t)\psi(\vec{r},t)$$

where Δ is the Laplacian operator $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.

F. Wave Packet

Wave packet is a concept introduced to describe a particle in quantum mechanics but with some analogy to the classical picture. To my opinion, it is not a necessary concept, but it is very useful to help people to establish some feeling about quantum mechanics. We introduce the wave packet here in order to quantitatively describe the Heisenberg's uncertainty principle.

Let us consider a free particle with zero potential energy. With V = 0, its Schrödinger equation becomes:

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\Delta\psi(\vec{r},t).$$

This differential equation is obviously satisfied by solutions of the form:

$$\psi(\vec{r},t) = A\exp[i(\vec{k}\cdot\vec{r}-\omega t)].$$

The principle of superposition of states tells us that every linear combination of the plane waves satisfying the above solution will also be a solution of the Schrödinger equation. Such a superposition can be written as:

$$\psi(\vec{r},t) = \frac{1}{(2\pi)^{3/2}}\int g(\vec{k})\exp[i(\vec{k}\cdot\vec{r}-\omega t)]d^3k.$$

Such a wave function, a superposition of plane waves, is called a three-dimensional "wave packet". For the sake of simplicity, we shall only study the case of a 1-D wave packet. The wave function is simplified to

$$\psi(x,t) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{+\infty} g(k)\exp[i(kx-\omega t)]dk$$

We are interested in the form of the wave packet at a given instant. If we choose this instant as the time origin, the wave function is written as:

$$\psi(x,0) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{+\infty} g(k)\exp(ikx)dk.$$

Here, g(k) is simply the Fourier transform of $\psi(x,0)$:

$$g(k) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{+\infty}\psi(x,0)\exp(-ikx)dx.$$

The validity of above wave function is not limited to the case of the free particle: whatever the potential, $\psi(x,0)$ can always be written in this form.

Let us try to understand qualitatively the behavior of $\psi(x,0)$ through the study of a very simple special case. Let $\psi(x,0)$, instead of being the superposition of an infinite number of plane waves, be the sum of only three plane waves. The wave vectors of these three waves are $k_0, k_0-\frac{\Delta k}{2}, k_0+\frac{\Delta k}{2}$, and their amplitudes are proportional to 1, ½, and ½. We then have:

$$\psi(x) = \frac{g(k_0)}{\sqrt{2\pi}}\left[e^{ik_0x}+\frac{1}{2}e^{i\left(k_0-\frac{\Delta k}{2}\right)x}+\frac{1}{2}e^{i\left(k_0+\frac{\Delta k}{2}\right)x}\right] = \frac{g(k_0)}{\sqrt{2\pi}}e^{ik_0x}\left[1+\cos\left(\frac{\Delta k}{2}x\right)\right].$$

$|\psi(x)|$ is maximum when x = 0. This is due to the fact that when x takes on this value, the three waves are in phase and interfere constructively as shown in Figure 3. As moving away from the value x = 0, the waves become more and more out of phase, and $|\psi(x)|$ decreases. The interference becomes completely destructive when the phase shift between e^{ik_0x} and $e^{i(k_0\pm\Delta k/2)x}$ is equal to ±π. $|\psi(x)|$ goes to zero when x = ± Δx/2, therefore, $\frac{\Delta k}{2}\frac{\Delta x}{2} = \pm\pi$. So we have $\Delta x\cdot\Delta k = 4\pi$.

When we return to the general wave packet, the wave function is the superposition of infinite number of plane waves within the interval $\left[k_0-\frac{\Delta k}{2}, k_0+\frac{\Delta k}{2}\right]$. Then we have

$$\Delta x \cdot \Delta k \sim 1.$$

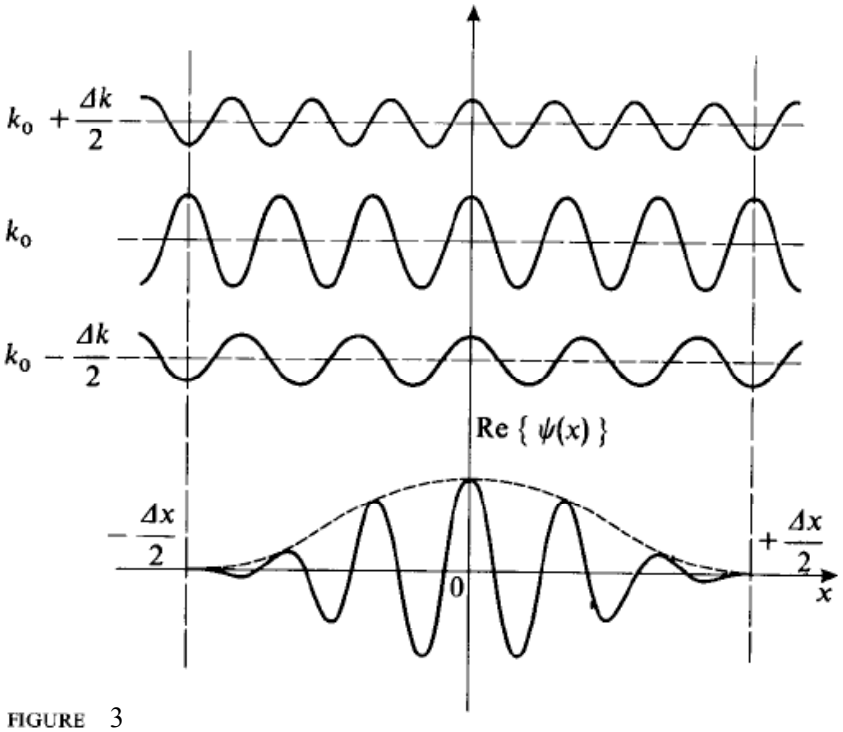

FIGURE 3

The real parts of the three waves whose sum gives the function $\psi(x)$ of (C-10). At $x = 0$, the three waves are in phase and interfere constructively. As one moves away from $x = 0$, they go out of phase and interfere destructively for $x = \pm \Delta x/2$.

In the lower part of the figure, Re $\{ \psi(x) \}$ is shown. The dashed-line curve corresponds to the function $\left[1 + \cos\left(\frac{\Delta k}{2} x\right)\right]$, which, according to (C-10), gives $|\psi(x)|$ (and therefore, the form of the wave packet).

G. Heisenberg's Uncertainty Principle (Quantitatively)

The relation derived above begins to show the uncertainty principle. To further illustrate it, let us consider a one-dimensional particle with certain momentum p_0. Since its momentum is certain, the uncertainty of its momentum $\Delta p = 0$. Then the corresponding wave function is a pure plane wave $\psi_{p_0}(x) = \exp[ip_0 x/\hbar]$. Thus, $\left|\psi_{p_0}(x)\right|^2 = 1$, i.e., the probability of the particle at any point in the space is equal and not dependent on x. In other words, the position of the particle is completely uncertain $\Delta x = \infty$.

Let us consider another example. If a one-dimensional particle has certain position, i.e., $\Delta x = 0$, then its corresponding wave function is $\psi_{x_0}(x) = \delta(x - x_0)$. The Fourier expansion gives $\psi_{x_0}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi_{x_0}(x) e^{-ipx/\hbar} dx = \frac{1}{\sqrt{2\pi\hbar}} e^{-ix_0 p/\hbar}$, therefore, $\left|\psi_{x_0}(p)\right|^2 = \frac{1}{2\pi\hbar}$, i.e., the probability of the particle to have any momentum is equal and not dependent on p. Thus, the particle momentum is completely uncertain $\Delta p = \infty$.

Recall the wave packet derived above, we have $\Delta x \cdot \Delta k \sim 1$. Using the de Broglie's relationship $p = \hbar k$, we have $\Delta x \cdot \Delta p \sim \hbar$. This is the Heisenberg's uncertainty principle. More accurate derivation gives the exact expression of the Heisenberg's uncertainty principle as

$$\Delta x \cdot \Delta p \geq \hbar / 2 .$$

Some other important expressions for the Heisenberg's uncertainty principle are

$$\Delta E \cdot \Delta t \geq \hbar / 2 .$$

III. The Equations of Motion

The equations of motion are to answer the question how the state of a system varies with time. Here the state can be expressed by a wave function or by a state vector. In 1926 Schrödinger proposed an equation of motion and successfully solved this problem. This is called Schrödinger equation.

The Schrödinger equation is a fundamental equation in quantum mechanics, just like the Newtonian equation in classical mechanics. The Schrödinger equation is a fundamental hypothesis. In principle, it cannot be derived or proven from more fundamental principles, but it can only be verified by experiments. However, we will still try to introduce the equation from a simply way to illustrate its meaning.

Let us consider the case of a free particle. The energy E and momentum p of the particle has the following relationship:

$$E = \vec{p}^{\,2} / 2m ,$$

where m is the mass of the particle. According to the de Broglie's relations, the angular frequency and wave vector of the matter wave associated with the particle are given by $\omega = E / \hbar$ and $\vec{k} = \vec{p} / \hbar$. In other words, this wave associated with the particle is a monochromatic plane wave:

$$\left|\psi(\vec{r},t)\right\rangle \propto \exp[i(\vec{k} \cdot \vec{r} - \omega t)] = \exp[i(\vec{p} \cdot \vec{r} - Et)/\hbar] .$$

Thus, we have the following relationship:

$$i\hbar \frac{\partial}{\partial t}\left|\psi\right\rangle = E\left|\psi\right\rangle,\ -i\hbar\nabla\left|\psi\right\rangle = \vec{p}\left|\psi\right\rangle, \text{ and } -\hbar^2\nabla^2\left|\psi\right\rangle = \vec{p}^{\,2}\left|\psi\right\rangle.$$

Therefore, we get:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\right)\left|\psi\right\rangle = \left(E - \frac{\vec{p}^{\,2}}{2m}\right)\left|\psi\right\rangle = 0$$

i.e.,

$$i\hbar \frac{\partial}{\partial t}\left|\psi\right\rangle = -\frac{\hbar^2}{2m}\nabla^2\left|\psi\right\rangle.$$

This is the equation of motion for a free particle. Furthermore, we consider the particle in a potential field $V(\vec{r})$. According to classical theory, the energy of the particle is given by

$$E = \vec{p}^{\,2} / 2m + V .$$

If we perform the following substitution:

$$E \rightarrow i\hbar \frac{\partial}{\partial t} \text{ and } \vec{p} \rightarrow \hat{p} = -i\hbar\nabla ,$$

then we obtain

$$i\hbar \frac{\partial}{\partial t}\left|\psi\right\rangle = \left[-\frac{\hbar^2}{2m}\nabla^2 + V\right]\left|\psi\right\rangle .$$

This is the equation of motion for the quantum mechanics, which was proposed by Schrödinger in 1926.

We may generalize the equation to the following form:

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle,$$

where H is the Hamiltonian operator that corresponds to the total energy of the system. H can be expressed in a general format as

$$H = \frac{\vec{p}^2}{2m} + V,$$

in which $\frac{\vec{p}^2}{2m}$ represents the kinetic energy of the system and V represents the potential energy of the system.

IV. Postulates of Quantum Mechanics

In this chapter, we shall study the postulates on which the quantum description of physical systems is based. We have already introduced them in a qualitative and partial way. These postulates will provide us with an answer to the following questions:

(1) How is the state of a quantum system at a given time described mathematically?

(2) Given this state, how can we predict the results of the measurement of various physical quantities?

(3) How can the state of the system at an arbitrary time t be found when the state at time t_0 is known?

1. Description of the state of a system

First Postulate: At a fixed time t_0, the state of a physical system is defined by specifying a state vector $|\psi(t_0)\rangle$ belonging to the state space.

It is important to note that, since the state space is a vector space, this first postulate implies a superposition principle: a linear combination of state vectors is still a state vector.

2. Description of physical quantities

Second Postulate: Every measurable physical quantity A is described by an operator $\hat{A}$ acting in the state space. This operator is an observable.

Unlike classical mechanics, quantum mechanics describes in a fundamentally different manner the state of a system and the associated physical quantities: a state is represented by an abstract vector, and a physical quantity is represented by an operator.

3. The measurement of physical quantities

(1) Possible Results

Third Postulate: The only possible result of the measurement of a physical quantity A is one of the eigenvalues of the corresponding observable A.

If the spectrum of A is discrete, the results that can be obtained by measuring A are quantized.

(2) Principle of Spectral Decomposition

Fourth Postulate (case of a discrete non-degenerate spectrum): When the physical quantity A is measured on a system in the normalized state $|\psi\rangle$, the probability $P(a_n)$ of obtaining the non-degenerate eigenvalue a_n of the corresponding observable A is $P(a_n) = |\langle u_n|\psi\rangle|^2$, where $|u_n\rangle$ is the normalized eigenvector of A associated with the eigenvalue a_n.

Let us assume the spectrum of A is entirely discrete. If all the eigenvalues a_n of A are non-degenerate, there is associated with each of them a unique eigenvector $|u_n\rangle$:

$$A|u_n\rangle = a_n|u_n\rangle.$$

Since A is an observable, the set of the $|u_n\rangle$, which we shall take to be normalized, constitutes a basis in the state space E, and the state vector $|\psi\rangle$ can be written:

$$|\psi\rangle = \sum_n c_n|u_n\rangle.$$

We postulate that the probability of finding a_n when A is measured is

$$P(a_n) = |c_n|^2 = |\langle u_n|\psi\rangle|^2.$$

Fourth Postulate (case of a discrete spectrum): When the physical quantity A is measured on a system in the normalized state $|\psi\rangle$, the probability $P(a_n)$ of obtaining the eigenvalue a_n of the corresponding observable A is : $P(a_n) = \sum_{i=1}^{g_n} |\langle u_n|\psi\rangle|^2$, where g_n is the degree of degeneracy of a_n and $\{|u_n^i\rangle\}$ (i = 1, 2, …, g_n) is an orthonormal set of vectors that forms a basis in the eigen-subspace E_n associated with the eigenvalue a_n of A.

Fourth Postulate (case of a continuous non-degenerate spectrum): When the physical quantity A is measured on a system in the normalized state $|\psi\rangle$, the probability $dP(\alpha)$ of obtaining a result included between α and $\alpha+d\alpha$ is equal to: $dP(\alpha) = |\langle v_\alpha|\psi\rangle|^2 d\alpha$, where $|v_\alpha\rangle$ is the eigenvector corresponding to the eigenvalue α of the observable A associate with the physical quantity A.

(3) Reduction of the State

Fifth Postulate: if the measurement of the physical quantity A on the system in the state $|\psi\rangle$ gives the result a_n, the state of the system immediately after the measurement is the normalized projection $\frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P|\psi\rangle}}$ of $|\psi\rangle$ onto the eigen-subspace associated with a_n.

4. Time evolution of systems

Sixth Postulate: The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle$$

where H(t) is the observable associate with the total energy of the system.

H is called the Hamiltonian operator of the system, as it is obtained from the classical Hamiltonian.

5. Quantization rules

We are now going to discuss how to construct, for a physical quantity A already defined in classical mechanics, the operator A that describes it in quantum mechanics.

With the position $\vec{r}(x,y,z)$ of the particle is associated with the observable $\vec{R}(X,Y,Z)$, and with the momentum $\vec{p}(p_x,p_y,p_z)$ of the particle is associated with the observable $\vec{P}(P_x,P_y,P_z)$, the components of $\vec{R}$ and $\vec{P}$ satisfy the canonical commutation relations: $[R_i,R_j]=[P_i,P_j]=0$ and $[R_i,P_j]=i\hbar\delta_{ij}$.

These postulates of quantum mechanics are consistent with the principle of uncertainty, the principle of superposition, and the principle of motion. They are a kind of expression of these principles in quantum physics language. These postulates are based on experiment results and have been confirmed by more experiments. They are the new ways for us to view and understand the real physical world.

Complement A_I

ORDER OF MAGNITUDE OF THE WAVELENGTHS ASSOCIATED WITH MATERIAL PARTICLES

De Broglie's relation:

$$\lambda = \frac{h}{p} \tag{1}$$

shows that, for a particle of mass m and speed v, the smaller m and v, the longer the corresponding wavelength.

To show that the wave properties of matter are impossible to detect in the macroscopic domain, take as an example a dust particle, of diameter $1\,\mu$ and mass $m \simeq 10^{-15}$ kg. Even for such a small mass and a speed of $v \simeq 1$ mm/s, formula (1) gives:

$$\lambda \simeq \frac{6.6 \times 10^{-34}}{10^{-15} \times 10^{-3}} \text{ meter} = 6.6 \times 10^{-16} \text{ meter} = 6.6 \times 10^{-6}\,\text{Å} \tag{2}$$

Such a wavelength is completely negligible on the scale of the dust particle.

Consider, on the other hand, a thermal neutron, that is, a neutron ($m_n \simeq 1.67 \times 10^{-27}$ kg) with a speed v corresponding to the average thermal energy at the (absolute) temperature T. v is given by the relation:

$$\frac{1}{2} m_n v^2 = \frac{p^2}{2m_n} \simeq \frac{3}{2} kT \tag{3}$$

where k is the Boltzmann constant ($k \simeq 1.38 \times 10^{-23}$ joule/degree). The wavelength which corresponds to such a speed is:

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{3m_n kT}} \tag{4}$$

For $T \simeq 300$ °K, we find :

$$\lambda \simeq 1.4\,\text{Å} \tag{5}$$

that is, a wavelength which is of the order of the distance between atoms in a crystal lattice. A beam of thermal neutrons falling on a crystal therefore gives rise to diffraction phenomena analogous to those observed with X-rays.

Let us now examine the order of magnitude of the de Broglie wavelengths associated with electrons ($m_e \simeq 0.9 \times 10^{-30}$ kg). If one accelerates an electron beam through a potential difference V (expressed in volts), one gives the electrons a kinetic energy:

$$E = qV = 1.6 \times 10^{-19}\, V \text{ joule} \tag{6}$$

($q = 1.6 \times 10^{-19}$ coulomb is the electron charge.) Since $E = \frac{p^2}{2m_e}$, the associated wavelength is equal to :

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_e E}} \tag{7}$$

that is, numerically:

$$\lambda = \frac{6.6 \times 10^{-34}}{\sqrt{2 \times 0.9 \times 10^{-30} \times 1.6 \times 10^{-19} V}} \text{ meter}$$

$$\simeq \frac{12.3}{\sqrt{V}} \text{ Å} \tag{8}$$

With potential differences of several hundreds of volts, one again obtains wavelengths comparable to those of X-rays, and electron diffraction phenomena can be observed with crystals or crystalline powders.

The large accelerators which are currently available are able to impart considerable energy to particles. This takes us out of the non-relativistic domain to which we have thus far confined ourselves. For example, electron beams are easily obtained for which the energy exceeds 1 GeV* $= 10^9$ eV (1 eV = 1 electron-volt $= 1.6 \times 10^{-19}$ joule), while the electron rest mass is equal to $m_e c^2 \simeq 0.5 \times 10^6$ eV. This means that the corresponding speed is very close to the speed of light c. Consequently, the non-relativistic quantum mechanics which we are studying here does not apply. However, the relations:

$$E = h\nu \tag{9-a}$$

$$\lambda = \frac{h}{p} \tag{9-b}$$

remain valid in the relativistic domain. On the other hand, relation (7) must be modified since, relativistically, the energy E of a particle of rest mass m_0 is no longer $p^2/2m_0$, but instead:

$$E = \sqrt{p^2c^2 + m_0^2c^4} \tag{10}$$

In the example considered above (an electron of energy 1 GeV), $m_e c^2$ is negligible compared to E, and we obtain:

$$\lambda \simeq \frac{hc}{E} = \frac{6.6 \times 10^{-34} \times 3 \times 10^8}{1.6 \times 10^{-10}} \text{ m} = 1.2 \times 10^{-15} \text{ m} = 1.2 \text{ fermi} \tag{11}$$

(1 fermi $= 10^{-15}$ m). With electrons accelerated in this way, one can explore the structure of atomic nuclei and, in particular, the structure of the proton; nuclear dimensions are of the order of a fermi.

COMMENTS:

(*i*) We want to point out a common error in the calculation of the wavelength of a material particle of mass $m_0 \neq 0$, whose energy E is known. This error consists of calculating the frequency ν using (9-a) and, then, by analogy with electromagnetic waves, of taking c/ν for the de Broglie

wavelength. Obviously, the correct reasoning consists of calculating, for example from (10) $\left($ or, in the non-relativistic domain, from the relation $E = \frac{p^2}{2m}\right)$ the momentum p associated with the energy E and then using (9-b) to find λ.

(*ii*) According to (9-a), the frequency ν depends on the origin chosen for the energies. The same is true for the phase velocity $V_\varphi = \frac{\omega}{k} = \nu\lambda$. Note, on the other hand, that the group velocity $V_G = \frac{\mathrm{d}\omega}{\mathrm{d}k} = 2\pi\frac{\mathrm{d}\nu}{\mathrm{d}k}$ does not depend on the choice of the energy origin. This is important in the physical interpretation of V_G.

Homework: Questions for self-study of fundamentals of quantum mechanics

1. Blackbody radiation
(1) What is this phenomenon?
(2) How did Planck explain it?
(3) What does it mean to the quantum theory of light?

2. Photoelectric effect
(1) What is it (describing the phenomenon)?
(2) How did Einstein explain it?
(3) What's its meaning to the wave-particle duality of light?

3. Compton scattering
(1) What is it (describing the phenomenon)?
(2) How did Compton explain it?
(3) What's its meaning to the wave-particle duality of light?

4. Young's double-slit interference experiment with light (please check with optic books)
(1) How does traditional light theory describe and explain the phenomenon?
(2) What determine the fringe interval and resolution?
(3) What dimensions of the slits and light source (both geometry and spectrally) are needed in order to produce the fringes?
(4) How do you explain the Young's interference from quantum mechanics point of view?

5. Fabry-Perot etalon is a very useful filter in optical remote sensing, including the lidar remote sensing. It uses multiple beam interference to produce very narrow fringes so that it can help compress the background light to improve the signal-to-noise ratio.
(1) Please check with optics books to find out how it works in traditional optics point of view, i.e., derive the multiple beam interference equations, and find out what determine the finess, free spectral range, and resolution.
(2) In reality, the return light for optical remote sensing (e.g., the airglow from the mesopause region) is very weak – photons nearly arriving at the etalon one by one. In this case, will you still have the etalon fringes? How do you explain it from the point of view of quantum mechanics?

6. What are the major points that quantum mechanics is essentially different from classical mechanical? Are there points that they are similar?

February 1, 2006

Notation and Representation in Quantum Mechanics

Professor Xinzhao Chu

Quantum mechanics (QM) uses some unique notations to present its mathematical formalism. Due to historical reasons along the development path of QM, different scientists also used different representations in QM and its application. Even today, different notations and representations are still mixed in modern QM research and applications. These caused lots of confusions for beginners. Apparently, we have felt this confusion in the class on Tuesday. For example, I prefer using Dirac notation, i.e., the vectors (ket or bra) to represent the state of a physical system, however, a Schrödinger wavefunction (as a function of the coordinates) is used when we try to solve the eigenvalue equation for a hydrogen atom. What is the relationship between the Dirac state vector and the Schrödinger wavefunction? To help understand these notation and representations, I put this short note together and hope to give you a clear "picture" about how to understand QM's mathematical formalism.

I. Mathematical Formalism of Quantum Mechanics

The state of a physical system is about the status of a motion of the physical system. Any physical system is undergoing a motion that can be undisturbed or disturbed. Each motion of the physical system possesses physical variables like energy, momentum, coordinates and angular momentum, etc, although they may not be determined to infinite accuracy in QM. Each undisturbed motion with information about energy, momentum, coordinates, angular momentum, etc forms a state of the physical system at an instant time. (The disturbed motion or the change of motion belongs to the problem of time evolution of the state.) In the mathematical formalism of quantum mechanics, the *state* of any physical system at an instant time is represented by a *state vector* $|\psi(t)\rangle$ in an abstract space, called the *state space* that is formed by all the possible states of the physical system. The length of the state vector is not important, but different directions of each vector define different states of the physical system. Because they are vectors, the linear combination of these vectors is still a vector, belonging to the same state space. This represents the principle of superposition of states – a unique and fundamental feature of quantum mechanics system. The state vector is an abstract and universal expression of the state of a physical system, which does not depend on any detailed representation.

Each physical system possesses many physical variables, but not all of them correspond to real numbers, i.e., not all of them are measurable. In QM, we only care about the physical quantities that are measurable. So measurable physical quantities must be real physical variables. Further study shows that the real physical variables whose eigenstates do not form complete sets are not measurable quantities. Therefore, the measurable physical quantities are real physical variables whose eigenstates form a complete set. Such physical variables (i.e., measurable physical quantities) are called *observables*. In the mathematical formalism of quantum mechanics, an observable (measurable physical quantity) is represented by an *operator* $\hat{A}$. When the operator $\hat{A}$ acts on a state vector $|\psi(t)\rangle$ in the state space, one of the eigenvalues corresponding to the

measurable physical quantity will be obtained. This interaction of an operator on a state vector corresponds to a measurement procedure. This is also a unique feature of quantum mechanics, i.e., the disturbance of measurements on a physical system is not negligible and further implying the indeterminacy and principle of uncertainty.

The time evolution of the state of a physical system is governed by the equations of motion, i.e., the *Schrödinger equation*. Thus, the state vector $|\psi(t)\rangle$ obeys:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle.$$

The method described above is the basic mathematical formalism for quantum mechanics. In the following the Dirac notation is introduced to describe this formalism.

II. Dirac Notation for Quantum Mechanics

1. Ket vector

Dirac notation uses a *ket vector* as the state vector. It is represented by the symbol $|\ \rangle$, inside which is placed a distinctive sign which enables us to distinguish the corresponding ket vector from all others, e.g., $|\psi\rangle$. *A state is specified by the direction of a ket vector and any length one may assign to the ket vector is irrelevant*. All the states of the physical system are in one-one correspondence with all the possible directions for a ket vector. This notation shows up very clearly the fundamental difference between the superposition of the quantum theory and any kind of classical superposition. In the case of a classical system for which a superposition principle holds, for instance a vibrating membrane, when one superposes a state with itself, the result is a different state, with a different amplitude of the oscillations. There is no physical characteristic of a quantum state corresponding to the magnitude of the classical oscillations, as distinct from their quality. Again, while there exists a classical state with zero amplitude of oscillation everywhere, namely the state of rest, there does not exist any corresponding state for a quantum system, the zero ket vector corresponding to no state at all.

2. Bra vector

Whenever we have a set of vectors in any mathematical theory, we can always set up a second set of vectors. In Dirac notation, the new vectors are bra vectors. A general bra vector is represented by the symbol $\langle\ |$, the mirror image of the symbol for a ket vector. We shall use the same label to specify a ket and the corresponding bra. Thus, the bra corresponding to $|\psi\rangle$ will be written $\langle\psi|$. There is a one-one correspondence between the bra vectors and the ket vectors, such that the bra vector corresponding to $|\varphi_1\rangle+|\varphi_2\rangle$ is $\langle\varphi_1|+\langle\varphi_2|$, and the bra vector corresponding to $c|\psi\rangle$ is $c^*\langle\psi|$, where $c*$ is the conjugate complex number to c. The correspondence from a ket vector to a bra vector is antilinear, and can be expressed as

$$\lambda_1|\varphi_1\rangle+\lambda_2|\varphi_2\rangle \Rightarrow \lambda_1^*\langle\varphi_1|+\lambda_2^*\langle\varphi_2|.$$

3. Scalar product

The bra vector and the ket vector have a scalar product written as $\langle\varphi|\psi\rangle$, and the scalar product corresponds to a number. So any complete bracket expression denotes a number and any incomplete bracket expression denotes a vector. The bra and ket vectors are complex quantities, since they can be multiplied by complex numbers and are then of the same nature as before, but they are complex quantities of a special kind that cannot be

split up into real and pure imaginary parts. For normal complex quantities, by taking half the sum of the quantity itself and its conjugate, one can get the real part of the complex quantity. But this method cannot be applied to the bra and ket vectors since they are of different natures and cannot be added together. We will use the words 'conjugate complex' to refer to numbers and other complex quantities that can be split up into real and pure imaginary parts, and the words 'conjugate imaginary' for bra and ket vectors that cannot be split.

On account of the one-one correspondence between bra vectors and ket vectors, any state of the physical system at a particular time may be specified by the direction of a bra vector, just as well as by the direction of a ket vector. The whole quantum theory will be symmetrical in essentials between bras and kets. In ordinary space, from any two vectors one can construct a number, i.e., their scalar product, which is a real number and is symmetrical between them. In the space of bra vector or the space of ket vectors, from any two vectors one can again construct a number – the scalar product of one with the conjugate imaginary of the other, but this number is complex and goes over into the conjugate complex number when the two vectors are interchanged.

These descriptions can be summarized to the following equations:

$\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$

$\langle\varphi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1\langle\varphi|\psi_1\rangle + \lambda_2\langle\varphi|\psi_2\rangle$

$\langle\lambda_1\varphi_1 + \lambda_2\varphi_2|\psi\rangle = \lambda_1^*\langle\varphi_1|\psi\rangle + \lambda_2^*\langle\varphi_2|\psi\rangle$

$\langle\psi|\psi\rangle$ is real and positive; zero if and only if $|\psi\rangle = 0$.

4. Orthogonal and normalized

There is a kind of perpendicularity in these spaces, which is a generalization of the perpendicularity in ordinary space. We shall call a bra and a ket vector *orthogonal* if their scalar product is zero, and two bras and two kets will be called orthogonal if the scaler product of one with the conjugate imaginary of the other is zero. Furthermore, we shall say that two states of the physical system are orthogonal if the vectors corresponding to these states are orthogonal.

The length of a ket vector $|\psi\rangle$ or of the conjugate imaginary bra vector $\langle\psi|$ is defined as the square root of the positive number $\langle\psi|\psi\rangle$. When we are given a state and wish to set up a bra or a ket vector to correspond to it, only the direction of the vector is given and the vector itself is undetermined to the extent of an arbitrary numerical factor. It is often convenient to choose this numerical factor so that the vector is of length unity. This procedure is called *normalization* and the vector so chosen is said to be *normalized*.

The vector is not completely determined even then, since one can still multiply it by any number of modulus unity, i.e., any number $e^{i\gamma}$ where γ is real, without changing its length. We shall call such a number a *phase factor*.

5. Linear operator

A linear operator $\hat{A}$ associates every ket vector $|\psi\rangle$ with another ket vector, and the correspondence is linear:

$$\hat{A}|\psi\rangle = |\psi'\rangle$$

$$\hat{A}\left(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle\right) = \lambda_1\hat{A}|\psi_1\rangle + \lambda_2\hat{A}|\psi_2\rangle$$

The product of two linear operators $\hat{A}$ and $\hat{B}$ is defined as

$$\left(\hat{A}\hat{B}\right)|\psi\rangle = \hat{A}\left(\hat{B}|\psi\rangle\right).$$

In general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. The commutator of $\hat{A}$ and $\hat{B}$ is $\left[\hat{A},\hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}$.

The matrix element of $\hat{A}$ between two ket vectors is the scalar product, given by

$$\langle\varphi|\left(\hat{A}|\psi\rangle\right) = \langle\varphi|\hat{A}|\psi\rangle.$$

6. Representations in state space

Dirac notations of the state are the abstract state vectors, which do not involve any concrete representations. In actual application of QM, a representation is usually chosen to express the state vectors for the convenience of calculation. Choosing a representation means choosing an orthonormal basis, either discrete or continuous, in the state space. Vectors and operators are then represented in this basis by numbers: components for the vectors, and matrix elements for the operators. The vectorial calculus then becomes a matrix calculus with these numbers. The choice of a representation is in theory arbitrary, as different representations are equivalent in principle. In reality, it obviously depends on the particular problem being studies: in each case, one chooses the representation that leads to the simplest calculations.

A set of ket vectors, discrete $\{|u_i\rangle\}$, is said to be orthonormal if the ket vectors of this set satisfy the orthonormalization relation:

$$\langle u_i|u_j\rangle = \delta_{ij}$$

Any ket vector belonging to the state space has a unique expansion on the complete orthonormal set:

$$|\psi\rangle = \sum_i c_i|u_i\rangle.$$

From the orthonormalization relationship of $\{|u_i\rangle\}$, it is easy to obtain $c_i = \langle u_i|\psi\rangle$, which means the projection of the state vector $|\psi\rangle$ on the basis of $\{|u_i\rangle\}$. When all c_i are determined, the state vector $|\psi\rangle$ is determined. Thus, the set of $\{c_i\} = \{\langle u_i|\psi\rangle\}$ is the representative of the state vector in the $\{|u_i\rangle\}$ representation.

From above, we have

$$|\psi\rangle = \sum_i c_i|u_i\rangle = \sum_i \langle u_i|\psi\rangle|u_i\rangle = \sum_i |u_i\rangle\langle u_i|\psi\rangle = \left(\sum_i |u_i\rangle\langle u_i|\right)|\psi\rangle.$$

Therefore, we have $\sum_i |u_i\rangle\langle u_i| = 1$.

III. Schrödinger Wavefunction

The representation shown above can have many different ones. In other words, the state vectors can be projected to different spaces as long as the space has a complete set of orthonormal basis ket vectors. For example, the vectors can be projected to the space of coordinates. They can also be projected to the space of momentum. In principle, these different spaces are equivalent. The physical foundation for this is that every physical system possesses features like coordinates, momentum, energy, angular momentum, etc. None of these quantities is special. If the state vectors can be expressed in one complete set of coordinates, they can certainly be expressed in a complete set of other observables, like momentum or energy, etc. Except time, measurable physical

quantities (e.g., coordinates, momentum, energy, angular momentum) are equal in the sense of mathematically expressing the state of a physical system.

The Schrödinger wavefunction used in the Schrödinger equation is a special representative of the state vectors in the space coordinate representation. It originated from the analogy to the classical mechanics. In Newtonian mechanics, a state of a physical system is expressed by a complete set of coordinates, momentum, and time: $\psi(\vec{r},\vec{p},t)$. If the coordinates and momentum are known at one instant time, then the state of this system at this instant is determined, and its past and future states are also determined through the equations of motion. In quantum mechanics, the principle of uncertainty does not permit the simultaneous determination of the coordinates and momentum to an infinite accuracy. Therefore, the quantities used to describe a quantum state are reduced and QM allows the state of the system to be expressed in a function of either coordinates or momentum, but not simultaneously. Schrödinger chose to express the state vector in the space of coordinates, i.e., $\psi(\vec{r},t)$. When in the eigenstates of energy, this wavefunction is independent of time, so it is reduced to $\psi(\vec{r})$ that is the wavefunction used in the eigenvalue equation of hydrogen atoms.

From the point of view of Dirac notation, the Schrödinger wavefunction $\psi(\vec{r})$ can be written as

$$\psi(\vec{r}) = \langle \vec{r} | \psi \rangle,$$

where $|\vec{r}\rangle$ is a component of the complete set of coordinate eigenstates $\{|\vec{r}\rangle\}$. Therefore, the Schrödinger wavefunction $\psi(\vec{r})$ is a component of the state vector $|\psi\rangle$ on the basis vector $|\vec{r}\rangle$ of the complete set $\{|\vec{r}\rangle\}$.

The probabilistic interpretation of the Schrödinger wavefunction $\psi(\vec{r},t)$ given by Born is that the probability $P(\vec{r},t)dv$ of finding the particle in the volume element dv at the point $\vec{r}$ at time t is given by

$$P(\vec{r},t)dv = |\psi(\vec{r},t)|^2 dv = |\langle \vec{r} | \psi \rangle|^2 dv.$$

Thus, the wavefunction $\psi(\vec{r}) = \langle \vec{r} | \psi \rangle$ gives the probability density of position.

The state vector $|\psi\rangle$ can also be expressed in a space of momentum $\{|\vec{p}\rangle\}$:

$$\psi(\vec{p}) = \langle \vec{p} | \psi \rangle$$

It is interpreted as $\psi(\vec{p})$ giving the probability distribution of momentum. $\psi(\vec{p})$ is the representative of the state vector $|\psi\rangle$ in the momentum representation.

These interpretations are consistent with the fourth postulate of QM in the note of "Fundamentals of Quantum Mechanics in Laser Remote Sensing", just here $|\vec{r}\rangle$ is the eigenstate with the eigenvalue of $\vec{r}$ for $\psi(\vec{r})$, and $|\vec{p}\rangle$ is the eigenstate with the eigenvalue of $\vec{p}$ for $\psi(\vec{p})$.

The reasons we prefer Dirac notation are twofold: (1) since it is independent of representations, it simplifies the mathematical formalism of QM; (2) it permits the generalization of the formalism in order to include physical quantities that cannot be given by a wavefunction, e.g., the spin.