

CHAPTER 1

1. INTRODUCTION

Quantum mechanics is the theory of **atomic** and **nuclear** systems. It has been developed out of classical physics, particularly the two main branches, Newtonian mechanics and Maxwell's electromagnetic theory. We start by outlining the concepts of classical theory. We then show how these concepts proved quite inadequate to describe atomic systems, and state the rules of thumb which were superposed on classical theory by Planck, Bohr, and de Broglie, to make up what is usually called **Old Quantum Theory**.

3. Inadequacy of Classical Theory

In summary, classical physics has completely failed to explain the following phenomena:

- | | |
|-----------------------------|---------------------------|
| (I) Blackbody radiation. | (II) Atomic line spectra. |
| (III) Photoelectric effect. | (IV) Compton effect. |

(c) Planck's Law

Planck succeeded in 1900 in avoiding the ultraviolet catastrophe and proposed description of blackbody radiation. In sharp contrast to Rayleigh's assumption that a wave can exchange **any** amount (continuum) of energy with matter, Planck considered energy exchange between radiation and matter must be **discrete**. He then **postulated** that energy of the radiation (of frequency ν) emitted from the walls of the cavity must be integer multiplied by $h\nu$:

$$E = nh\nu, \quad n = 1, 2, 3, \dots \quad (1.6)$$

h is a universal constant, and $h\nu$ is the energy of a "**quantum**" of radiation,

ν represents frequency of the oscillating charge in the cavity's walls as well as the frequency of the radiation emitted from the walls.

In other words, Planck said that $E = nh\nu$, so that E can only certain allowed values as shown in Figure 1.3:

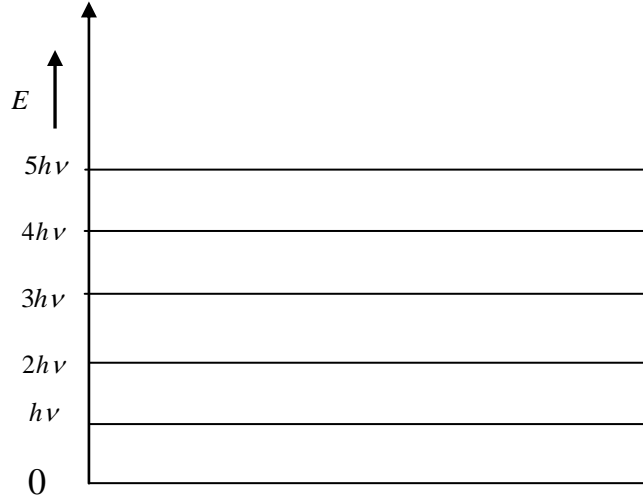


FIG. 1.3 Quantization of energy.

Planck used Boltzmann distribution to give the number N of oscillators with energy $nh\nu$, $N = N_0 e^{(-nh\nu/kT)}$. This expression gives the total of oscillators as $\sum_{n=0}^{\infty} N_0 e^{(-nh\nu/kT)}$. Since each oscillator has an energy $nh\nu$, the total energy equals to $\sum_{n=0}^{\infty} nh\nu N_0 e^{(-nh\nu/kT)}$. This total energy divided by the total number of oscillators yields the average energy of an oscillator:

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} nh\nu N_0 \exp(-nh\nu/kT)}{\sum_{n=0}^{\infty} N_0 \exp(-nh\nu/kT)} = \frac{0 + h\nu \exp(-h\nu/kT) + 2h\nu \exp(-2h\nu/kT) + \dots}{1 + \exp(-h\nu/kT) + \exp(-2h\nu/kT) + \dots}$$

where $\exp(x)$ means e^x . If we let $y = \exp(-h\nu/kT)$, we then have

$$\langle E \rangle = h\nu y \left(\frac{1 + 2y + 3y^2 + \dots}{1 + y + y^2 + \dots} \right) = h\nu y \left[\frac{(1-y)^{-2}}{(1-y)^{-1}} \right] = \frac{h\nu y}{1-y} = \frac{h\nu}{(1/y - 1)},$$

$$\Rightarrow \langle E \rangle = \frac{h\nu}{e^{h\nu/kT} - 1} \quad (1.7)$$

Planck used Eq. (1.7) instead of the energy-equipartition average of kT which Rayleigh and Jeans had used. The result was

$$u(\nu) d\nu = \bar{E} G(\nu) d\nu = \frac{8\pi\nu^3}{c^3} \frac{h}{e^{h\nu/kT} - 1} d\nu \quad (1.8)$$

which agree with the experimental findings (see FIG. 1.1).

(II) Atomic Structure and Spectroscopy

(a) The Bohr Model for hydrogen atom

The failure of classical concepts when applied to the motion of electrons appeared most clearly in connection with the hydrogen atom. The experiments of Rutherford showed that an atom can be regarded as a negatively charged electron orbiting around a relatively massive, positively charged nucleus (for hydrogen a single proton). With the neglect of radiation, this system is exactly analogous to the motion of a planet round the sun, with the gravitational attraction between the masses being replaced by the Coulomb attraction between the charges. Then Bohr postulated two main rules:

(i) The magnitude of the angular momentum L of the electron is an integer multiple of \hbar , where $\hbar = h/2\pi$.

$$L = n\hbar, \quad n = 1, 2, 3, \dots \quad (1.9)$$

This requirement of discrete values for the angular momentum immediately leads to discrete values of the energy E_n .

(ii) The radiation occurs when the electron makes discontinuous jumps from an orbit of energy E_n to one of energy $E_{n'}$, say, and the resulting angular frequency $\omega_{nn'}$ is determined by

$$\hbar\omega_{nn'} = |E_n - E_{n'}|. \quad (1.10)$$

We now apply these rules to the hydrogen atom whose electron, mass m , revolves around a nucleus (taken as fixed) in a circular orbit of radius r , with angular velocity ω .

Centripetal force=Coulomb force.

$$\frac{mv^2}{r} = \frac{e_M^2}{r^2}, \quad \text{where} \quad e_M^2 = \frac{e^2}{4\pi\epsilon_0}$$

$$mr\omega^2 = \frac{e_M^2}{r^2} \Rightarrow \omega^2 = \frac{e_M^2}{mr^3} \quad (1.11)$$

Bohr's condition (i) is

$$L = I\omega = mr^2\omega = n\hbar, \quad n = 1, 2, \dots$$

$$\text{then } \omega = \frac{n\hbar}{mr_n^2} \Rightarrow \omega^2 = \frac{n^2\hbar^2}{m^2r_n^4}, \quad (1.12)$$

Equating Eq. (1.11) with Eq. (1.12), leads to

$$r_n = \frac{\hbar^2 n^2}{me_M^2} = a_0 n^2 \quad (1.13)$$

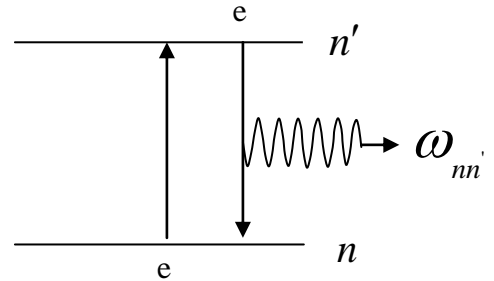
where the Bohr radius is $a_0 = \frac{\hbar^2}{me_M^2}$ which is the radius of lowest orbit.

$$\begin{aligned} E_n &= KE + PE = \frac{1}{2}mv^2 - \frac{e_M^2}{r} = \frac{1}{2}mr^2\omega^2 - \frac{e_M^2}{r} \\ &= \frac{1}{2}r(mr\omega^2) - \frac{e_M^2}{r} = \frac{1}{2}\frac{e_M^2}{r} - \frac{e_M^2}{r}, \\ E_n &= -\frac{1}{2}\frac{e_M^2}{r_n} = -\frac{1}{2}\frac{e_M^2}{a_0}\frac{1}{n^2} \Rightarrow E_1 = -\frac{e_M^2}{2a_0} \end{aligned} \quad (1.14)$$

The possible angular frequencies of radiation are

$$\omega_{nn'} = \frac{E_n - E_{n'}}{\hbar} = \frac{1}{2}\frac{e_M^2}{\hbar a_0} \left(\frac{1}{n^2} - \frac{1}{n'^2} \right). \quad (1.15)$$

The series with $n = 2$ and $n' = 3, 4, 5$ has four lines which lie in the visible region and form part of Balmer series.



CHAPTER 2

OPERATORS and WAVE FUNCTIONS

2.1 Definitions and Operator Equations

Before proceeding to the setting up of quantum mechanics, we now develop that part of the mathematical theory of operators which plays an essential role in the later work. Loosely speaking, an operator, which we denote by \hat{A} , is any mathematical entity which operates on any function of x , say, and turns it into another function. The simplest example of an operator is to take \hat{A} itself to be a function of x , the operation being multiplication. Thus we might have

$$\hat{A}(x) = x. \quad (2.1)$$

The operator x operating on *any* function $\Psi(x)$ produces the new function $x\Psi(x)$. In other words $\hat{A}\Psi(x) = x\Psi(x)$

A less trivial example is the operation of differentiation, the operator \hat{A} being any function of $\partial/\partial x$, i.e., $\hat{A} = \partial/\partial x$

Then, for *any* function $\Psi(x)$,

$$\hat{A}\Psi = \frac{\partial\Psi}{\partial x} \quad (2.2)$$

Another example is $\hat{A} = \frac{\partial}{\partial x} x$, then

$$\begin{aligned} \hat{A}\Psi &= \left(\frac{\partial}{\partial x} x \right) \Psi = \frac{\partial}{\partial x} (x\Psi) = \Psi + x \frac{\partial\Psi}{\partial x}, \\ \therefore \hat{A}\Psi &= (1 + x \frac{\partial}{\partial x}) \Psi \end{aligned} \quad (2.3)$$

Since the final equality (2.3) is valid for *any* Ψ , we can cancel the factor Ψ on the right, and write the operator equation,

$$\left(\frac{\partial}{\partial x} x\right) = \left(1 + x \frac{\partial}{\partial x}\right) \quad (2.4)$$

In general, an operator equation,

$$\hat{A} = \hat{B} + \hat{C},$$

implies that $\hat{A}\Psi = \hat{B}\Psi + \hat{C}\Psi$.

2.2 The Eigenvalue Equation

To each operator $\hat{A}(x, \partial/\partial x)$ belong a set of numbers, a_n , and a set of functions, $\Psi_n(x)$, defined by the equation,

$$\hat{A} \Psi_n(x) = a_n \Psi_n(x), \quad (2.5)$$

where a_n is an **eigenvalue**, and $\Psi_n(x)$ is the corresponding **eigenfunction**. The eigenfunctions of an operator are thus those special functions which remain unaltered under the operation of the operator, apart from multiplication by the eigenvalue. The above equation is the eigenvalue equation of the operator \hat{A} .

2.3 Commutation Relations

We consider the successive operation of two operators. We define the

commutator of two operators \hat{A} and \hat{B} to be

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}, \quad (2.8)$$

which is the difference between operating first with \hat{B} and then with \hat{A} , and first with \hat{A} and then with \hat{B} . In general this is not zero,

$$[\hat{A}, \hat{B}] \neq 0, \quad (2.9)$$

but is some new operator. To show this, it is sufficient to consider the simple case

$$\hat{A} = x, \quad \hat{B} = \frac{\partial}{\partial x}. \quad (2.10)$$

Then for any $\Psi(x)$,

$$\begin{aligned} \left[x, \frac{\partial}{\partial x} \right] \Psi(x) &= \left(x \frac{\partial}{\partial x} - \frac{\partial}{\partial x} x \right) \Psi(x) = \left(x \frac{\partial \Psi}{\partial x} - \frac{\partial}{\partial x} x \Psi \right) = \left(x \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial}{\partial x} x - x \frac{\partial \Psi}{\partial x} \right) \\ \therefore \left[x, \frac{\partial}{\partial x} \right] \Psi(x) &= -\Psi(x) \end{aligned} \quad (2.11)$$

Since this is true for any $\Psi(x)$, we have the operator equation

$$\left[x, \frac{\partial}{\partial x} \right] = -1. \quad (2.12)$$

An equation which determines the commutator of two operators is called a commutation relation. The special cases, such as (2.12), in which the commutator of two operators is a number, play a particularly important role in the subsequent theory.

2.4 Wave Functions

The quantity with which quantum mechanics is concerned is the wave function Ψ of a body. While Ψ itself has no physical interpretation, the square of its absolute magnitude $|\Psi|^2$ evaluated at a particular place at a particular time is proportional to the probability of finding the body there at that time. The linear momentum, angular momentum, and energy of the body are other quantities that can be established from Ψ . The problem of quantum mechanics is to determine Ψ for a body when its freedom of motion is limited by the action of external forces.

Wave functions are usually complex with both real and imaginary parts. A probability, however, must be a positive real quantity. The probability density $|\Psi|^2$ for a complex Ψ is therefore taken as the product $\Psi^* \Psi$ of Ψ and its complex conjugate Ψ^* . The complex conjugate of any function is obtained

by replacing $i (= \sqrt{-1})$ by $-i$ wherever it appears in the function. Every complex function Ψ can be written-in the form

$$\text{Wave function } \Psi = A + iB$$

where A and B are real functions. The complex conjugate Ψ^* of Ψ is

$$\Psi^* = A - iB, \quad (\text{complex conjugate})$$

$$\text{and so } |\Psi|^2 = \Psi^* \Psi = A^2 - i^2 B^2 = A^2 + B^2,$$

since $i^2 = -1$. Hence $\Psi^* \Psi$ is always a positive real quantity, as required.

2.6 Expectation Values

How to extract information from a wave function

Once Schrödinger's equation has been solved for a particle in a given physical situation, the resulting wave function $\Psi(x, y, z, t)$ contains all the information about the particle that is permitted by the uncertainty principle. Except for those variables that are quantized this information is in the form of probabilities and not specific numbers.

As an example, let us calculate the **expectation value** $\langle x \rangle$ of the position of a particle confined to the x -axis that is described by the wave function $\Psi(x, t)$. This is the value of x we would obtain if we measured the positions of a great many particles described by the same wave function at some instant t and then averaged the results.

To make the procedure clear, we first answer a slightly different question: What is the average position x of a number of identical particles distributed along the x -axis in such a way that there are N_1 particles at x_1 , N_2 particles at x_2 , and so on? The average position in this case is the same as the center of mass of the distribution, and so

$$\bar{x} = \frac{N_1 x_1 + N_2 x_2 + N_3 x_3 + \dots}{N_1 + N_2 + N_3 + \dots} = \frac{\sum N_i x_i}{\sum N_i} \quad (2.14)$$

When we are dealing with a single particle, we must replace the number N_i of particles at x_i by the probability P_i that the particle be found in an interval dx at x_i .

This probability is

$$P_i = |\Psi_i|^2 dx$$

where Ψ , is the particle wave function evaluated at $x = x_i$. Making this substitution and changing the summations to integral, we see that the expectation value of the position of the single particle is

$$\langle x \rangle = \frac{\int_{-\infty}^{\infty} x |\Psi|^2 dx}{\int_{-\infty}^{\infty} |\Psi|^2 dx} \quad (2.15)$$

2.7 Hermitian Operators

When an operator satisfy the following relation, it is called Hermitian operator

$$\int_{-\infty}^{\infty} \Psi_1^* \hat{A} \Psi_2 dx = \int_{-\infty}^{\infty} (\Psi_1 \hat{A})^* \Psi_2 dx \quad (2.16)$$

Now let's explore whether $\hat{A} \equiv \frac{d}{dx}$ is Hermitian or not.

We substitute \hat{A} in the left side of equation (2.16) and integrate by parts, then we get:

$$\int_{-\infty}^{\infty} \Psi_1^* \frac{d}{dx} \Psi_2 dx = [\Psi_1^* \Psi_2]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \Psi_2 \left(\frac{d}{dx} \Psi_1^* \right) dx,$$

$$[\Psi_1^* \Psi_2]_{-\infty}^{\infty} = 0$$

$$\therefore \int_{-\infty}^{\infty} \Psi_1^* \frac{d}{dx} \Psi_2 dx = - \int_{-\infty}^{\infty} \Psi_2 \left(\frac{d}{dx} \Psi_1^* \right) dx.$$

Since the d/dx operator does not satisfy Eq.(2.16), it is not Hermitian operator.

Theorem: The eigenvalues of Hermitian operators are always real.

Suppose that $\Psi_i(x)$ is one of the eigenvalues of the Hermitian operator \hat{A} with an eigenvalue a_i , i.e., :

$$\hat{A}\Psi_i(x) = a_i\Psi_i(x) \quad (2.17)$$

Taking the complex conjugate:

$$\hat{A}^*\Psi_i^*(x) = a_i^*\Psi_i^*(x) \quad (2.18)$$

We multiply Eq.(2.17) by Ψ_i^* from the left side and Eq.(2.18) by Ψ_i from the right side, and we integrate for all the accepted values of x :

$$\int_{-\infty}^{\infty} \Psi_i^* \hat{A} \Psi_i dx = \int_{-\infty}^{\infty} a_i \Psi_i^* \Psi_i dx,$$

$$\int_{-\infty}^{\infty} \hat{A}^* \Psi_i^* \Psi_i dx = \int_{-\infty}^{\infty} a_i^* \Psi_i^* \Psi_i dx,$$

Since that \hat{A} is a Hermitian operator, the left sides of the last two equations are equal

$$\int_{-\infty}^{\infty} a_i \Psi_i^* \Psi_i dx = \int_{-\infty}^{\infty} a_i^* \Psi_i^* \Psi_i dx,$$

$$(a_i - a_i^*) = 0$$

$$\therefore a_i = a_i^* \quad (2.19)$$

which means that a_i is real.

CHAPTER 3

QUANTUM MECHANICS

3.1 Operations of Observation

Classical mechanics breaks down when applied to small systems, provided they are small enough. It can be applied satisfactorily to the stars in their courses and to the motion of a golf ball, but fails completely when applied to atoms. It would appear that when we get down to atomic systems the word "small" takes on an absolute, not just a relative, meaning. The understanding of the absolute significance of smallness is the basic clue to the understanding of quantum mechanics.

In the physics of classical systems—systems to which classical concepts can be applied successfully—it is tacitly assumed that the operations of observation do not appreciably disturb their motion. In applying Maxwell's equations for instance, it is assumed that the currents and fields involved can all be measured without altering their values, or upsetting the development of the observed system. More precisely, it is assumed that any disturbances which are caused by measurement (the variation in current, for example, due to applying a voltmeter) can be corrected for exactly, at least in principle

The simplest type of observation is to look at something. This involves shining light on it which, as we have seen, means striking it with photons. If the position is to be determined accurately, the wavelength of the corresponding waves must be sufficiently short, their frequency correspondingly high, and the momentum of the photons consequently above a certain limit. A blow with such a photon may appreciably disturb the observed system if it is small enough. It is conceivable that these disturbances could also be allowed for, but if not, we have immediately an absolute meaning to size. This idea has been expressed precisely by Dirac, who postulated in general that **"there is a limit to the fineness of our powers of**

observation and the smallness of the accompanying disturbance—a limit which is inherent in the nature of things, and can never be surpassed by improved technique". If the system is large enough for these unavoidable disturbances to be negligible, the tacit assumption of classical physics applies, and the system may be expected to obey classical laws. If, on the other hand, the system is such that these disturbances are appreciable, it is "small" in an absolute sense, and a new theory is required for dealing with it.

The quantum physicist thus appears, if not as a bull in a china shop, at least as a man with his eyes shut, liable to knock down anything he touches, trying to obtain a clear picture of the delicate objects which surround him. Our problem is to set up a physical theory of information collected in this clumsy manner. The surprising thing is that it can be done at all, not that it takes a form which is fundamentally different from classical theory.

The first point is that since the operations of observation affect the physical systems, they may be expected to appear explicitly in the theory. These operations have two main properties:

(i) To each type of observation (e.g. observation of energy, momentum or position) there belongs a set of numbers -the possible results of the observation. We know already, from the energy levels of hydrogen, that these numbers may run over a continuous range, as in classical theory, or take on a set of discrete values.

(ii) Suppose we have two types of observation A and B . (For example, A might mean observation of position and B that of momentum.) We denote observation B followed by observation A , by AB . Then BA denotes the same types of observation carried out in the opposite order. Since each observation may disturb, and hence affect, the result of the other, the two procedures may well yield different results.

We write this symbolically as

$$AB - BA \neq 0$$

The value of this expression must be related to the magnitude of the unavoidable disturbances. It is at this point, and with this interpretation, that we expect some new constant to enter the theory, to give a quantitative rather than a qualitative meaning to our absolute definition of smallness. From our experience with Old Quantum Theory it is an obvious conjecture that this new constant will turn out to be Planck's constant, \hbar .

3.2 Operators and Observations: Interpretive Postulates

The reader will hardly have failed to notice that the physical properties of observations correspond exactly to the mathematical properties of operators developed in Chapter 2. To each belongs a set of numbers, and the effect of any pair may depend on the order in which they are applied. We thus make the general assumption that the observations A are represented by operators, \hat{A} , there being one operator for each observable property—the energy, the position, etc. The functions on which the operators operate represent the state of the system, and are known as state functions (or wave functions). If the state function is an eigenfunction it is referred to as an eigenstate. More precisely we make the following interpretive assumptions (to be discussed below):

I(i): The possible results of an observation \hat{A} are the corresponding eigenvalues a_n .

I(ii): An observation \hat{A} on a system in an eigenstate Ψ_n certainly leads to the result a_n .

I(iii): The average value of repeated observations \hat{A} on a set of systems, each one in an arbitrary state $\Psi(x)$, is

$$\bar{a} = \frac{\int_{-\infty}^{\infty} \Psi^*(x) \hat{A} \Psi(x) dx}{\int_{-\infty}^{\infty} \Psi^*(x) \Psi(x) dx}. \quad (3.1)$$

The first of these assumptions is almost inevitable, since we must clearly identify the numbers associated with the operation \bar{A} , namely, the results of the observation, with the numbers associated with the corresponding operator \hat{A} . The assumption I(ii) is also very plausible since there is a close parallel between the structure of the eigenvalue equation (2.5) and the ideal physical observation, which is an operation, \hat{A} , on the system Ψ_n , which leaves it unchanged apart from the production of a number—the result of the observation, a_n .

The postulate I(iii) deals with a more difficult situation. If the system is in a general state $\Psi(x)$, an observation \hat{A} must, by I(i), have as its result one of the eigenvalues of \bar{A} . Repeated observations \hat{A} on a set of systems, each in the state $\Psi(x)$, will produce a statistical distribution of the different eigenvalues, and I(iii) asserts what the average value of this distribution will be. The average must be some number constructed from the operator \hat{A} and the state $\Psi(x)$. It must further be consistent with I(ii). In the special case

$$\Psi(x) = u_n(x) \quad (3.2)$$

$$\bar{a} = \frac{\int_{-\infty}^{\infty} u_n^*(x) \hat{A} u_n(x) dx}{\int_{-\infty}^{\infty} u_n^*(x) u_n(x) dx} = a_n. \quad (3.3)$$

3.3 Physical Postulates

The interpretive postulates given above set up the machinery for a mathematical representation of quantum observations (accompanied by inevitable disturbances).

We now come to two postulates with more direct physical content.

(a) The Correspondence Principle

It is clear that there is one condition that quantum mechanics must satisfy. In the limit of the observed systems becoming large and the disturbances becoming negligible, it must go over into classical mechanics. To ensure this we make the first physical postulate.

P(i): The essentially definitive relations between physical variables in classical mechanics, which do not involve derivatives, are also satisfied by the corresponding quantum operators. Thus if \hat{x} and \hat{p} are the position and momentum operators then the operator for the z-component of angular momentum, for example, is

$$\hat{l}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \quad (3.4)$$

For a particle, of definite mass, in a classical potential $V(x)$, the energy operator (Hamiltonian) in terms of the position and momentum operators is the sum of the kinetic and potential energy terms,

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}). \quad (3.5)$$

In particular, the Hamiltonian for a quantum harmonic oscillator of angular frequency ω is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m}{2}\omega^2\hat{x}^2. \quad (3.6)$$

(b) The Complementarity Principle

We must now make more precise the general considerations at the end of 3.1. If \hat{A} and \hat{B} represent the observation of particular observables, the inequality $[\hat{A}, \hat{B}] \neq 0$ means physically that there may be a mutual disturbance between the two observations. This must be replaced by some equality for particular observables.

The notion of a photon in Old Quantum Theory and the argument given in 3.1 suggest that there is a direct connection between the mutual disturbances of position and momentum measurements, and that these should be related to \hbar . We thus postulate

$$[\hat{x}, \hat{p}] = \alpha \hbar, \quad (3.7)$$

where α is a number to be determined. The simplest, but not the only, representation of the operator x is an ordinary algebraic variable. We thus replace x by x ,

$$\hat{x} \rightarrow x \quad (3.8a)$$

$$\text{According to } \left[x, \frac{\partial}{\partial x} \right] = -1,$$

$$\left[x \frac{\partial}{\partial x} - \frac{\partial}{\partial x} x \right] = -1,$$

Multiplying by $-\alpha \hbar$,

$$\begin{aligned} (-\alpha \hbar) x \frac{\partial}{\partial x} - (-\alpha \hbar) \frac{\partial}{\partial x} x &= \alpha \hbar \\ \left[x, -\alpha \hbar \frac{\partial}{\partial x} \right] &= \alpha \hbar, \end{aligned} \quad (3.9)$$

Thus if x is represented by (3.8a), hence (3.7) suggests that

$$\hat{p} \rightarrow -\alpha \hbar \frac{\partial}{\partial x} \quad (3.8b)$$

The eigenvalue equation, (2.5), for the momentum operator with eigenvalue

$$\hat{p} \Psi_p(x) = \left(-\alpha \hbar \frac{\partial}{\partial x} \right) \Psi_p(x) = p \Psi_p(x). \quad (3.10)$$

The eigenfunctions are thus

$$\Psi_p = \exp \left[-\frac{px}{\alpha \hbar} \right] \quad (3.11)$$

If we take

$$\alpha = i \quad (3.12)$$

We then have the space part of a de Broglie wave $\Psi = A e^{-(i/\hbar)(Et - px)}$ appearing automatically as the state function of a particle of definite momentum. This is just the type of fundamental relation between particle and wave which is required. We therefore make the second physical postulate,

$$\mathbf{P(ii)} \quad [\hat{x}, \hat{p}] = i\hbar \quad (3.13)$$

This leads directly, from (3.8a) and (3.8b), to the important representation of these operators.

$$\begin{aligned} \hat{x} &\rightarrow x \\ \hat{p} &\rightarrow -i\hbar \frac{\partial}{\partial x} \end{aligned}$$

(3.14)

(3.14)

which is known as the Schrödinger representation.

3.4 The Schrödinger Equation: Time-Dependent Form

A basic physical principle that cannot be derived from anything else

In quantum mechanics the wave function Ψ corresponds to the wave variable y in the wave equation $y = A e^{-i(\omega t - kx)}$ of wave motion in general. However, Ψ , unlike y ,

is not itself a measurable quantity and may therefore be complex. For this reason we assume that Ψ for a particle moving freely in the $+x$ direction is specified by

$$\Psi = A e^{-i(\omega t - kx)} \quad (3.15)$$

where $k = \frac{2\pi}{\lambda}$ is the wave number and $\omega = 2\pi\nu$ is the angular frequency. This is convenient since we already know what ω and k are in terms of the total energy E and momentum p of the particle being described by Ψ . Because $E = \hbar\omega$ and according to de Broglie postulate: $\lambda = \frac{h}{p} \Rightarrow \frac{\lambda}{2\pi} = \frac{h/2\pi}{p} \Rightarrow \frac{1}{k} = \frac{\hbar}{p}$ or $p = k\hbar$, so we can write the energy and momentum as

$$\begin{aligned} E &= \hbar\omega \\ p &= k\hbar \end{aligned}$$

(3.16)

Substituting Eq. (3.16) in Eq. (3.15)

$$\Psi = A e^{-(i/\hbar)(Et - px)} \quad (3.17)$$

Equation (3.17) describes the wave equivalent of an unrestricted particle of total energy E and momentum p moving in the $+x$ direction, just as a harmonic displacement wave moving freely along a stretched string.

The expression for the wave function Ψ given by Eq. (3.17) is correct only for freely moving particles. However, we are most interested in situations where the motion of a particle is subject to various restrictions. An important concern, for example, is an electron bound to an atom by the electric field of its nucleus. What we must now do is to obtain the fundamental differential equation for Ψ , which we can then be solved for Ψ in a specific situation. This equation, which is Schrödinger's equation, can be arrived at in various ways, but it *cannot* be rigorously derived from existing physical principles: **the equation represents something new**. What will be done here is to show one route to the wave equation for Ψ and then to

discuss the significance of the result. We begin by differentiating Eq. (3.17) for Ψ twice with respect to x , which gives

$$\frac{\partial^2 \Psi}{\partial x^2} = -\frac{p^2}{\hbar^2} \Psi \quad \Rightarrow \quad p^2 = -\hbar^2 \frac{\partial^2 \Psi}{\partial x^2} \quad (3.18)$$

Differentiating Eq. (3.17) once with respect to t gives

$$\frac{\partial \Psi}{\partial t} = -\frac{iE}{\hbar} \Psi \quad \Rightarrow \quad E\Psi = i\hbar^2 \frac{\partial \Psi}{\partial t} \quad (3.19)$$

At speeds small compared with that of light, the total energy E of a particle is the sum of its kinetic energy T and its potential energy V , where V is in general a function of position x and time t :

$$E = T + V = \frac{p^2}{2m} + V(x, t) \quad (3.20)$$

The function V represents the influence of the rest of the universe on the particle. Of course, only a small part of the universe interacts with the particle to any extent; for instance, in the case, of the electron in a hydrogen atom, only the electric field of the nucleus must be taken into account.

Multiplying both sides of Eq. (3.21) by the wave function Ψ gives

$$\frac{p^2 \Psi}{2m} + V\Psi = E\Psi \quad (3.21)$$

Now we substitute for $E\Psi$ and $p^2\Psi$ from Eqs. (3.18) and (3.19) to obtain the time-dependent form of Schrödinger's equation

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad (3.22)$$

In three dimensions the time-dependent form of Schrödinger's equation is

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad (3.23)$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ and the particle's potential energy V is some function of x , y , z , and t .

3.5 The Schrödinger Equation: Time-Independent Form

In a great many situations the potential energy of a particle does not depend on time explicitly; the forces that act on it, and hence V , vary with the position of the particle only. When this is true, Schrödinger's equation may be simplified by removing all reference to t .

We begin by noting that the one-dimensional wave function Ψ of an unrestricted particle may be written

$$\Psi = Ae^{-(i/\hbar)(Et)} e^{+(i/\hbar)(px)} = \psi e^{-(i/\hbar)(Et)} \quad (3.24)$$

where $\psi = Ae^{-(i/\hbar)(px)}$

Evidently Ψ is the product of a time-dependent function $e^{-(i/\hbar)(Et)}$ and a position-dependent function ψ . As it happens, the time variations of *all* wave functions of particles acted on by forces independent of time have the same form as that of an unrestricted particle. Substituting the Ψ of Eq. (3.24) into the time-dependent form of Schrödinger's equation, we find that

$$-\frac{\hbar^2}{2m} e^{-(i/\hbar)(Et)} \frac{\partial^2 \psi}{\partial x^2} + V \psi e^{-(i/\hbar)(Et)} = E \psi e^{-(i/\hbar)(Et)} \quad (3.25)$$

Dividing through by the common exponential factor gives

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{2m}{\hbar^2} (E - V) \psi = 0 \quad (3.26)$$

Equation (3.26) is the steady-state form of Schrödinger's equation. In three dimensions it is

$$\nabla^2\psi + \frac{2m}{\hbar^2}(E - V)\psi = 0 \quad (3.27)$$

An important property of Schrödinger's steady-state equation is that, if it has one or more solutions for a given system, each of these wave functions corresponds to a specific value of the energy E . Thus energy quantization appears in wave mechanics as a natural element of the theory, and energy quantization in the physical world is revealed as a universal phenomenon characteristic of all stable systems.

3.7 Particle in a box

How boundary conditions and normalization determine wave functions.

To solve Schrodinger's equation, even in its simpler steady-state form, usually requires elaborate mathematical techniques. For this reason the study of quantum mechanics has traditionally been reserved for advanced students who have the required proficiency in mathematics. However, since quantum mechanics is the theoretical structure whose results are closest to experimental reality, we must explore its methods and applications to understand modern physics. As we shall see, even a modest mathematical background is enough for us to follow the trains of thought that have led quantum mechanics to its greatest achievements.

The simplest quantum-mechanical problem is that of a particle trapped in a box with infinitely hard walls. We may specify the particles motion by saying that it is restricted to traveling along the x -axis between $x = 0$ and $x = L$ by infinitely hard walls. A particle does not lose energy when it collides with such walls, so that its total energy stays constant. From a formal point of view the potential energy $V(x)$ of the particle is infinite on both sides of the box, while $V(x)$ is a constant—say 0 for convenience—on the inside (Fig. 3.1).

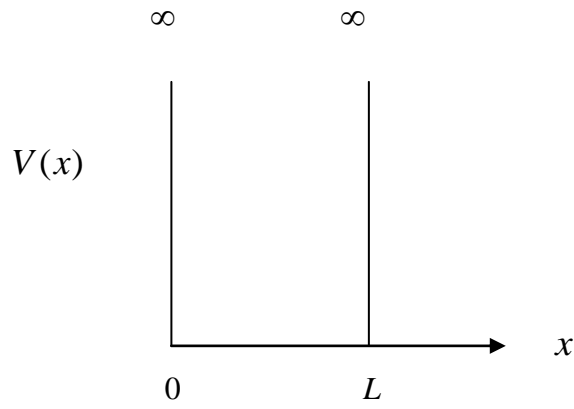


FIG. 3.1: A square potential well with infinitely high barriers at each end corresponds to a box with infinitely hard walls.

Because the particle cannot have an infinite amount of energy, it cannot exist outside the box, and so its wave function Ψ is 0 for $x \leq 0$ and $x \geq L$. Our task is to find what Ψ is within the box, namely, between $x = 0$ and $x = L$.

Within the box Schrodinger's equation becomes

$$\frac{d^2\Psi}{dx^2} + \frac{2m}{\hbar^2} E\Psi = 0. \quad (3.19)$$

since $V(x)=0$ there. (The total derivative $d^2\Psi/dx^2$ is the same as the partial derivative $\partial^2\Psi/\partial x^2$ because Ψ is a function only of x in this problem.)

Equation (3.19) has the form $\frac{d^2 y}{dx^2} + k^2 y = 0$ and the solution of this kind of equations is given mathematically by $y = A \sin kx + B \cos kx$. In the same way, the solution of Eq.(3.19) can be given by

$$\Psi = A \sin \frac{\sqrt{2mE}}{\hbar} x + B \cos \frac{\sqrt{2mE}}{\hbar} x \quad (3.20)$$

where we have used $p^2 = k^2 \hbar^2 = 2mE$. Eq.(3.20) we can verify by substitution back into Eq.(3.19). A and B are constants to be evaluated. This solution is subject to the boundary conditions that $\Psi = 0$ for $x = 0$ and for $x = L$. Since $\cos 0 = 1$, the second term cannot describe the particle because it does not vanish at $x = 0$. Hence we

conclude that $B = 0$. Since $\sin 0 = 0$, the sine term always yields $\Psi = 0$ at $x = 0$, as required, but Ψ should be 0 at $x = L$ and this *only* when

$$k_n L = \frac{\sqrt{2mE}}{\hbar} L = n\pi \quad (3.21)$$

This result comes about because the sines of the angles $\pi, 2\pi, 3\pi$, are all 0.

From Eq. (3.21) it is clear that the energy of the particle can have only certain values, which are the eigenvalues mentioned in the previous section. These eigenvalues, constituting the energy levels of the system, are found by solving Eq. (3.21) for E_n , which gives

$$\text{Particle in a box } E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2, \quad n=1, 2, 3, \dots \quad (3.22)$$

If we set $L = a$, Equation (3.22) becomes

$$E_n = \frac{\pi^2 \hbar^2}{2ma^2} n^2, \quad (3.23)$$

The essential feature of discrete energy spectrum has appeared naturally from the formalism. This may be compared with the Bohr formula for hydrogen

$$E_n^H = -\frac{\hbar^2}{2ma_0^2} \frac{1}{n^2},$$

Considering the crudeness of the model the similarity is striking. The discrepancy by the factor of factor of π is typical of one-dimensional approximations to three dimensional systems. The difference in sign and n^2 being in the denominator is due to the fact that for hydrogen the levels are measured from the top of the potential well downwards. For the square well, we measure from the bottom up.

CHAPTER 4

ONE-DIMENSIONAL MOTION

4.1 The Potential Step

Before considering the quantum theory of the harmonic oscillator and hydrogen atom, we compare the quantum and classical predictions in some simple one-dimensional cases

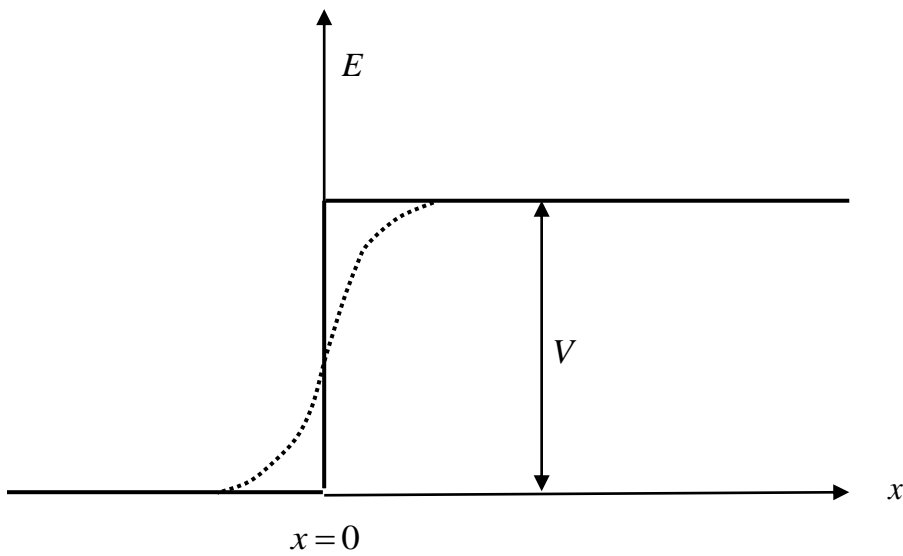


FIG 4.1. The energy diagram of a potential step. The dotted curve gives the realistic situation. The full line is an idealized situation, for which calculations are easier.

The simplest system is the motion of a particle in a potential of the form illustrated in Fig. 4.1 by the dotted curve. Since the force $F(x)$ is

$$F(x) = -\frac{\partial V}{\partial x} \quad (4.1)$$

this represents a particle which moves freely except in the neighborhood of the origin, where it is subjected to a force towards the left. If the particle has total energy E_0 , and kinetic energy T , then

$$E_0 = T(x) + V(x) \quad (4.2)$$

There are two cases (either $E_0 > V$ or $E_0 < V$) to be considered and we look at them both classically and quantum mechanically.

Case(i) $E_0 > V$ (Classical)

Particles coming from the left approach the potential barrier with kinetic energy T_0 , and momentum p_0 , given by

$$T_0 = E_0 = \frac{p_0^2}{2m}.$$

As the particles move through the region of the potential barrier, they are slowed up by the force, and kinetic energy is converted into potential energy. They have sufficient energy to penetrate the barrier and there is **total transmission**. The particles emerge to the right with kinetic energy, T_1 and momentum, p_1 where

$$T_1 = \frac{p_1^2}{2m} = E_0 - V. \quad (4.3)$$

Case (i) $E_0 > V$ (Quantel)

Schrödinger's equation is given by $-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V \Psi = E \Psi$.

For the left side of the origin ($V=0$), we rewrite Schrödinger's equation as

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} = E\Psi.$$

Multiply by $\frac{-2m}{\hbar^2}$ to get

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2mE}{\hbar^2} \Psi = 0.$$

Define $k_0^2 = \frac{2mE_0}{\hbar^2}$, to obtain

$$\left(\frac{\partial^2}{\partial x^2} + k_0^2 \right) \Psi_L(x) = 0, \quad x < 0 \quad (4.4)$$

the suffix L denoting the solutions to left of the origin. We are interested in the situation in which a particle approaches from the left, and may then be either transmitted or reflected. We thus look for a solution of the form

$$\Psi_L = e^{+ik_0x} + Ae^{-ik_0x}, \quad (4.5)$$

(incident) + (reflected).

We have arbitrarily normalized the coefficient of the incident wave to **unity**.

For the right side of the origin ($V \neq 0$), Schrödinger's equation becomes

$$\left(\frac{\partial^2}{\partial x^2} + k_1^2 \right) \Psi_R(x) = 0, \quad x > 0 \quad (4.6)$$

$$\text{where } k_1^2 = \frac{2m(E_0 - V)}{\hbar^2},$$

The suffix R denoting the solutions to right of the origin ($x > 0$).

The solution of Eq. (4.6) is given by

$$\Psi_R(x) = B e^{+ik_1 x}, \quad (4.7)$$

(transmitted).

We want to find the possible energy values, E_0 , of the system, and the reflected and transmitted intensities, determined through A and B , respectively. The continuity conditions at $x = 0$ are

$$1 + A = B \quad (\Psi \text{ continuous}), \quad (4.8)$$

$$k_0 - k_0 A = k_1 B \quad (\Psi' \text{ continuous}) \quad (4.9)$$

these can be solved for any value of E_0 , and imply

$$A = \frac{k_0 - k_1}{k_0 + k_1}, \quad B = \frac{2k_0}{k_0 + k_1} \quad (4.10)$$

The relative probability of finding the particle at a point x , for $x < 0$, is

$$P = |\Psi|^2 = |\Psi_L + \Psi_R|^2 = |e^{ik_0 x} + A e^{-ik_0 x}|^2 = 1 + |A|^2 + 2A \cos 2k_0 x. \quad (4.11)$$

The final oscillating term is not of much physical interest and can be removed by averaging over a range of length large compared with $2\pi/k_0$. The important new qualitative feature of the quantum theory is that, since

$$|A|^2 \neq 0 \quad (4.12)$$

there is a **non-vanishing reflected beam**.

CHAPTER 5

THE HARMONIC OSCILLATOR

5.1 Classical Theory

According to classical theory a harmonic oscillator is a particle, mass m , moving under the action of a force

$$F = -kx \quad \text{Hook's law} \quad (5.1)$$

where k is the force constant.

This force is given also by Newton's law

$$F = ma = m \frac{d^2 x}{dt^2} \quad \text{Newton's law} \quad (5.2)$$

Equating Eq.(1.5) with Eq.(5.2)

$$m \frac{d^2 x}{dt^2} = -kx \Rightarrow \frac{d^2 x}{dt^2} + \frac{k}{m} x = 0$$

let $\frac{k}{m} = \omega^2$, the equation of motion is then

$$\frac{d^2 x}{dt^2} + \omega^2 x = 0$$

with solution

$$x = A \cos \omega t \quad (5.3)$$

which represents an oscillatory motion of angular frequency ω , and amplitude A .

The potential is related to the force by

$$F = -\frac{\partial V}{\partial x}, \quad \Rightarrow \quad \int_0^x dV = -\int_0^x F dx = -\int_0^x -kx dx \quad \Rightarrow \quad V = \frac{1}{2} kx^2$$

so that

$$V(x) = \frac{1}{2} m \omega^2 x^2 \quad (5.4)$$

The shape of the potential energy is presented in FIG.5.1

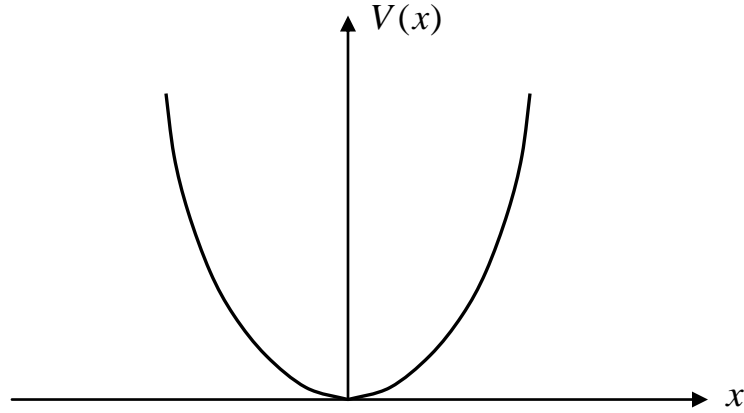


FIG.5.1 Potential energy of the simple harmonic oscillator.

5.2 Quantum Theory: The Eigenvalues

We now consider the quantum theory of such a system. Since the classical motion is bound for all values, the entire quantum energy spectrum should consist of discrete values. The energy eigenvalue equation is

$$\hat{H} u(x) = E u(x) \quad (\text{Schrödinger Equation, here } u(x) \equiv \Psi(x)) \quad (5.6)$$

with

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m \omega^2 x^2. \quad (5.7)$$

where \hat{H} is the **Hamiltonian operator** for harmonic oscillator.

Schrödinger Equation for harmonic oscillator is given by

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m \omega^2 x^2 \right] u_n(x) = E_n u_n(x). \quad (5.8)$$

If this is multiplied by $-2/\hbar\omega$ we get

$$\left[\frac{\hbar}{m\omega} \frac{\partial^2}{\partial x^2} - \frac{m\omega}{\hbar} x^2 \right] u_n(x) = -\frac{2E_n}{\hbar\omega} u_n(x). \quad (5.9)$$

Introducing the dimensionless variables

$$y = \left(\frac{m\omega}{\hbar} \right)^{1/2} x \quad (5.10)$$

$$\varepsilon = E_n / \hbar\omega$$

Schrödinger Equation becomes

$$\left(\frac{\partial^2}{\partial y^2} - y^2 \right) u_n(y) = -2\varepsilon_n u_n(y). \quad (5.11)$$

This equation may be solved by the standard techniques, which are employed below for angular momentum and the hydrogen atom. Instead, we use the **factorization method**, which is particularly elegant for this problem and brings to the fore a new type of operator, which in the long run plays a very important role in the theory. Since

$$\left(\frac{\partial}{\partial y} + y \right) \left(\frac{\partial}{\partial y} - y \right) u_n(y) = \left(\frac{\partial^2}{\partial y^2} - y^2 - 1 \right) u_n(y). \quad (5.12)$$

where $\left(\frac{\partial}{\partial y} - y \right)$ is the **raising ladder operator** and $\left(\frac{\partial}{\partial y} + y \right)$ is the **lowering ladder operator**.

Using Eq. (5.12), Eq. (5.11) may be re-written as follows:

$$\left(\frac{\partial}{\partial y} + y \right) \left(\frac{\partial}{\partial y} - y \right) u_n(y) = [-2\varepsilon_n - 1] u_n(y). \quad (5.13a)$$

Alternatively, it may be written as follows

$$\left(\frac{\partial}{\partial y} - y \right) \left(\frac{\partial}{\partial y} + y \right) u_n(y) = [-2\varepsilon_n + 1] u_n(y). \quad (5.13b)$$

Multiply (5.13a) by $\left(\frac{\partial}{\partial y} - y \right)$, then

$$\left(\frac{\partial}{\partial y} - y\right)\left(\frac{\partial}{\partial y} + y\right)\left(\frac{\partial}{\partial y} - y\right)u_n(y) = [-2\varepsilon_n - 1]\left(\frac{\partial}{\partial y} - y\right)u_n(y). \quad (5.14)$$

Then, either

$$\left(\frac{\partial}{\partial y} - y\right)u_n(y) = 0, \quad (5.15)$$

or

$$\left(\frac{\partial}{\partial y} - y\right)u_n(y) = u_{n+1}(y), \text{ say,} \quad (5.16)$$

and (5.14) can be written

$$\left(\frac{\partial}{\partial y} - y\right)\left(\frac{\partial}{\partial y} + y\right)u_{n+1}(y) = [-2(\varepsilon_n + 1) + 1]u_{n+1}(y). \quad (5.17)$$

This is (5.13b) for u_{n+1} , provided

$$\varepsilon_n + 1 = \varepsilon_{n+1} \quad (5.18)$$

The only solution to Eq. (5.15) is

$$u(y) = e^{y^2/2}$$

This diverges for large y and is therefore not a possible state. Thus given any solution $u_n(y)$, eigenvalue ε_n , it is always possible to generate a new state $u_{n+1}(y)$ by (5.16) with eigenvalue $\varepsilon_n + 1$.

Similarly, multiplying (5.13b) by $\left(\frac{\partial}{\partial y} + y\right)$,

$$\left(\frac{\partial}{\partial y} + y\right)\left(\frac{\partial}{\partial y} - y\right)\left(\frac{\partial}{\partial y} + y\right)u_n(y) = [-2\varepsilon_n + 1]\left(\frac{\partial}{\partial y} + y\right)u_n(y). \quad (5.19)$$

Now either

$$\left(\frac{\partial}{\partial y} + y\right) u_n(y) = 0, \quad (5.20)$$

or

$$\left(\frac{\partial}{\partial y} + y\right) u_n(y) = u_{n-1}(y), \text{ say,} \quad (5.21)$$

In the latter case (5.19) can be written

$$\left(\frac{\partial}{\partial y} + y\right) \left(\frac{\partial}{\partial y} - y\right) u_{n-1}(y) = [-2(\varepsilon_n - 1) - 1] u_{n-1}(y).$$

which is (5.13a) for u_{n-1} provided

$$\varepsilon_n - 1 = \varepsilon_{n-1} \quad (5.22)$$

Thus given any solution u_n , eigenvalue ε_n , it is possible to generate a new state of lower energy, $u_{n-1}(y)$ determined by (5.21) with eigenvalue $\varepsilon_n - 1$ unless u_n is the ground state, u_0 .

For $n = 0$, the solution to (5.15) is as follows

$$\begin{aligned} \frac{du_0}{dy} = -yu_0 &\Rightarrow \int_0^{u_0} \frac{du_0}{u_0} = -\int_0^y y dy \Rightarrow \ln u_0 = -\frac{y^2}{2}, \\ &\Rightarrow u_0 = e^{-y^2/2} \end{aligned} \quad (5.23)$$

Further, by (5.21) and (5.13b), the ground state energy is

$$2\varepsilon_0 - 1 = 0, \quad (5.24)$$

Combining these results, (5.20) and (5.13b), the eigenvalues are

$$\varepsilon_0 = \frac{1}{2}, \quad \varepsilon_1 = 1 + \frac{1}{2}, \quad \varepsilon_2 = 2 + \frac{1}{2}, \quad \dots \quad \varepsilon_n = n + \frac{1}{2}$$

but $\varepsilon = E_n / \hbar\omega$,

$$\therefore E_0 = \frac{1}{2}\hbar\omega, \quad E_1 = (1 + \frac{1}{2})\hbar\omega, \quad E_2 = (2 + \frac{1}{2})\hbar\omega, \quad \dots$$

$$\text{or } \boxed{E_n = (n + \frac{1}{2}) \hbar \omega,} \quad n = 0, 1, 2, \dots \quad (5.25)$$

a discrete set for all energies, in accordance with the general argument given above.

The expression for the energy levels is one of the most important in quantum mechanics. It justifies Planck's explanation of the interaction of radiation with matter, provided matter can be regarded as a collection of oscillators, each one emitting or absorbing radiation of its own frequency. The energy exchange is then restricted by the oscillator eigenvalues to take place in units of $\hbar \omega$, which is just Planck's hypothesis.

5.3 The raising and lowering ladder operator

The successive eigenfunctions can be generated from $u_0(x)$ by repeated application of (5.16), so that, for example

$$\begin{aligned} u_1(y) &= a_+ u_0(y) = \left(\frac{\partial}{\partial y} - y \right) u_0(y) = \left(\frac{\partial}{\partial y} - y \right) e^{-y^2/2} = \frac{\partial}{\partial y} e^{-y^2/2} - y e^{-y^2/2} \\ &= -2y e^{-y^2/2} = -2y u_0(x). \end{aligned} \quad (5.26)$$

This kind of polynomials is called Hermite polynomials. Hermite polynomials is

given by $H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2}$. We list a few of the Hermite polynomials here:

$$H_0(y) = 1$$

$$H_1(y) = 2y$$

$$H_2(y) = 4y^2 - 2$$

$$H_3(y) = 8y^3 - 12y$$

$$H_4(y) = 16y^4 - 48y^2 + 12$$

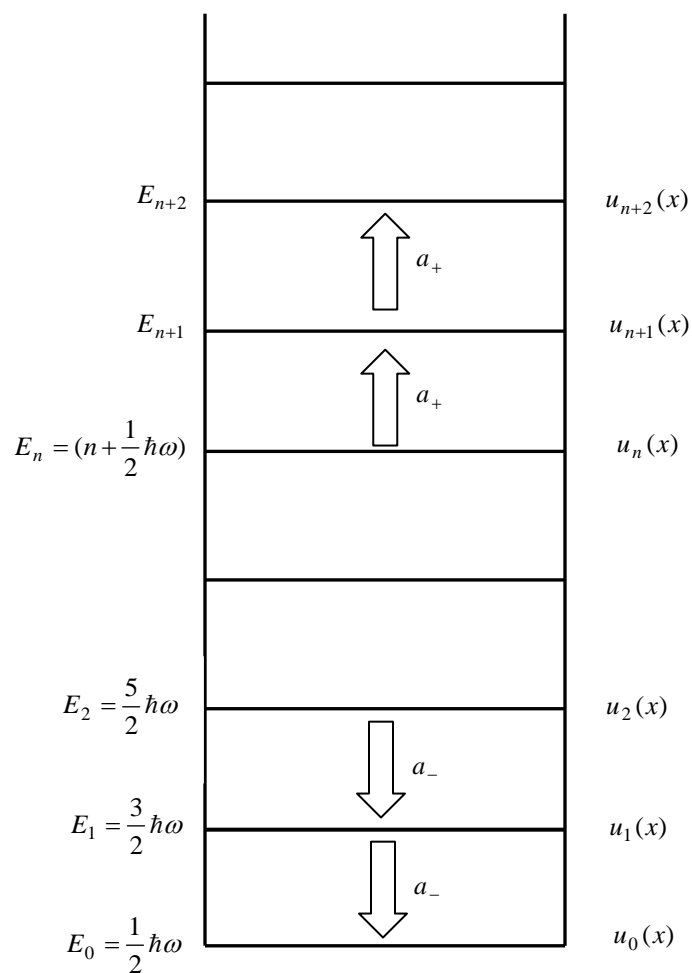


FIG.5.2 The ladder of stationary states for the simple harmonic oscillator.

CHAPTER 6

Quantum Theory of Hydrogen Atom

The first problem that Schrödinger tackled with his new wave equation was that of the hydrogen atom. He found the mathematics heavy going, but was rewarded by the discovery of how naturally quantization occurs in wave mechanics: "It has its basis in the requirement that a certain spatial function be finite and single-valued." In this chapter we shall see how Schrödinger's quantum theory of the hydrogen atom achieves its results, and how these results can be interpreted in terms of familiar concepts.

6.1 Schrödinger's Equation for Hydrogen Atom

A hydrogen atom consists of a proton, a particle of electric charge $+e$, and an electron, a particle of charge $-e$ which is 1836 times lighter than the proton. For the sake of convenience we shall consider the proton to be stationary, with the electron moving about in its vicinity but prevented from escaping by the protons electric field.

Schrödinger's equation for the electron in three dimensions, which is what we must use for the hydrogen atom, is

$$\nabla^2\Psi + \frac{2m}{\hbar^2}(E - V)\Psi = 0 \quad (6.1)$$

The potential energy V here is the electric potential energy

$$V = -\frac{e^2}{4\pi\epsilon_0 r} \quad (6.2)$$

of a charge $-e$ when it is the distance r from another charge $+e$.

Since V is a function of r rather than of x, y, z , we cannot substitute Eq. (6.2) directly into Eq. (6.1). There are two alternatives. One is to express V in terms of the cartesian coordinates x, y, z by replacing r by $\sqrt{x^2 + y^2 + z^2}$. The other is to

express Schrödinger's equation in terms of the spherical polar coordinates r , θ , ϕ defined in FIG. 6.1. Owing to the symmetry of the physical situation, doing the latter is appropriate here. The spherical polar coordinates r , θ , ϕ of the point P shown in Fig. 6.1 have the following interpretations:

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\theta = \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}} = \cos^{-1} \frac{z}{r}$$

$$\phi = \tan^{-1} \frac{y}{x}$$

$$\cos \theta = \frac{z}{r} \Rightarrow z = r \cos \theta$$

$$\cos \phi = \frac{x}{\rho} \Rightarrow x = \rho \cos \phi$$

$$\sin \phi = \frac{y}{\rho} \Rightarrow y = \rho \sin \phi$$

$$\rho = r \sin \theta$$

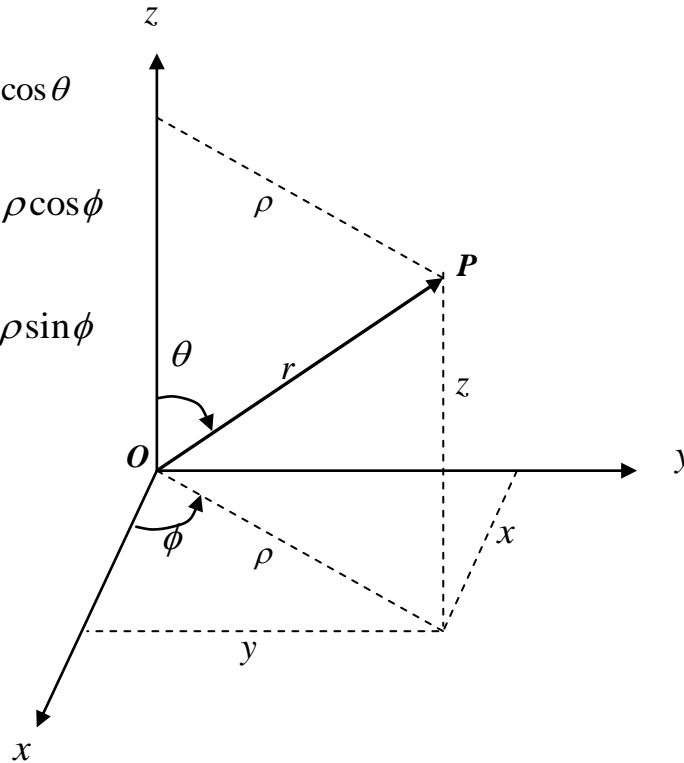


FIG 6.1 Spherical polar coordinates, $\theta \equiv$ zenith angle, $\phi \equiv$ azimuth angle,

$r = \sqrt{x^2 + y^2 + z^2}$ = length of radius vector from origin O to a point P .

In spherical polar coordinates Schrödinger's equation is written

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Psi}{\partial \phi^2} + \frac{2m}{\hbar^2} (E - V) \Psi = 0 \quad (6.3)$$

Substituting Eq. (6.2) for the potential energy V and multiplying the entire equation by $r^2 \sin^2 \theta$, we obtain

$$\sin^2 \theta \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Psi}{\partial \theta} \right) + \frac{\partial^2 \Psi}{\partial \phi^2} + \frac{2mr^2 \sin^2 \theta}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi = 0 \quad (6.4)$$

Equation (6.4) is the partial differential equation for the wave function Ψ of the electron in a hydrogen atom. Together with the various conditions Ψ must obey, namely that Ψ be normalizable and that Ψ and its derivatives be continuous and single-valued at each point r, θ, ϕ , this equation completely specifies the behavior of the electron. In order to see exactly what this behavior is, we must solve Eq. (6.4) for Ψ .

When Eq. (6.4) is solved, it turns out that three quantum numbers are required to describe the electron in a hydrogen atom, in place of the single quantum number of the Bohr theory. In the Bohr model, the electron's motion is basically one-dimensional, since the only quantity that varies as it moves is its position in a definite orbit. One quantum number is enough to specify the state of such an electron, just as one quantum number is enough to specify the state of a particle in a one-dimensional box.

A particle in a three-dimensional box needs three quantum numbers for its description, since there are now three sets of boundary conditions that the particle's wave function Ψ must obey: Ψ must be 0 at the walls of the box in the x, y , and z directions independently. In a hydrogen atom the electron's motion is restricted by the inverse-square electric field of the nucleus instead of by the walls of a box, but the electron is nevertheless free to move in

three dimensions, and it is accordingly not surprising that three quantum numbers govern its wave function also.

6.2 Separation of Variables

The advantage of writing Schrödinger's equation in spherical polar coordinates for the problem of the hydrogen atom is that in this form it may be separated into three independent equations, each involving only a single coordinate. Such a separation is possible here because the wave function $\Psi(r, \theta, \phi)$ has the form of a product of three different functions: $R(r)$, which depends on r alone; $\Theta(\theta)$ which depends on θ alone; and $\Phi(\phi)$, which depends on ϕ alone. Of course, we do not really know that this separation is possible yet, but we can proceed by assuming that

$$\Psi(r, \theta, \phi) = R(r) \Theta(\theta) \Phi(\phi) \quad (6.5)$$

and then seeing if it leads to the desired separation. The function $R(r)$ describes how the wave function Ψ of the electron varies along a radius vector from the nucleus, with θ and ϕ constant. The function $\Theta(\theta)$ describes how Ψ varies with zenith angle θ along a meridian on a sphere centered at the nucleus, with r and ϕ constant. The function $\Phi(\phi)$ describes how Ψ varies with azimuth angle ϕ along a parallel on a sphere centered at the nucleus, with r and θ constant. From Eq. (6.5), which we may write more simply as

$$\Psi(r, \theta, \phi) = R \Theta \Phi$$

we see that

$$\frac{\partial \Psi}{\partial r} = \Theta \Phi \frac{\partial R}{\partial r} = \Theta \Phi \frac{dR}{dr}, \quad \frac{\partial \Psi}{\partial \theta} = R \Phi \frac{\partial \Theta}{\partial \theta} = R \Phi \frac{d\Theta}{d\theta}, \quad \frac{\partial \Psi}{\partial \phi} = R \Theta \frac{\partial \Phi}{\partial \phi} = R \Theta \frac{d\Phi}{d\phi}$$

The change from partial derivatives to ordinary derivatives can be made because we are assuming that each of the functions R , Θ , and Φ depends only on the respective variables r , θ , and ϕ .

When we substitute $\Psi = R\Theta\Phi$ in Schrödinger's equation for the hydrogen atom and divide the entire equation by $R\Theta\Phi$, we find that

$$\frac{\sin^2 \theta}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{\Theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} + \frac{2mr^2 \sin^2 \theta}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi = 0 \quad (6.6)$$

The third term of Eq. (6.6) is a function of azimuth angle ϕ only, whereas the other terms are functions of r and θ only. Let us rearrange Eq. (6.6) to read

$$\frac{\sin^2 \theta}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{\Theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \frac{2mr^2 \sin^2 \theta}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi = -\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} \quad (6.7)$$

This equation can be correct only if both sides of it are equal to the same constant, since they are functions of *different* variables. As we shall see, it is convenient to call this constant m_l . The differential equation for the function ϕ is therefore

$$-\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} = m_l^2 \quad (6.8)$$

Next we substitute m_l^2 for the right-hand side of Eq. (6.7), divide the entire equation by $\sin^2 \theta$, and rearrange the various terms, which yields

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{2mr^2}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi = \frac{m_l^2}{\sin^2 \theta} - \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \quad (6.9)$$

Again we have an equation in which different variables appear on each side, requiring that both sides be equal to the same constant, this constant is called $l(l+1)$, once more for reasons that will be apparent later. The equations for the functions θ and R are therefore

$$\frac{m_l^2}{\sin^2 \theta} - \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = l(l+1) \quad (6.10)$$

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{2mr^2}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) = l(l+1) \quad (6.11)$$

Equations (6.8), (6.10), and (6.11) are usually written

$$\frac{d^2\Phi}{d\phi^2} + m_l^2\Phi = 0 \quad \text{Azimuth Equation} \quad (6.12)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \left[l(l+1) - \frac{m_l^2}{\sin^2 \theta} \right] \Theta = 0 \quad \text{Zenith Equation} \quad (6.13)$$

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \left[\frac{2m}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) - \frac{l(l+1)}{r^2} \right] R = 0 \quad \text{Radial Equation} \quad (6.14)$$

Each of these is an ordinary differential equation for a single function of a single variable. Only the equation for R depends on the potential energy $V(r)$.

6.3 Magnetic Quantum Numbers

The first of these equations, Eq. (6.12), is readily solved. The result is

$$\Phi(\phi) = Ae^{im_l\phi} \quad (6.15)$$

As we know, one of the conditions that a wave function—and hence Φ , which is a component of the complete wave function Ψ —must obey is that it have a single value at a given point in space. It is clear that Φ and $\Phi + 2\pi$ both identify the same meridian plane. Hence it must be true or $\Phi = \Phi + 2\pi$.

$$Ae^{im_l\phi} = Ae^{im_l(\phi+2\pi)}$$

which can happen only when m_l is 0 or a positive or negative integer ($\pm 1, \pm 2, \pm 3, \dots$)

The constant m_l is known as the **magnetic quantum number** of the hydrogen atom.

6.4 The Angular Equation (Zenith Equation)

We may rewrite Eq. (6.13) as

$$\sin\theta \frac{d}{d\theta} \left(\sin\theta \frac{d\Theta}{d\theta} \right) + [l(l+1) - m_l^2] \Theta = 0$$

The solution of this equation is given by

$$\Theta(\theta) = A P_l^m(\cos\theta), \quad (6.16)$$

$$\text{where } P_l^m(x) \equiv (1-x^2)^{|m|/2} \left(\frac{d}{dx} \right)^{|m|} P_l(x), \quad (6.17)$$

and $P_l^m(x)$ is the l th **Legendre** polynomial. We may define these polynomials by the **Rodrigues** formula

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx} \right)^l (x^2 - 1)^l, \quad (6.18)$$

$$\text{for } l=0, \quad P_0(x) = \frac{1}{2^0 0!} \left(\frac{d}{dx} \right)^0 (x^2 - 1)^0 = 1,$$

$$\text{for } l=1, \quad P_1(x) = \frac{1}{2^1 1!} \left(\frac{d}{dx} \right)^1 (x^2 - 1)^1 = x,$$

$$\text{for } l=2, \quad P_2(x) = \frac{1}{2^2 2!} \left(\frac{d}{dx} \right)^2 (x^2 - 1)^2 = \frac{1}{2} (3x^2 - 1).$$

HomeWork: Show that (a) $P_1^1(\cos\theta) = \sin\theta$, (b) $P_1^0(\cos\theta) = \cos\theta$,

(c) $P_2^2(\cos\theta) = 3\sin^2\theta$, (d) $P_2^1(\cos\theta) = 3\sin\theta\cos\theta$, (e) $P_2^0(\cos\theta) = (3\cos^2\theta - 1)$.

The normalized angular wave functions are called **spherical harmonics**:

$$Y_{lm}(\theta, \phi) = \varepsilon \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} e^{im\phi} P_l^m(\cos \theta) \quad (6.19)$$

where $\varepsilon = (-1)^m$ for $m \geq 0$ and $\varepsilon = 1$ for $m \leq 0$.

Example 6.1 Use Equations (6.17), (6.18), and (6.19) to construct Y_{00} , and Y_{21} .

Check they are normalized and orthogonal.

Solution

$$\text{Eq. (6.19)} \Rightarrow Y_{00} = \frac{1}{\sqrt{4\pi}} P_0^0(\cos \theta) ; \text{Eq. (6.18)} \Rightarrow P_0^0(x) = P_0(x) ;$$

$$\text{Eq. (6.17)} \Rightarrow P_0(x) = 1$$

$$\Rightarrow Y_{00} = \frac{1}{\sqrt{4\pi}}$$

$$Y_{21} = -\sqrt{\frac{5}{4\pi}} \frac{1}{3 \cdot 2} e^{i\phi} P_2^1(\cos \theta) ; P_2^1(x) = \sqrt{1-x^2} \frac{d}{dx} P_2(x) ;$$

$$P_2(x) = \frac{1}{2^2 2!} \left(\frac{d}{dx} \right)^2 (x^2 - 1)^2 = \frac{1}{8} \frac{d}{dx} [2(x^2 - 1)2x] = \frac{1}{2} [x^2 - 1 + x(2x)] = \frac{1}{2} (3x^2 - 1).$$

$$P_2^1(x) = \sqrt{1-x^2} \frac{d}{dx} \left[\frac{3}{2} x^2 - \frac{1}{2} \right] = \sqrt{1-x^2} 3x ; P_2^1(\cos \theta) = 3 \cos \theta \sin \theta .$$

$$\Rightarrow Y_{21} = -\sqrt{\frac{15}{8\pi}} e^{i\phi} \sin \theta \cos \theta$$

$$\text{Normalization test } \iint |Y_{00}|^2 \sin \theta d\theta d\phi = \frac{1}{4\pi} \left[\int_0^\pi \sin \theta d\theta \right] \left[\int_0^{2\pi} d\phi \right] = \frac{1}{4\pi} (2)(2\pi) = 1.$$

$\therefore Y_{00}$ is normalized

$$\text{Normalization test } \iint |Y_{21}|^2 \sin \theta d\theta d\phi = \frac{15}{8\pi} \int_0^\pi \sin^2 \theta \cos^2 \theta \sin \theta d\theta \int_0^{2\pi} d\phi$$

$$= \frac{15}{4} \int_0^\pi \cos^2 \theta (1 - \cos^2 \theta) \sin \theta d\theta = \frac{15}{4} \left[-\frac{\cos^3}{3} + \frac{\cos^5}{5} \right]_0^\pi = \frac{15}{4} \left[\frac{2}{3} - \frac{2}{5} \right] = \frac{5}{2} - \frac{3}{2} = 1$$

$\therefore Y_{21}$ is normalized

Orthogonality

$$\iint Y_{00}^* Y_{21} \sin \theta d\theta d\phi = \frac{1}{\sqrt{4\pi}} \sqrt{\frac{15}{8\pi}} \left[\int_0^\pi \sin \theta \cos \theta \sin \theta d\theta \right] \left[\int_0^{2\pi} e^{i\phi} d\phi \right] = \left[\frac{\sin^3 \theta}{3} \right]_0^\pi \left[\frac{e^{i\phi}}{i} \right]_0^{2\pi} = 0.$$

$\therefore Y_{00}$ and Y_{21} are orthogonal to each other.

Some spherical harmonics are presented in FIG. 6.2, The graphs represent the associated Legendre functions, $P_l^m(\cos \theta)$. In these plots r tell you the magnitude of the function in the direction of θ .

In general the normalized hydrogen functions are

$$Y_{nlm} = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} e^{-r/na} \left(\frac{2r}{na}\right)^l \left[L_{n-l-1}^{2l+1}(2r/na) \right] Y_{lm}(\theta, \phi)$$

$$\text{where } L_{q-p}^p(x) = (-1)^p \left(\frac{d}{dx}\right)^p L_q(x) \quad (6.20)$$

is the associated **Laguerre** polynomials

$$\text{and } L_q(x) = e^x \left(\frac{d}{dx}\right)^q (e^{-x} x^q) \quad (6.21)$$

is the q th **Laguerre** polynomial.

$$Y_{lm}(\theta, \phi) = \Theta(\theta) \Phi(\phi)$$

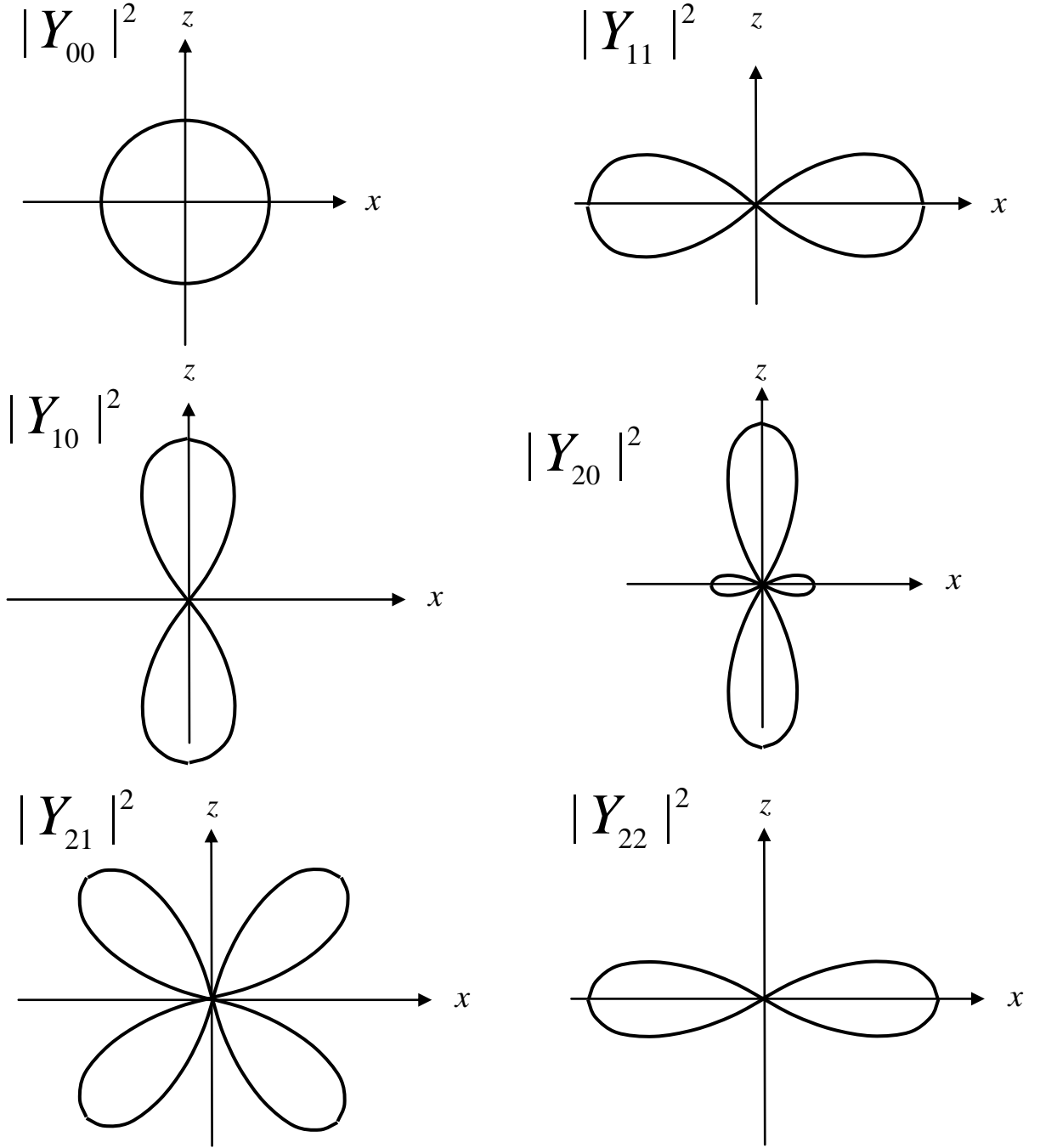


FIG.6.2 Distribution of $|Y(\theta, \phi)|^2$. The sketches represent sections of the distribution made in z - x plane. The three-dimensional distribution are obtained by rotating the figures about the z -axis.

Example 6.2 Find the ground state electron energy E_1 by substituting the radial wave function $R(r)$ that correspond to $n=1$, $l=0$ into Eq. (6.14).

Solution

From Table 6.2 we see that $R = \left(\frac{2}{a_0^{3/2}} \right) e^{-r/a_0}$. Hence $\frac{dR}{dr} = \left(-\frac{2}{a_0^{5/2}} \right) e^{-r/a_0}$

$$\begin{aligned} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) &= -\frac{2}{a_0^{5/2}} \left[\frac{1}{r^2} \frac{d}{dr} r^2 e^{-r/a_0} \right] = -\frac{2}{a_0^{5/2}} \frac{1}{r^2} \left[2r e^{-r/a_0} + r^2 \left(-\frac{1}{a_0} \right) \right] \\ &= \left(\frac{2}{a_0^{7/2}} - \frac{4}{ra_0^{5/2}} \right) e^{-r/a_0} \end{aligned}$$

Substituting in Eq. (6.14) with $E = E_1$ and $l=0$ gives

$$\left[\left(\frac{2}{a_0^{7/2}} + \frac{4mE_1}{\hbar^2 a_0^{3/2}} \right) + \left(\frac{me^2}{\pi\epsilon_0 \hbar a_0^{3/2}} - \frac{4}{a_0^{5/2}} \right) \frac{1}{r} \right] e^{-r/a} = 0$$

each parenthesis must equal to 0 for the entire equation to equal 0. For the second parenthesis this gives

$$\frac{me^2}{\pi\epsilon_0 \hbar a_0^{3/2}} - \frac{4}{a_0^{5/2}} = 0$$

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2} = \frac{\hbar^2}{me_M^2} \quad \text{where} \quad e_M^2 = \frac{e^2}{4\pi\epsilon_0}$$

which is the Bohr radius $a_0 = r_1$ given by Eq. (1.13)

The first parenthesis gives

$$\frac{2}{a_0^{7/2}} + \frac{4mE_1}{\hbar^2 a_0^{3/2}} = 0 \quad \Rightarrow$$

$$E_1 = -\frac{\hbar^2}{2ma_0^2} = -\frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} = \frac{1}{2} \cdot \frac{e^2}{4\pi\epsilon_0} \cdot \frac{me^2}{4\pi\epsilon_0\hbar^2}$$

$$= -\frac{e_M^2}{2a_0}$$

which agrees with Bohr's result [see Eq. (1.14)].

Probability of Finding the Electron

The probability density of the electron at the point r, θ, ϕ is proportional to $|\Psi|^2$, but the actual probability of finding it in the infinitesimal volume element dV there is $|\Psi|^2 dV$. In spherical polar coordinates (Fig. 6.4),

$$\text{Volume element } dV = r^2 \sin \theta dr d\theta d\phi \quad (6.25)$$

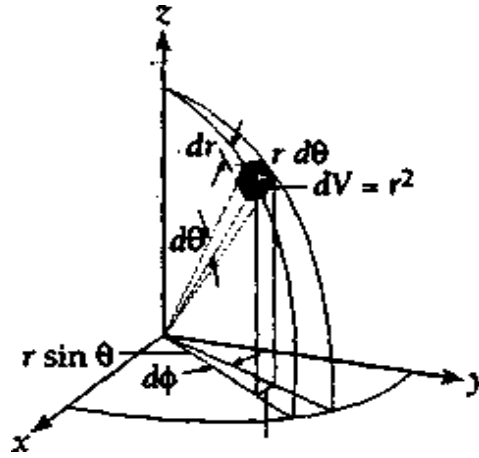


FIG 6.4 Volume element dV in spherical polar coordinates.

As Θ and Φ are normalized functions, the actual probability $P(r) dr$ of finding the electron in a hydrogen atom somewhere in the spherical shell between r and $r + dr$ from the nucleus is

$$P(r) dr = r^2 |R|^2 dr \int_0^\pi |\Theta|^2 \sin \theta d\theta \int_0^{2\pi} |\Phi|^2 d\phi,$$

$$= r^2 |R|^2 dr. \quad (6.26)$$

Equation (6.26) is plotted in Fig. 6.5 for the same states whose radial functions R were shown in Fig. 6.8. The curves are quite different as a rule. We note immediately that P is not a maximum at the nucleus for s states, as R itself is, but has its maximum a definite distance from it.

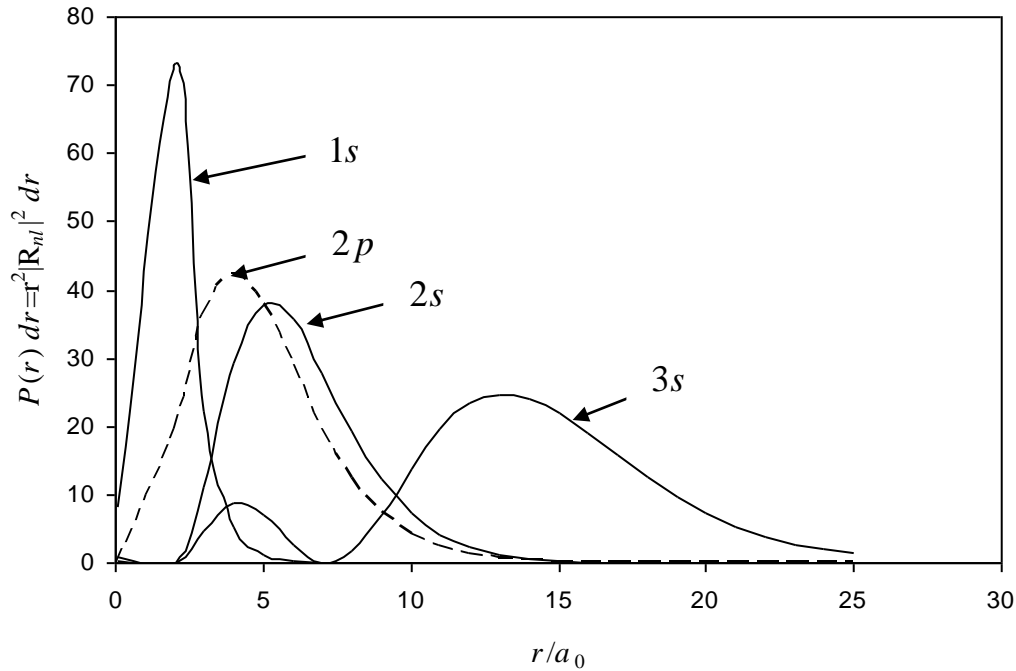


FIG 6.5 The probability of finding the electron in a hydrogen atom at a distance between r and $r + dr$ from the nucleus for the quantum states $1s, 2s, 2p, 3s$.

The most probable value of r for a $1s$ electron turns out to be exactly a_0 , the orbital radius of a ground-state electron in the Bohr model. However, the average value of r for a $1s$ electron is $1.5a_0$, which is puzzling at first sight because the energy levels are the same in both the quantum-mechanical and Bohr atomic models. This apparent discrepancy is removed when we recall that the electron energy depends upon $1/r$ rather than upon r directly, and the average value of $1/r$ for a $1s$ electron is exactly $1/a_0$.

Example 6.3 What is the most probable value of r , in the ground state of hydrogen atom?

Solution

First we must figure out the probability that the electron would be between r and $r + dr$.

$$\Psi_{100} = \frac{1}{\sqrt{\pi} a_0^{3/2}} e^{-r/a_0}$$

$$P = |\Psi_{100}|^2 4\pi r^2 dr = \frac{4}{a_0^3} e^{-2r/a_0} r^2 dr = p(r) dr$$

$$p(r) = \frac{4}{a_0^3} \left[2r e^{-2r/a_0} + r^2 \left(-\frac{2}{a_0} e^{-2r/a_0} \right) \right] = \frac{8r}{a_0^3} e^{-r/a_0} \left(1 - \frac{r}{a_0} \right) = 0 \Rightarrow \boxed{r = a_0}$$

Example 6.4

How much more likely is a 1s electron in a hydrogen atom to be at the distance a_0 from the nucleus than at the distance $a_0/2$?

Solution

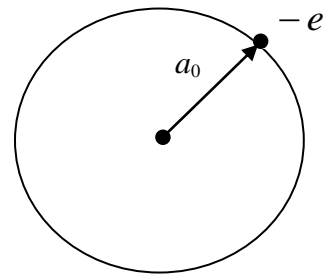
According to Table 6.2 the radial wave function for a 1s electron is $R = \frac{2}{a_0^{3/2}} e^{-r/a_0}$

From Eq. (6.25) we have for the ratio of the probabilities that an electron in a hydrogen atom be at the distances r_1 and r_2 from the nucleus

$$\frac{P_1}{P_2} = \frac{r_1^2 |R_1|^2}{r_2^2 |R_2|^2} = \frac{r_1^2 e^{-r_1/a_0}}{r_2^2 e^{-r_2/a_0}}$$

Here $r_1 = a_0$ and $r_2 = a_0/2$, so

$$\frac{P_{a_0}}{P_{a_0/2}} = \frac{a_0^2 e^{-2}}{(a_0/2)^2 e^{-1}} = 4 e^{-1} = 1.47$$



The electron is 47 percent more likely to be a_0 from the nucleus than half that distance (see FIG 6.5).

CHAPTER 7

Quantum Mechanics Formalism

7.1 The Hilbert Space and Wave Functions

7.1.1 The Linear Vector Space

A linear vector space consists of two sets of elements and two algebraic rules:

- a set of vectors ψ, ϕ, χ, \dots
- a rule for vector **addition** and a rule of scalar **multiplication**

(a) Addition rule

The addition rule has the following properties

1. If ψ and ϕ are vectors (elements) of a space, their sum, $\psi + \phi$, is also a vector of the same space.
2. Commutativity: $\psi + \phi = \phi + \psi$.
3. Associativity: $(\psi + \phi) + \chi = \psi + (\phi + \chi)$.
4. Existence of a zero or neutral vector: for each vector ψ must exist a zero vector (0) such that $0 + \psi = \psi + 0 = \psi$.
5. Existence of a symmetric or inverse vector: each vector ψ must have a symmetric vector ($-\psi$) such that $\psi + (-\psi) = (-\psi) + \psi = 0$.

(b) Multiplication rule

The multiplication of vectors by a scalar (scalars can be real or complex numbers) has these properties

1. The product of a scalar with a vector gives another vector. In general, if ψ and ϕ are two vectors of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.

2. Distributivity with respect to addition:

$$a(\psi + \phi) = a\psi + a\phi, \quad (a + b)\psi = a\psi + b\phi \quad (7.1)$$

3. Associativity with respect to multiplication of scalars:

$$a(b\psi) = (ab)\psi \quad (7.2)$$

4. For each element ψ there must exist a unitary element I and a zero scalar such that $I \cdot \psi = \psi \cdot I = \psi$ and $0 \cdot \psi = \psi \cdot 0 = 0$ (7.3)

7.1.2 The Hilbert Space

A Hilbert space \mathcal{H} consists of a set of vectors ψ, ϕ, χ, \dots and a set of scalars a, b, c, \dots which satisfy the following four properties:

(a) \mathcal{H} is a linear space

The properties of a linear space are considered in the previous section.

(b) \mathcal{H} has a defined a scalar product that is strictly positive

The scalar product of an element ψ with another element ϕ is a scalar, a complex number, denoted by (ψ, ϕ) where $(\psi, \phi) = \text{complex number}$. The scalar product satisfies the following properties:

1. The scalar product of ψ with ϕ is equal to complex conjugate of the scalar product of ϕ with ψ i.e:

$$(\psi, \phi) = (\phi, \psi)^* \quad (7.4)$$

Note : $(\psi, \phi) \neq (\phi, \psi)$, where $(\psi, \phi) = \int \psi^*(x) \phi(x) dx$, while $(\phi, \psi) = \int \phi^*(x) \psi(x) dx$.

2. The scalar product of ψ with ϕ is linear with respect to ϕ :

$$(\psi, a\phi_1 + b\phi_2) = a(\psi, \phi_1) + b(\psi, \phi_2) \quad (7.5)$$

The scalar product of a state vector ψ with itself is a positive real number:

$$(\psi, \psi) = \|\psi\|^2 \geq 0 \text{ where } (\psi, \psi) = \int \psi^*(x) \psi(x) dx \quad (7.6)$$

where the equality holds only for $\psi = 0$.

(c) \mathcal{H} is separable

There exists a Cauchy sequence $\psi_n \in \mathcal{H}$ ($n = 1, 2, \dots$) such that for every ψ of \mathcal{H} and $\varepsilon > 0$, there exists at least one ψ_n of the sequence for which

$$\|\psi - \psi_n\| < \varepsilon. \quad (7.7)$$

(d) \mathcal{H} is complete

Every Cauchy sequence of elements $\psi \in \mathcal{H}$ converges to an element of \mathcal{H} . That is, for any ψ_n the relation

$$\lim_{n, m \rightarrow \infty} \|\psi_n - \psi_m\| = 0, \quad (7.8)$$

defines a unique limit ψ of \mathcal{H} such that

$$\lim_{n, m \rightarrow \infty} \|\psi - \psi_n\| = 0, \quad (7.9)$$

7.1.3 Dimension and Basis of a Vector Space

A set of N vectors $\phi_1, \phi_2, \phi_3 \dots \phi_N$ is said to be **linearly independent** if and only if the solution of the equation

$$\sum_{i=1}^N a_i \phi_i = 0, \quad (7.10)$$

is $a_1 = a_2 = \dots = a_N = 0$. But if there exists a set of scalars, which are not all zero, so that one of the vectors can be expressed as a linear combination of the others,

$$\phi = \sum_{i=1}^N a_i \phi_i, \quad (7.11)$$

the set $\{\phi_i\}$ is said to be **linearly dependent**.

The **dimension** of a space vector is given by the **maximum number** of linearly independent vectors the space can have. For instance, if the maximum number of linearly independent

vectors a space has is N (i.e., $\phi_1, \phi_2, \dots, \phi_N$) this space is said to be N -dimensional. In this case, any vector ψ of the vector space can be expressed as a linear combination:

$$\psi = \sum_{i=1}^N a_i \phi_i, \quad (7.12)$$

The **basis** of a vector space consists of a set of the maximum possible number of linearly independent vectors belonging to that space. These vectors, $\phi_1, \phi_2, \dots, \phi_N$, to be denoted in short by $\{\phi_i\}$, are called the **base vectors**. Although the set of these linearly independent vectors is arbitrary, it is convenient to choose them *orthonormal*; that is, their scalar products satisfy the relation $(\phi_i, \phi_j) = \delta_{ij}$ (δ_{ij} is the Kronecker delta function defined as $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$). The basis is said to be *orthonormal* if it consists of a set of orthonormal vectors. Moreover, the basis is said to be *complete* if it spans the entire space; that is, there is no need to introduce any additional base vector. The expansion coefficients a_i in (7.12) are called the *components* of the vector ψ in the basis. Each component is given by the scalar product of ψ of corresponding base vector, $a_j(\phi_j, \psi)$.

Examples of linear vector spaces

Let us give two examples of linear spaces that are Hilbert spaces: one having a *finite (discrete)*, the other an *infinite (continuous)* basis.

- The first one is the three-dimensional **Euclidean** vector space; the basis of this space three linearly independent vectors, usually denoted by $\vec{i}, \vec{j}, \vec{k}$. Any vector of Euclidean space can be written in terms of the base vectors as $\vec{A} = a_1 \vec{i} + a_2 \vec{j} + a_3 \vec{k}$, where a_1, a_2 , and a_3 are the components of \vec{A} in the basis; each component can determined by taking the scalar product of \vec{A} with the corresponding base vector: $a_1 = \vec{i} \cdot \vec{A}$, $a_2 = \vec{j} \cdot \vec{A}$ and $a_3 = \vec{k} \cdot \vec{A}$. Note that the scalar product in the Euclidean is real and hence symmetric. The

norm in this space is the usual length of \vec{A} (i.e., $\|\vec{A}\| = A$). Note also that whenever $\vec{A} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k} = 0$, we have $a_1 = 0$, $a_2 = 0$ and $a_3 = 0$ and that none of the unit vectors \vec{i} , \vec{j} , \vec{k} can be expressed as a linear combination of the other two.

- The second example is the space of the entire complex functions $\psi(x)$; the dimension of this space is infinite for it has an infinite number of linearly independent basis vectors.

Example 7.1

Check whether the following sets of functions are linearly independent or dependent on the real x -axis

- (a) $f(x) = 4$, $g(x) = x^2$, $h(x) = e^{2x}$
- (b) $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$
- (c) $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$
- (d) $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^2$, $h(x) = 2x + 3x^2 - 8x^3$

Solution

(a) The first set is clearly linearly independent since $a_1f(x) + a_2g(x) + a_3h(x) = 0$
 $\Rightarrow 4a_1 + a_2x^2 + a_3e^{2x} = 0$ implies that $a_1 = a_2 = a_3 = 0$ for any value of x .

(b) The functions $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$ are also linearly independent since $a_1x + a_2x^2 + a_3x^3 = 0$ implies that $a_1 = a_2 = a_3 = 0$ no matter what the value of x . For instance taking $x = -1, 1, 3$ the following system of three equations

$$-a_1 + a_2 + a_3 = 0, \quad -a_1 + a_2 + a_3 = 0, \quad 2a_1 + 9a_2 + 27a_3 = 0 \quad (2.13)$$

yields $a_1 = a_2 = a_3 = 0$

(c) functions $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$ are not linearly independent, since $g(x) = 5f(x) + 0 \times h(x)$.

(d) functions $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^2$, $h(x) = 2x + 3x^2 - 8x^3$ are not

independent since $h(x) = 3f(x) - 2g(x)$

7.2 Dirac Notation

The physical state of a system is represented in quantum mechanics by elements of a linear space. These elements are called state vectors. We can represent the state by means of function expansions. This is analogous to specifying a Euclidean vector by its components in various coordinate systems. *The meaning of a vector is of the coordinate system chosen to represent its components.* Similarly, the state of a microscopic system has a meaning independent of the basis in which it is expanded.

To free state vectors from coordinate meaning, Dirac introduced what was to become an invaluable notation in quantum mechanics; he denoted the state vector ψ by what he called **ket** vector $|\psi\rangle$, its complex conjugate ψ^* by a **bra** $\langle\psi|$ and the scalar (inner) product $(\phi|\psi)$ by a **bra-ket** $\langle\phi|\psi\rangle$:

$$\psi \rightarrow |\psi\rangle \quad (7.14)$$

$$\psi^* \rightarrow \langle\psi| \quad (7.15)$$

$$(\phi, \psi) \rightarrow \langle\phi|\psi\rangle \quad (7.16)$$

when a ket is multiplied by a complex number, we get also a ket. Also when a bra is multiplied by a complex number, we get also a bra.

In quantum mechanics we deal with wave functions $\psi(\vec{r}, t)$ but in the more general formalisms we deal with ket vectors $|\psi\rangle$. Like wave functions, kets are elements of the Hilbert space. *For every ket there exists a unique bra and vice versa.* The bra vectors belong to a Hilbert space, \mathcal{H}^* which is said to be the *dual* space of the Hilbert space \mathcal{H} of the ket vectors.

Properties of kets, bras and bra-kets

1. Every ket has a corresponding bra

Every *ket* $|\psi\rangle$, corresponds to a *bra* $\langle\psi|$, which is given by:

$$(|\psi\rangle)^* = \langle\psi|, \quad \text{or} \quad (a|\psi\rangle)^* = a^* \langle\psi| \quad (7.17)$$

where a is a complex number. There is a one-to-one correspondence between bras and kets.

We can also write

$$|a\psi\rangle = a|\psi\rangle, \quad \text{or} \quad \langle a\psi| = a^* \langle\psi| \quad (7.18)$$

2. Properties of the scalar product

Since in quantum mechanics the scalar product is a complex number, the ordering matters a lot. We must be careful to distinguish a scalar product from its complex conjugate $\langle\phi|\psi\rangle$ is not the same thing as $\langle\psi|\phi\rangle$

$$\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle \quad (7.19)$$

This property becomes clearer if we apply it to (2.20):

$$\langle\phi|\psi\rangle^* = \left(\int \phi^*(\vec{r}, t) \psi(\vec{r}, t) \right)^* = \int \psi^*(\vec{r}, t) \phi(\vec{r}, t) = \langle\psi|\phi\rangle \quad (7.20)$$

The only case where $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle$ occurs is when $|\psi\rangle$ and $|\phi\rangle$ are real. Let us mention some additional properties of the scalar product:

$$\langle\psi|a_1\psi_1 + a_2\psi_2\rangle = a_1\langle\psi|\psi_1\rangle + a_2\langle\psi|\psi_2\rangle \quad (7.21)$$

$$\langle a_1\psi_1 + a_2\psi_2|\psi\rangle = a_1^*\langle\psi_1|\psi\rangle + a_2^*\langle\psi_2|\psi\rangle \quad (7.22)$$

• The norm is real and positive

For any state vector $|\psi\rangle$ of the Hilbert space H , the norm $\langle\psi|\psi\rangle$ is real and positive;

$\langle\psi|\psi\rangle$ to zero only for the case $|\psi\rangle = 0$. If the state $|\psi\rangle$ is normalized then $\langle\psi|\psi\rangle = 1$.

- **Schwarz inequality**

If any two states $|\psi\rangle$ and $|\phi\rangle$ the Hilbert space, we can show that

$$|\langle\psi|\phi\rangle|^2 \leq \langle\psi|\psi\rangle \langle\phi|\phi\rangle \quad (7.23)$$

If $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent (i.e., proportional: $|\psi\rangle = a|\phi\rangle$ where a is a scalar, this relation becomes an equality. This inequality is analogous to the following relation of the real Euclidean space

$$|\vec{A} \cdot \vec{B}|^2 \leq |\vec{A}|^2 |\vec{B}|^2. \quad (7.24)$$

- **Triangle inequality**

$$\sqrt{\langle\psi+\phi|\psi+\phi\rangle} \leq \sqrt{\langle\psi|\psi\rangle} + \sqrt{\langle\phi|\phi\rangle} \quad (7.25)$$

If $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent, $|\psi\rangle = a|\phi\rangle$ and if the proportionality scalar a is real positive, the triangle inequality becomes an equality. The counterpart of this inequality in Euclidean space is given by $|\vec{A} + \vec{B}| \leq |\vec{A}| + |\vec{B}|$.

- **Orthonormal states**

Two kets, $|\psi\rangle$ and $|\phi\rangle$, are said to be orthogonal if they have a vanishing scalar product:

$$\langle\psi|\phi\rangle = 0 \quad (7.26)$$

- **Forbidden quantities**

If $|\psi\rangle$ and $|\phi\rangle$, are belong to the same vector space, products of the type $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are forbidden. They are nonsensical, since $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are neither kets nor bras.

- **Physical meaning of the scalar product**

The scalar product can be interpreted in two ways First, by analogy with the scalar product vectors in the Euclidean space, where $\vec{A} \cdot \vec{B}$ represents the projection \vec{B} on \vec{A} , the product $\langle\phi|\psi\rangle$ also represents the projection of $|\psi\rangle$ onto $|\phi\rangle$. Second, in the case of normalized states and according to Born's probabilistic interpretation, the quantity $\langle\phi|\psi\rangle$ represents the

probability amplitude that the system's state $|\psi\rangle$ will, after performing a measurement on the system, be found to be in another state $|\phi\rangle$.

Example 7.2 (Bra-ket Algebra)

Consider the states $|\psi\rangle = 3i|\phi_1\rangle - 7i|\phi_2\rangle$ and $|\chi\rangle = -|\phi_1\rangle + 2i|\phi_2\rangle$ where $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthonormal.

- (a) Calculate $|\psi + \chi\rangle$ and $\langle\psi + \chi|$
- (b) Calculate the scalar products $\langle\psi|\chi\rangle$ and $\langle\chi|\psi\rangle$. Are they equal?
- (c) Show that the states $|\psi\rangle$ and $|\chi\rangle$ satisfy the Schwarz inequality.
- (d) Show that the states $|\psi\rangle$ and $|\chi\rangle$ satisfy the triangle inequality.

Solution

(a) The calculation of $|\psi + \chi\rangle$ is straightforward

$$\begin{aligned} |\psi + \chi\rangle &= |\psi\rangle + |\chi\rangle = (3i|\phi_1\rangle - 7i|\phi_2\rangle) + (-|\phi_1\rangle + 2i|\phi_2\rangle) \\ &= (-1 + 3i)|\phi_1\rangle + 5i|\phi_2\rangle. \end{aligned} \quad (7.27)$$

This leads at once to the expression of $\langle\psi + \chi|$

$$\langle\psi + \chi| = (-1 + 3i)^*\langle\phi_1| + (-5i)^*\langle\phi_2| = (-1 - 3i)\langle\phi_1| + 5i\langle\phi_2|. \quad (7.28)$$

(b) Since $\langle\phi_1|\phi_1\rangle = \langle\phi_2|\phi_2\rangle = 1$, $\langle\phi_1|\phi_2\rangle = \langle\phi_2|\phi_1\rangle = 0$, and since the corresponding to the kets $|\psi\rangle = 3i|\phi_1\rangle - 7i|\phi_2\rangle$ and $|\chi\rangle = -|\phi_1\rangle + 2i|\phi_2\rangle$ are given

by $\langle\psi| = -3i\langle\phi_1| + 7i\langle\phi_2|$ and $\langle\chi| = -\langle\phi_1| + 2i\langle\phi_2|$, the scalar products are

$$\begin{aligned} \langle\psi|\chi\rangle &= (-3i\langle\phi_1| + 7i\langle\phi_2|)(-\langle\phi_1| + 2i\langle\phi_2|) \\ &= (-3i)(-1)\langle\phi_1|\phi_1\rangle + (7i)(2i)\langle\phi_2|\phi_2\rangle \\ &= -14 - 3i, \end{aligned}$$

$$\begin{aligned}
\langle \chi | \psi \rangle &= (-\langle \phi_1 - 2i\langle \phi_2 | \rangle (3i | \phi_1 \rangle - 7i | \phi_2 \rangle)) \\
&= (-1)(3i)\langle \phi_1 | \phi_1 \rangle + (-2i)(-7i)\langle \phi_2 | \phi_2 \rangle \\
&= -14 - 3i .
\end{aligned}$$

We see that $\langle \psi | \chi \rangle$ is equal to the complex conjugate of $\langle \chi | \psi \rangle$.

(c) Let us first calculate $\langle \psi | \psi \rangle$ and $\langle \chi | \chi \rangle$

$$\langle \psi | \psi \rangle = (3i\langle \phi_1 | + 7i\langle \phi_2 | \rangle (3i | \phi_1 \rangle - 7i | \phi_2 \rangle) = (-3i)(3i) + (7i)(-7i) = 58 ,$$

$$\langle \chi | \chi \rangle = (-\langle \phi_1 | - 2i\langle \phi_2 | \rangle (- | \phi_1 \rangle + 2i | \phi_2 \rangle) = (-1)(-1) + (-2i)(2i) = 5 .$$

Since $\langle \psi | \chi \rangle = -14 + 3i$ we have $|\langle \psi | \chi \rangle|^2 = 14^2 + 3^2 = 205$.

Combining the values $|\langle \psi | \chi \rangle|^2$, $\langle \psi | \psi \rangle$, $\langle \chi | \chi \rangle$, we see the Schwarz inequality (7.23) is satisfied:

$$205 < (58)(5) \Rightarrow |\langle \psi | \chi \rangle|^2 < \langle \psi | \psi \rangle \langle \chi | \chi \rangle .$$

(d) First, let us use (7.27) and (7.28) to calculate

$$\begin{aligned}
\langle \psi + \chi | \psi + \chi \rangle &= [(-1 - 3i)\langle \phi_1 | + 5i\langle \phi_2 | \rangle] [(-1 + 3i) | \phi_1 \rangle - 5i | \phi_2 \rangle] \\
&= (-1 - 3i)(-1 + 3i) + (5i)(-5i) \\
&= 35 .
\end{aligned}$$

Since $\langle \psi | \psi \rangle = 58$, and $\langle \chi | \chi \rangle = 5$, we infer that the triangle inequality (7.25) is satisfied:

$$\sqrt{35} < \sqrt{58} + \sqrt{5} \Rightarrow \sqrt{\langle \psi + \chi | \psi + \chi \rangle} < \sqrt{\langle \psi | \psi \rangle} + \sqrt{\langle \chi | \chi \rangle}$$

7.3 Linear Transformations and matrix representation

Suppose you take every vector (in three-space) and multiply it by 17(or 22, say,) or you rotate every vector by 39° (say) about the z-axis, or you reflect every vector in the x-y plane these are all examples **of linear transformations**. A linear transformation \hat{A} takes each vector in a vector space and "transforms" it into some other vector $\hat{A}|\psi\rangle=|\phi\rangle$, with the proviso that the operation is **linear**:

$$\hat{A} (a |\psi\rangle + b |\phi\rangle) = a (\hat{A} |\psi\rangle) + b (\hat{A} |\phi\rangle) \quad (7.29)$$

for any vectors $|\psi\rangle, |\phi\rangle$ and any scalars a, b .

If you know what a particular linear transformation does to a set **of basis** vectors, you can easily figure out what it does to *any* vector. For suppose that

$$\hat{A} |\psi_1\rangle = A_{11} |\psi_1\rangle + A_{21} |\psi_2\rangle + \cdots + A_{n1} |\psi_n\rangle,$$

$$\hat{A} |\psi_2\rangle = A_{12} |\psi_1\rangle + A_{22} |\psi_2\rangle + \cdots + A_{n2} |\psi_n\rangle,$$

...

$$\hat{A} |\psi_n\rangle = A_{1n} |\psi_1\rangle + A_{2n} |\psi_2\rangle + \cdots + A_{nn} |\psi_n\rangle,$$

or, more compactly,

$$\hat{A} |\psi_j\rangle = \sum_{i=1}^n A_{ij} |\psi_i\rangle \quad (7.30)$$

If $|\psi\rangle$ is an arbitrary vector:

$$|\psi\rangle = a_1 |\psi_1\rangle + a_2 |\psi_2\rangle + \cdots + a_n |\psi_n\rangle \quad (7.31)$$

Evidently \hat{A} takes a vector with components a_1, a_2, \dots, a_n into a vector with components

$$a'_i = \sum_{j=1}^n A_{ij} a_j \quad (7.32)$$

If the basis is orthonormal, it follows from Equation 7.30 that

$$A_{ij} = \langle \psi_i | \hat{A} | \psi_j \rangle \quad (7.33)$$

It is convenient to display these complex numbers in the form of a **matrix**:

$$\hat{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & . & & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \quad (7.34)$$

The study of linear transformations, then, reduces to the theory of matrices. The *sum* of two linear transformations ($\hat{A} + \hat{B}$) is defined in the natural way:

$$(\hat{A} + \hat{B}) | \psi \rangle = \hat{A} | \psi \rangle + \hat{B} | \psi \rangle \quad (7.35)$$

this matches the usual rule for adding matrices (you add their corresponding elements):

$$\hat{C} = \hat{A} + \hat{B} \quad \Leftrightarrow \quad C_{ij} = A_{ij} + B_{ij} \quad (7.36)$$

And now, some useful matrix terminology: The **transpose** of a matrix (which we shall write with a superscript T : i.e., we write A^T) is the same set of elements, but with rows and columns interchanged:

$$A^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & . & & \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \quad (7.37)$$

Notice that the **transpose** of a **column matrix** $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ is a **row matrix**:

$$a^T = (a_1 \quad a_2 \quad \cdots \quad a_n)$$

A square matrix is symmetric if it is equal to its transpose (reflection in the **main** diagonal — upper left to lower right — leaves it unchanged); it is antisymmetric if this operation reverses the sign:

$$\hat{A}^T = \hat{A} \Rightarrow \text{Symmetric matrix}, \quad \hat{A}^T = -\hat{A} \Rightarrow \text{Antisymmetric matrix} \quad (7.38)$$

To construct the (complex) **conjugate** of a matrix (which we denote, as usual, with an asterisk: A^*), you take the complex conjugate of every element:

$$\hat{A}^* = \begin{pmatrix} A_{11}^* & A_{12}^* & \cdots & A_{1n}^* \\ A_{21}^* & A_{22}^* & \cdots & A_{2n}^* \\ \vdots & \cdot & \cdot & \cdot \\ A_{n1}^* & A_{n2}^* & \cdots & A_{nn}^* \end{pmatrix} \quad (7.39)$$

A matrix is **real** if all its elements are real and **imaginary** if they are all imaginary:

$$\text{If } A^* = A \Rightarrow A \text{ is Real matrix ; } \hat{A}^* = -\hat{A} \Rightarrow \hat{A} \text{ is Imaginary matrix} \quad (7.40)$$

The **Hermitian conjugate** (or **adjoint**) of a matrix (indicated by a dagger: A^\dagger) is the transposed conjugate:

$$\hat{A}^\dagger = (\hat{A}^T)^* \quad (7.41)$$

A square matrix is **Hermitian** (or **self-adjoint**) if it is equal to its Hermitian conjugate; if Hermitian conjugation introduces a minus sign, the matrix is **skew Hermitian** (or **anti-Hermitian**):

$$\hat{A}^\dagger = \hat{A} \Rightarrow \text{Hermitian matrix}, \quad \hat{A}^\dagger = -\hat{A} \Rightarrow \text{Skew Hermitian matrix} \quad (7.42)$$

With this notation the inner product of two vectors (with respect to an orthonormal basis

$\langle \psi | \phi \rangle = (a_1^* b_1 + a_2^* b_2 + \cdots + a_n^* b_n)$, can be written very neatly in matrix form:

$$\langle \psi | \phi \rangle = a^\dagger b. \quad (7.43)$$

(Notice that each of the three operations discussed in this paragraph, if applied twice, returns you to the original matrix.)

Matrix multiplication is not, in general, commutative ($\hat{A}\hat{B} \neq \hat{B}\hat{A}$); the difference between the two orderings is called the **commutator**:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (7.44)$$

The transpose of a product is the product of the transposes *in reverse order*.

$$(\hat{A}\hat{B})^T = \hat{B}^T \hat{A}^T \quad (7.45)$$

and the same goes for Hermitian conjugates:

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger \quad (7.46)$$

Inverse and Unitary Operators

A matrix has an inverse only if it is square and its determinant is nonzero; a matrix that has an inverse is called a nonsingular and a matrix that has no inverse is called a singular matrix.

The elements A_{mn}^{-1} of inverse matrix is given by

$$A_{mn}^{-1} = \frac{\text{cofactors of } A_{mn}}{\det. \text{ of } A_{mn}} \quad \text{or} \quad A_{mn}^{-1} = \frac{B^T}{\det. \text{ of } A_{mn}}$$

where B is the matrix of cofactors (also called the minor)

Trace of an operator

The trace $Tr(\hat{A})$ of an operator is given, within an orthonormal basis $\{|\phi_n\rangle\}$, by the expression

$$Tr(\hat{A}) = \sum_n \langle \phi_n | \hat{A} | \phi_n \rangle = \sum_n A_{nn}$$

The trace of a matrix is equal to the sum of its diagonal elements

$$Tr(\hat{A}) = Tr \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = A_{11} + A_{22} + A_{33} + \cdots,$$

Properties of the trace

$$Tr(\hat{A}^\dagger) = (Tr(\hat{A}))^*,$$

$$Tr(\alpha\hat{A} + \beta\hat{B} + \gamma\hat{C} + \cdots) = \alpha Tr(\hat{A}) + \beta Tr(\hat{B}) + \gamma Tr(\hat{C}) + \cdots,$$

$$Tr(\hat{A}\hat{B}) = Tr(\hat{B}\hat{A}).$$

Example 7.3 Calculate inverse matrix of $\hat{A} = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & i & -2 \end{pmatrix}$

Solution

Since the determinant of \hat{A} is $\det(\hat{A}) = -4 + 16i$, we have $A^{-1} = B^T / (-4 + 16i)$, where the elements of the cofactor matrix $B_{nm} = (-1)^{n+m}$ times the determinant of the submatrix obtained from \hat{A} by removing the n^{th} row and the m^{th} column. In this way, we have

$$B_{11} = (-1)^{1+1} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = (-1)^2 \begin{vmatrix} 1 & 5 \\ -i & -2 \end{vmatrix} = -2 + 5i \quad (7.47)$$

$$B_{12} = (-1)^{1+2} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} = (-1)^3 \begin{vmatrix} 3 & 5 \\ 0 & -2 \end{vmatrix} = 6$$

$$B_{21} = (-1)^3 \begin{vmatrix} i & 0 \\ -i & -2 \end{vmatrix} = 2i, \quad B_{22} = (-1)^4 \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4$$

$$B_{23} = (-1)^5 \begin{vmatrix} 2 & i \\ 0 & -i \end{vmatrix} = 2i, \quad B_{31} = (-1)^4 \begin{vmatrix} i & 0 \\ 1 & 5 \end{vmatrix} = 5i$$

$$B_{32} = (-1)^5 \begin{vmatrix} 2 & 0 \\ 3 & -2 \end{vmatrix} = -10, \quad B_{33} = (-1)^6 \begin{vmatrix} 2 & i \\ 3 & 1 \end{vmatrix} = 2 - 3i$$

$$\text{hence } \hat{B} = \begin{pmatrix} -2 + 5i & 6 & -3i \\ 2i & -4 & 2i \\ 5i & -10 & 2 - 3i \end{pmatrix}$$

Taking the transpose of \hat{B} , we obtain

$$\begin{aligned}
 A^{-1} &= \frac{1}{-4+16i} B^T = \frac{-1-4i}{68} \begin{pmatrix} -2+5i & 2i & 5i \\ 6 & -4 & -10 \\ -3i & 2i & 2-3i \end{pmatrix} \\
 &= \frac{1}{68} \begin{pmatrix} 22+3i & 8-2i & 20-5i \\ -6-24i & 4+16i & 10+40i \\ -12+3i & 8-2i & -14-5i \end{pmatrix}.
 \end{aligned}$$